CALCULUS OF VARIATIONS
and
TENSOR CALCULUS

U. H. Gerlach

January 2008

Beta Edition

Chapter 1

FUNDAMENTAL IDEAS

Lecture 1

1.1 Multivariable Calculus as a Prelude to the Calculus of Variations.

Calculus of several variables deals with the behaviour of functions whose domain is spanned by a finite number of coordinates.

The utility of this calculus stems from the fact that it provides, among others, methods for finding the "critical" points of a given function, namely those points where it has an extremum:

$$\frac{\partial f}{\partial x^i}(x_0^1, x_0^2, \ldots, x_0^n) = 0$$

The critical point $\vec{x}_0 = (x_0^1, x_0^2, \ldots, x_0^n)$ is a point where the function has a maximum, a minimum or an inflection behaviour. Which type of these extremal behaviours the function has in the neighborhood of its critical point can, in general, be inferred from the properties of the (symmetric) second derivative matrix,

$$\left[\frac{\partial^2 f(\vec{x}_0)}{\partial x^i \partial x^j}\right] \quad \left.\begin{matrix} i \\ j \end{matrix}\right\} 1, \ldots n$$

at the critical point $\vec{x}_0$.

The *calculus of variations* extends these methods from a finite to an infinite dimensional setting.

The distinguishing feature is that each point in this infinite dimensional setting is a *function*. Furthermore, the task of finding the extremum of a function becomes the task of finding the extremum of a "functional".

A "functional" is a map whose domain is the infinite dimensional space, each of whose points is a function. A "critical" point in this space is an "optimal" function which maximizes, minimizes, or in general, extremizes the given functional.

1.2 Some Typical Problems in the Calculus of Variations.

Nearly all problems in mathematical engineering, physics and geometry are closely related in one form or another to the calculus of variations. It turns out that these problems can be expressed in terms of some sort of optimization principle which says that some functional must be maximized, minimized or, in general, extremized. Let us consider some typical examples of such problems.

Problem 1. *The Brachistochrone Problem.*

Given: A particle moves along a curve $y(x)$ without friction. The particle starts with zero velocity from point $(x_1, y_1) = (0,0)$ and advances, because of a vertical constant gravitational field, to the point (x_2, y_2).

Question: For what curve is the travel time a minimum? Let us formulate the problem more precisely. The travel time is given by the integral

$$\text{time} = \int_1^2 dt = \int_{(x_1 y_1)}^{(x_2,y_2)} \frac{d(\text{path length})}{\text{speed}} \equiv \int_1^2 \frac{ds}{v} \ .$$

For a typical curve $y(x)$ with slope $\frac{dy}{dx}$ the speed v and the path length differentials are obtained from the following considerations:

(a) *time independent* gravitational field *implies* that the *total energy* of a system *is conserved*, i.e. independent of time. Hence

$$\text{T.E.} \equiv \text{K.E.}+\text{P.E.} = \text{constant} \ ,$$

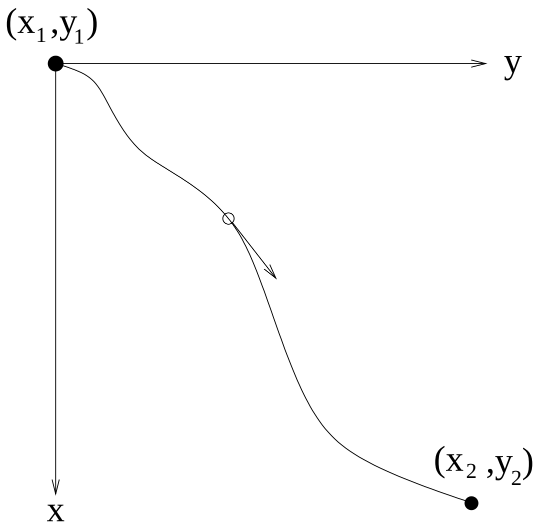


Figure 1.1: Particle sliding down a bent wire.

or

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

Here we have chosen the P.E. energy to be zero at $x = 0$:

$$\text{P.E.}(x = 0) = 0$$

Thus

$$v = \sqrt{2gx} \ .$$

Comment: The fact that the total energy is independent of time,

$$\text{T.E.=constant} \ ,$$

is an illustration of a general principle, the principle of energy conservation. It applies to all dynamical systems evolving in a time-invariant environment, e.g. a freely swinging pendulum or vibrating body. It does not apply to dynamical systems subjected to externally applied driving forces which depend on time.

(b) The element of path length is

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

The total travel time is therefore

$$t_{12} = \int_{(x_1 y_1)}^{(x_2,y_2)} \frac{\sqrt{1+\left(\frac{dy}{dx}\right)^2}}{\sqrt{2gx}}\, dx\ ,$$

which is a function of the function $y(x)$, i.e. t_{12} is a "functional".

Find a function $y(x)$ such that the "functional" t_{12}=minimum!

Problem 2. *Minimum Surface of Revolution.*

Given: Some curve between two fixed points (x_1, y_1) and (x_2, y_2) in the plane $z = 0$. A surface of revolution is formed by revolving the curve around the y-axis.

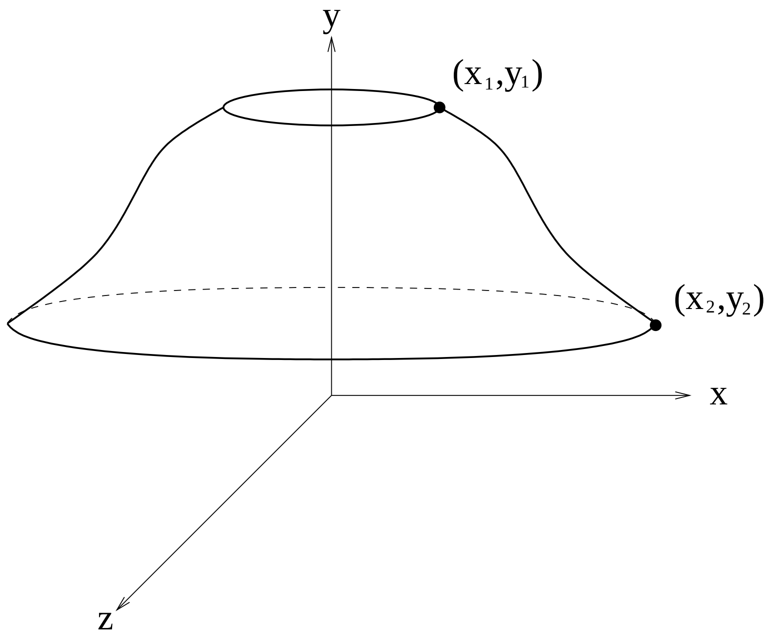


Figure 1.2: Surface area of revolution generated by a curve in the $x-y$ plane.

Question: For what curve $y(x)$ is the surface area a minimum?

One must find a function $y(x)$ such that the functional

$$\begin{aligned}\text{Area}[y] &= \int_1^2 2\pi x ds = 2\pi \int_{(x_1,y_1)}^{(x_2,y_2)} x\sqrt{1+\left(\frac{dy}{dx}\right)^2}dx \\ &= \text{minimum!}\end{aligned}$$

Problem 3. *Fermat's Principle of Least Time.*

Given: An inhomogeneous medium through which light (or sound) can propagate with a velocity that depends on position,

$$v(x,y) = \frac{c}{n(x,y)} \ .$$

Here $n(x,y)$ is the position dependent refractive index, which is assumed to be given.

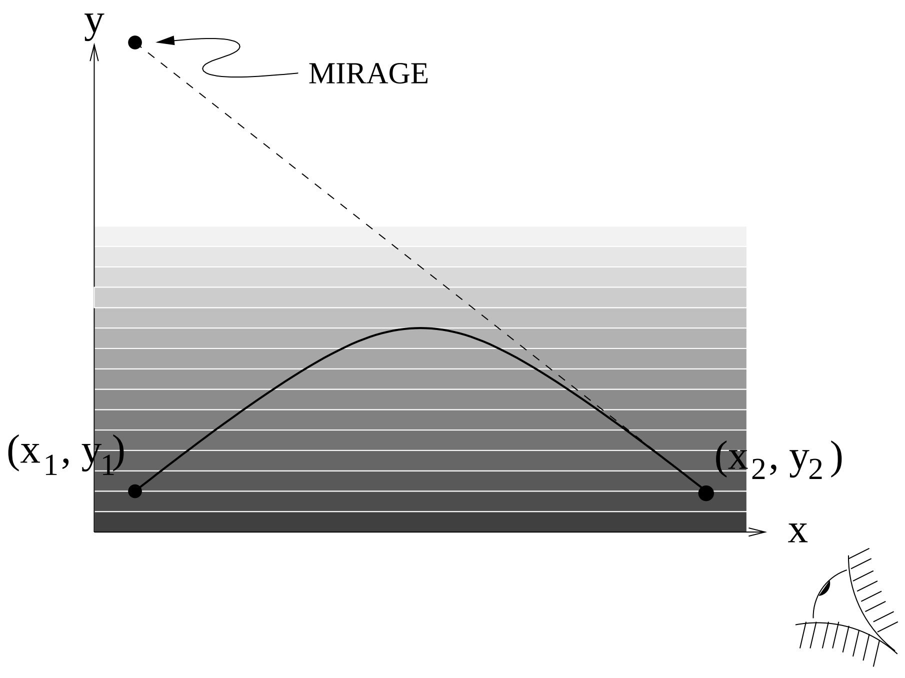


Figure 1.3: Optical path as determined by Fermat's principle of least time. In the lighter regions the propagation speed of light is faster. Thus the optimal (least time) path includes such regions in its domain of passage.

Statement of Fermat's Principle:
A light (or a radio, or a sound) beam takes that path $y(x)$ through two points (x_1, y_1) and (x_2, y_2) for which the travel time is a minimum.

This principle subsumes many natural phenomena, including 1.) mirages, 2.) ionospheric bending of radio beacons, 3.) twinkling of stars, 4.) sonar beams refracted by underwater temperature and salinity changes, 5.) light beam traversing an optical interface, etc.

The application of Fermat's principle consists of finding $y(x)$ such that

$$\begin{aligned} t_{12} &= \int_1^2 \frac{ds}{v(x,y)} \\ &= \int_{x_1}^{x_2} n(x,y)\frac{\sqrt{1+(y')^2}}{c}\,dx = \text{minimum!} \end{aligned}$$

1.3 Methods for Solving Problems in Calculus of Variations.

There are three known ways of finding an optimal function which minimizes (or maximizes) the given integral expression for the functional.

1. the method of finite differences,
2. the method of variations,
3. the direct method via minimizing sequences.

We shall briefly discuss the first, extensively develop the second, and postpone discussion of the third until later.

1.3.1 Method of Finite Differences.

In order to appreciate the basic meaning of problems in the calculus of variations and methods for solving them, it is important to see how they are related to the calculus of n variables. We are given a functional of the form

$$J[y] = \int_a^b F(x,y,y')dx \quad \text{with } y(a) = A \;\text{ and }\; y(b) = B.$$

This is a map which assigns a number to each curve $y(x)$ passing through the two points (a, A) and (b, B). One may find the curve which minimizes the integral J by the following approximation scheme:

1. Subdivide the interval $[a, b]$ into $n+1$ equal parts by using the points

$$x_0 = a,\ x_1, \ldots, x_n,\ x_{n+1} = b$$

2. Replace the curve $y(x)$ with the polygonal line whose vertices are

$$(x_0, A), (x_1, y(x_1)), \ldots, (x_n, y(x_n)), (x_{n+1}, B)$$

3. Approximate the integral $J[y]$ by the finite sum

$$J(y_1, \ldots, y_n) = \sum_{i=1}^{n+1} F\left(x_i, y_i, \frac{y_i - y_{i-1}}{\Delta x}\right)\ \Delta x$$

where

$$y_i = y(x_i) \text{ and } \Delta x = x_i - x_{i-1}$$

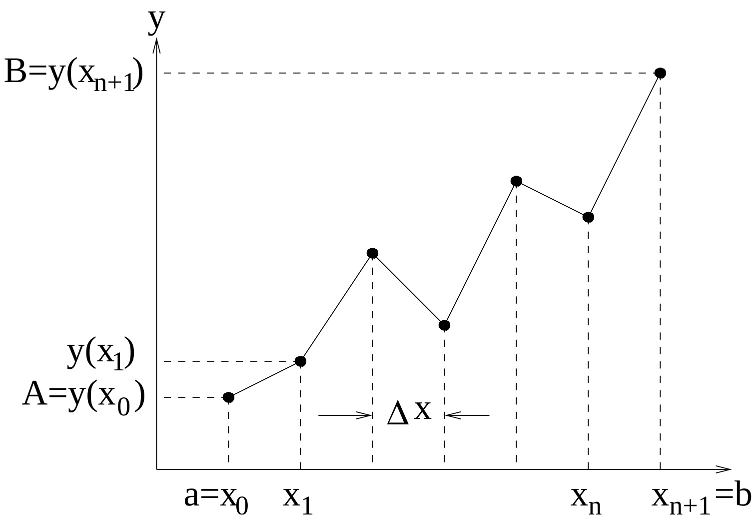


Figure 1.4: Curve approximated by a polygonal line.

Thus each polygonal line is uniquely determined by the ordinate y_1, $y_2, \ldots, y_n$ of the vertices because the end points are assumed to be *fixed*. Consequently, the *functional* $J[y]$ may be considered as *being approximated* by the function $J(y_1, \ldots, y_n)$, and the problem of finding the optimal curve $y(x)$ which minimizes $J[y]$ may be approximated by the problem of finding the critical point $(y_1^*, \ldots, y_n^*)$ which minimizes $J(y_1, \ldots, y_n)$.

This minimization process consists of solving the following n simultaneous equations

$$\frac{\partial J}{\partial y_j} = 0 \qquad j = 1, \ldots, n$$

Using the expression for J, namely

$$\begin{aligned} J(y_1, \ldots, y_n) &= F\left(x_1, y_1, \frac{y_1 - y_0}{\Delta x}\right) \Delta x + F\left(x_2, y_2, \frac{y_2 - y_1}{\Delta x}\right) \Delta x \\ \ldots &+ F\left(x_j, y_j, \frac{y_j - y_{j-1}}{\Delta x}\right) \Delta x + F\left(x_{j+1}, y_{j+1}, \frac{y_{j+1} - y_j}{\Delta x}\right) \Delta x \\ + \ldots &+ F\left(b, B, \frac{B - y_n}{\Delta x}\right) \Delta x, \end{aligned}$$

let us calculate the partial derivative w.r.t y_j:

$$\begin{aligned} \frac{\partial J}{\partial y_j} &= \frac{\partial}{\partial y_j} \{F(x_j, y_j, \underbrace{\frac{y_j - y_{j-1}}{\Delta x}}_{y'_{j-1}}) \Delta x \quad + \quad F(x_{j+1}, y_{j+1} \underbrace{\frac{y_{j+1} - y_j}{\Delta x}}_{y'_j}) \Delta x\} \\ &= F_{y_j}(x_j, y_j, \frac{y_j - y_{j-1}}{\Delta x}) \Delta x \quad + \quad F_{y'_j}(x_j, y_j, y'_{j-1}) \frac{1}{\Delta x} \Delta x \\ &\qquad + \quad F_{y'_j}(x_{j+1}, y_{j+1}, y'_j) \frac{-1}{\Delta x} \Delta x \end{aligned}$$

As $\Delta x \to 0$; x_j, $x_{j-1} \to x$; $y_j \to y$ the number of subintervals increases without limit and the expression approaches zero. But if we divide by Δx, then we have

$$\lim_{\Delta x \to 0} \frac{1}{\Delta x} \frac{\partial J}{\partial y_j} \equiv \frac{\delta J}{\delta y}$$

Remark: The expression $\Delta x \partial y_j$ has a direct geometrical significance as the area between the given curve and the varied (dashed) curve. The limit of the

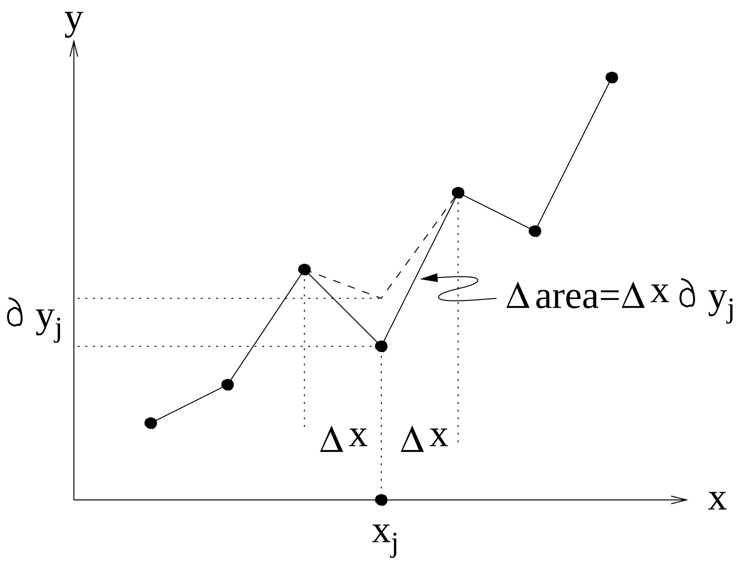


Figure 1.5: The area between the solid and the dashed curve is $\Delta x \partial y_j$.

differential quotient is the variational derivative of J,

$$\frac{\delta J}{\delta y} = \frac{\partial F(x, y, y')}{\partial y} - \frac{d}{dx} \frac{\partial F(x, y, y')}{\partial y'}$$

The process of minimizing the functional $J[y]$ by first approximating it as $J(y_1, \ldots, y_n)$, as function of the n independent variables $y_1, \ldots, y_n$ is called the *method of finite differences.*

Leonard Euler used it to solve problems in the calculus of variations. By replacing smooth curves with polygonal lines he solved such problems as multivariable problems in n-dimensions. By letting $n \to \infty$, he obtained the exact solution.

Thus the functional $J[y]$ may be considered a function of infinitely many variables, and the *calculus of variations may be considered an extension* of multivariable calculus *into a space of infinite dimensions.*

Lecture 2

1.4 The Method of Variations.

The method of finite differences first approximates an infinite dimensional problem by one of finite dimensions. Solving the finite dimensional problem consists of finding the critical point of a scalar function having a finite number independent variables. The coordinates of this critical point, say $\{y_i = y(x_i) : i = 1, \ldots, n\}$, are the vertices of a polygonal curve, the finite approximation to the curve that optimizes the scalar functional. One obtains the exact solution, the optimal curve, to the infinite dimensional problem by letting n, the dimension of the approximation space, become infinite. There exist numerical optimization techniques for finding the critical point scalar function on a finite dimensional approximation space. One obtains the exact solution to the infinite problem by letting n, the dimension of the approximation space, become infinite. But in that case the numerical optimization technique becomes computationally very expensive.

However, one can dispense with the intermediate approximation process entirely by dealing with the infinite dimensional problem directly. In this strategy one does not obtain the optimal curve directly. Instead, one obtains a differential equation which this curve must satisfy. However, finding solutions to a differential equation is a highly developed science. Hence dispensing with the intermediate approximation process is an enormous advantage if, as is usually the case, one can solve the differential equation.

It is difficult to overstate the importance of the method which allows us to do this. It is called the method of variations, which we shall now describe.

1.4.1 Variants and Variations

The basic problem of the calculus of variation is this:

Given: (a) $F(x; q, p)$ a function of three variables and (b) a pair of points (x_1y_1) and (x_2y_2).

Determine: That "optimal" function $y(x)$ which passes through (x_1y_1) and (x_2y_2) such that

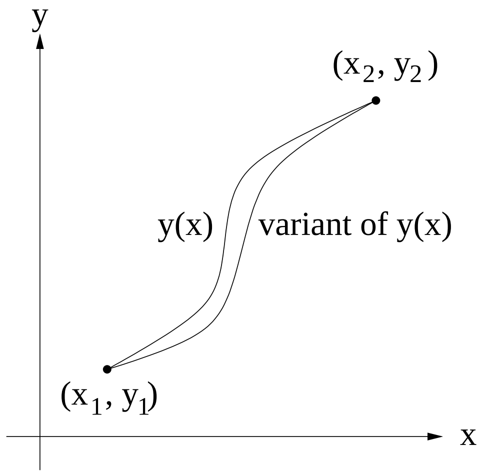
$$J[y] = \int_{x_1}^{x_2} F(x; y(x), y'(x))dx$$

has the largest (or smallest or, in general, extremal value for $y(x)$ as compared to any *variant* of $y(x)$ also passing through (x_1y_1) and (x_2y_2).

Remark 1: The dictionary definition of "variant" is as follows:

variant – something that differs in form only slightly from something else.

Remark 2: In order to determine whether $y(x)$, which passes through (x_1, y_1)



and (x_2, y_2), is an "optimal" curve, one must compare the functional J evaluated for all variants of $y(x)$, including $y(x)$ itself.
Remark 3: We shall *assume* that an optimal curve exists, and then deduce as a consequence the conditions that $y(x)$ must satisfy. In other words, we shall determine the *necessary conditions* for $y(x)$ to exist. We are *not* developing *sufficient conditions* for the existence of an optimal curve.

To test whether a curve is optimal by comparing values of J, pick a curve suspected of being an optimal one and deform it slightly. This yields a *variant*, a *deformed curve*, of the original curve, namely

$$\begin{aligned} y &= f(x, \alpha_i) \\ &\equiv f(x, 0) + \alpha_1 h_1(x) + \alpha_2 h_2(x) + \dots \\ &\equiv f(x, 0) + \delta f(x) \end{aligned}$$

Here

$$y = f(x, 0) \equiv f(x)$$

is the *optimal curve* under consideration, while

$$\delta f(x) = \alpha_1 h_1(x) + \alpha_2 h_2(x) + \dots$$

is the *variation* in the function f.

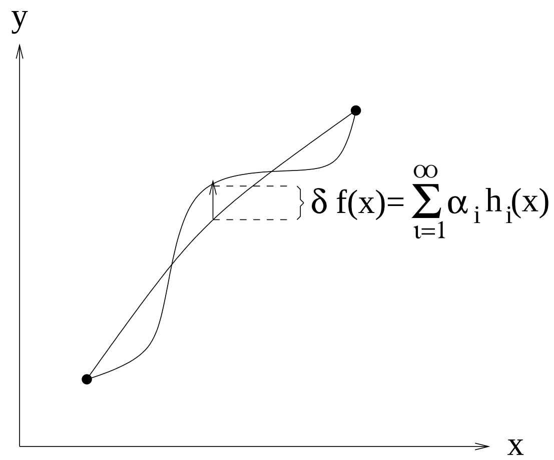


Figure 1.6: Arbitrary variation of f as a superposition of basis functions.

Definition. The fact that all variants of f pass through the same pair of endpoints we express by saying that they are *admissible* functions; in other words admissible functions have the property that their variations vanish at the endpoints,

$$\boxed{\delta f(x_1) = \delta f(x_2) = 0}.$$

Because of the linear independence of the h_i's, this is equivalent to

$$\boxed{h_i(x_1) = h_i(x_2) = 0 \quad i = 1, 2 \ldots}$$

Relative to the set of basis functions $\{h_i\}$ the variants

$$\begin{array}{ll}
A: & f(x,0) - .2h_1(x) \\
B: & f(x,0) - .1h_1(x) \\
C: & f(x,0) \\
D: & f(x,0) + .1h_1(x) \\
E: & f(x,0) + .2h_1(x) \\
F: & f(x,0) + .2h_1(x) + .1h_2(x) \\
G: & f(x,0) + .2h_1(x) + .2h_2(x)
\end{array}$$

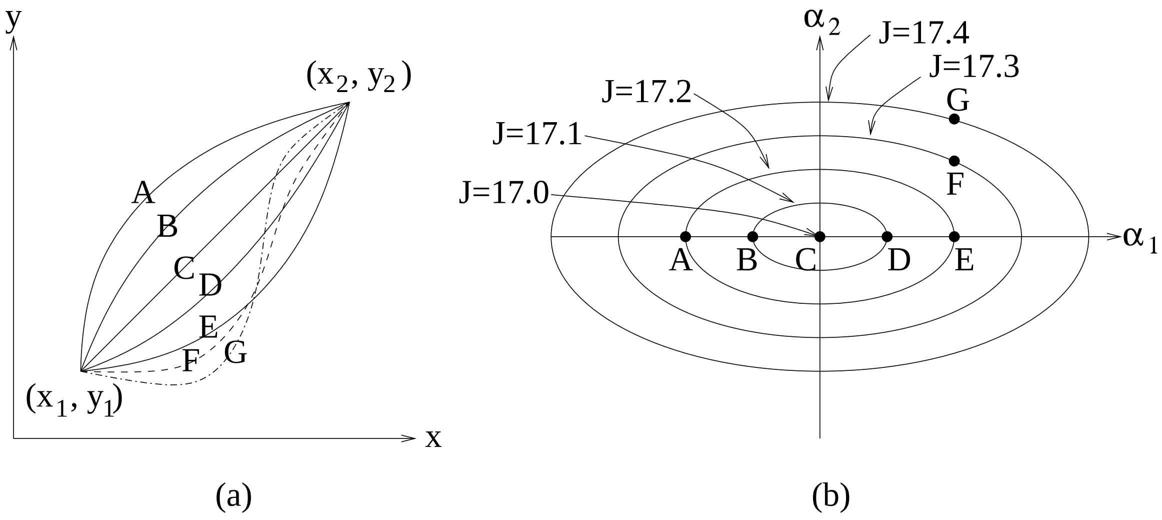


Figure 1.7: (a) Optimal curve (C) and some of its variants in x-y space. (b) Isograms of J on the function space of variants.

are represented as points

$$\begin{array}{ll} A: & (-.2, 0, \dots) \\ B: & (-.1, 0, \dots) \\ C: & (0, 0, \dots) \\ D: & (.1, 0, \dots) \\ E: & (.2, 0, \dots) \\ F: & (.2, .1, \dots) \\ G: & (.2, .2, \dots) \end{array}$$

in the function space *coordinatized* by α_1, α_2, etc. as depicted in Figure 1.7.

We now proceed to compare the values of the integral to be optimized. For all variants the value of J is obtained from

$$J[y] = \int_{x_1}^{x_2} F\left(x; y, \frac{dy}{dx}\right) dx.$$

The difference between a function and its variant is called the *variation* in the function.

a) For fixed x, the *variation* in the function f is

$$\delta f(x) = \sum \alpha_i h_i(x)$$

b) the variation in the derivative of f is

$$\begin{aligned} \delta\left(\frac{df}{dx}\right) &\equiv \frac{df(x,\alpha_i)}{dx} - \frac{df(x,0)}{dx} = \frac{d}{dx}\left(\delta f(x)\right) \\ &= \alpha_1 h_1'(x) + \alpha_2 h_2'(x) + \ldots \end{aligned}$$

c) the variation in the integrand F is

$$\begin{aligned} \delta F &\equiv F(x; f(x,\alpha_i), f'(x,\alpha_i)) - F(x; f(x,0), f'(x,0)) \\ &= \frac{\partial F}{\partial y}\,\delta f + \frac{\partial F}{\partial y'}\,\delta\left(\frac{df}{dx}\right) \\ &= \frac{\partial F}{\partial y}\,\delta f + \frac{\partial F}{\partial y'}\,\frac{d}{dx}(\delta f) \end{aligned}$$

d) Using these variations one determines the *variation* in the *integrand* with the result

$$\begin{aligned} \Delta J &\equiv J(\alpha_i) - J(0) \\ &= \int_{x_1}^{x_2} [F(x; f(x,\alpha_i), f'(x,\alpha_i)) - F(x; f(x,0), f'(x,0)]dx \qquad (1.1) \\ &= \int_{x_1}^{x_2} \left[F_y \delta f + F_{y'}\frac{d}{dx}(\delta f)\right] dx + \text{neglegible higher order terms in } f \end{aligned}$$

The Taylor series expansion theorem was used to obtain this result. It depends on the variation δf and its derivative. However, we would like to have the result depend only on δf, and not on its derivative.

This desire can be realized by integrating the second term by parts. One obtains

$$\Delta J = \int_{x_1}^{x_2} \left[F_y - \frac{d}{dx}(F_{y'})\right] \delta f(x)dx + F_{y'}\delta f(x)\Big|_{x_1}^{x_2} + \text{neglegible terms}$$

We now make use of the fact that all variants of f pass through the same end points, i.e. that all variations δf have the property

$$\delta f(x_1) = \delta f(x_2) = 0,$$

which we specified already on P 1.22. Consequently, the boundary contribution to ΔJ vanishes. We now consider the *principal linear part* of ΔJ, which is linear in δf, which we designate by δJ, and which is given by

$$\delta J = \int_{x_1}^{x_2} \left[F_y - \frac{d(F_{y'})}{dx} \right] \delta f(x) dx. \tag{1.2}$$

This is the *1st variation* of the functional J at $y = f(x,0)$. It is a linear functional which is also called the *differential* of J.

1.4.2 Variational Derivative

The fact that an arbitrary variant of $f(x,0)$ is given by

$$f(x,\alpha_i) = f(x,0) + \sum_{i=1}^{\infty} \alpha_i h_i(x)$$

implies that the variations of f,

$$\delta f(x) = \sum_{i=1}^{\infty} \alpha_i h_i(x) \tag{1.3}$$

form a linear vector space which is spanned by the basis vectors $\{h_i(x) : i = 1, 2, \dots\}$. Taking a cue from multivariable calculus, we can use the differential δJ of J, and identify the *directional derivatives* of J along the respective directions h_k in the space of admissible (see P 1.19) functions.

These directional derivatives, we recall, are give with the help of Eqs. (1.1), (1.2) and (1.3) by

$$\begin{aligned} \left.\frac{\partial J(\alpha_i)}{\partial \alpha_k}\right|_{\alpha_i=0} &= \lim_{\alpha_k \to 0} \frac{\Delta J(0,\dots,\alpha_k,\dots)}{\alpha_k} \\ &= \int_{x_1}^{x_2} \left[F_y - \frac{d}{dx}(F_{y'}) \right] h_k(x) dx \qquad (1.4) \\ & \quad k = 1, 2, \dots \end{aligned}$$

In order to establish the parallel with multivariable calculus, let us designate

$$F_y - \frac{d}{dx}\, F_{y'} \equiv \frac{\delta J}{\delta y(x)} \qquad x_1 < x < x_2 \tag{1.5}$$

as the (*first*) *variational derivative of* $J[y]$.

It is clear that this derivative depends on x.

Reminder:

Recall that in multivariable calculus the gradient $\vec{\nabla} f$ of a function is related to its *directional derivative* $D_{\vec{e}_k} f$ along the same basis vector e_k by the equation

$$D_{e_k} f = (\vec{\nabla} f) \cdot \vec{e}_k$$

One may look at the right hand side of the directional derivative, Eq. (1.4), in the same way. It follows that the variational derivative of the functional J,

$$\frac{\delta J}{\delta y(x)} = F_y - \frac{d}{dx}\, F_{y'} \qquad x_1 < x < x_2,$$

corresponds to what in multivariable calculus is the gradient of a function. The obvious difference between the two is that the number of components of a gradient in multivariable calculus is finite and they are labeled by integers. By contrast the "gradient" of a function J, i.e. its variational derivative, has a continuous infinity of components which are labeled by the values of x $(x_1 < x < x_2)$.

1.4.3 Euler's Differential Equation

The usefulness of the first variational arises when one considers an "optimal" curve, say

$$y = f(x) \tag{1.6}$$

for which J has a maximal (or minimum, or extremal) value. In such a circumstance all the directional derivates, Eq. (1.4), vanish, i.e.

$$0 = \frac{\partial J(\alpha_i)}{\partial \alpha_k} = \int_{x_1}^{x_2} \frac{\delta J}{\delta y(x)} h_k(x)dx \qquad k = 1, 2, \ldots$$

It follows that for $x_1 < x < x_2$

$$\frac{\delta J}{\delta y(x)} \equiv F_y(x; y(x), y'(x)) - \frac{d}{dx} F_{y'}(x; y(x), y'(x)) = 0. \tag{1.7}$$

This is *Euler's differential equation.* This is the equation which the curve, Eq. (1.6) must satisfy if it extremizes the functional J.

Why must one have $\dfrac{\delta J}{\delta y(x)} = 0$ for $x_1 < x < x_2$? Suppose $\dfrac{\delta J}{\delta y(x)} \neq 0$ in the neighborhood of some point x.
Then from the set

$$\{h_k(x) : k = 1, 2, \dots\}$$

we choose or construct a localized "blip" function

$$\delta f(x) = \sum_{k=1}^{\infty} \beta_k h_k(x) \ .$$

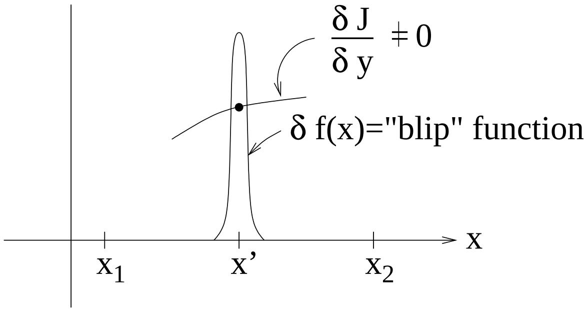


Figure 1.8: A "blip" function is a variation which is differentiable but is nonzero only in a neighborhood around a point x'.

An example of a smooth blip function which vanishes for

$$r \equiv |x - x'| \geq \epsilon$$

is

$$\delta f(x) = \begin{cases} \exp\left(-\dfrac{\epsilon^2}{\epsilon^2 - r^2}\right) & \text{for } r < \epsilon \\ 0 & \text{for } r \geq \epsilon. \end{cases}$$

Its use would yield

$$0 \neq \int_{x_1}^{x_2} \frac{\delta J}{\delta y(x)} \ \delta f(x) dx = \sum_{k=1}^{\infty} \frac{\partial J}{\partial \alpha_k} \beta_k$$

This contradicts the fact that

$$0 = \frac{\partial J}{\partial \alpha_k} = 0 \qquad \forall k.$$

The conclusion is therefore that Euler's differential Eq.(1.7) is indeed the equation which an optimal curve must satisfy. Explicitly, with $F = F(x; y(x), y'(x))$, this equation is

$$\boxed{\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} = 0} \ ,$$

or

$$y'' F_{y'y'} + y' F_{y'y} + F_{y'x} - F_y = 0 \ .$$

This Euler's differential equation is one of second order. It constitutes a necessary condition which $y(x)$ must satisfy if it is to be an optimal curve for the functional $J[y]$.

The Euler equations for the three extremum problems in Section 2 are as follows:

1. Brachistochrone:

$$F = \sqrt{\frac{1+y'^2}{x}} \qquad \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\frac{1}{\sqrt{x}}\right) = 0$$

2. Least area:

$$F = x\sqrt{1+y'^2} \qquad xy'' + y'(1+y'^2) = 0$$

3. Fermat:

$$F = n(x,y)\sqrt{1+y'^2} \qquad ny'' = (n_y - n_x y')(1+y'^2)$$

1.5 Solved Example

Let us solve the Brachistochrone problem for which $F = \sqrt{\frac{1+y'^2}{x}}$. The fact that $\frac{\partial F}{\partial y} = 0$ implies that the Euler equation is

$$0 = \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \ .$$

Consequently

$$\frac{y'}{\sqrt{1+y'^2}}\ \frac{1}{\sqrt{x}} = \text{const} \equiv \left(\frac{1}{2a}\right)^{1/2} .$$

Square the result to obtain

$$\frac{1}{1+y'^2}\ \frac{y'^2}{x} = \frac{1}{2a}$$

or

$$y'^2 2a = x + xy'^2 .$$

Thus

$$\left(\frac{dy}{dx}\right)^2 = \frac{x}{2a-x} = \frac{x^2}{2ax-x^2} . \tag{1.8}$$

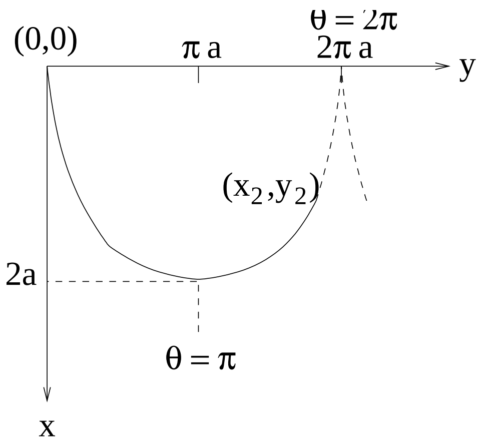


Figure 1.9: The solution curve for the brachistochrone problem is a cycloid.

Equation (1.8) can be integrated. One obtains

$$\begin{aligned} y(x) &= \int_{x_1}^{x} \frac{xdx}{(a^2-a^2+2ax-x^2)^{1/2}} \\ &= \int_{x_1}^{x} \frac{xdx}{\sqrt{a^2-(a-x)^2}} \end{aligned}$$

We make the substitution $a - x = a\cos\theta$, or

$$x = a(1 - \cos\theta)$$

to obtain

$$y = \int_{\theta_1}^{\theta} \frac{a(1-\cos\theta)a\sin\theta}{a\sin\theta} d\theta \ .$$

Thus we have

$$\begin{aligned} y &= a(\theta - \sin\theta) + const. \ \leftarrow \text{zero to make } y_1 = 0 \text{ the starting height.} \\ x &= a(1 - \cos\theta) \ . \end{aligned}$$

This is a θ-parametrized cycloid whose amplitude a must be adjusted so that the curve goes through (x_2, y_2).

Lecture 3

1.6 Integration of Euler's Differential Equation.

The Brachistochrone problem illustrates the simplification which occurs if the variational integrand F is *independent* of y. In that case

$$\frac{\partial F}{\partial y'} = \text{const.} \tag{1.9}$$

along the curve, and one has to deal only with a 1st order differential equation. Equation (1.9) is called an *integral of motion*. If F is *independent* of x then an analogous simplification is possible. This is seen best from the "Second Form" of Euler's Equation. We wish to obtain another *integral of motion*. We note that

$$\frac{dF(x,y,y')}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\, y' + \frac{\partial F}{\partial y'}\, y''$$

and

$$\frac{d}{dx}\left(y' \frac{\partial F}{\partial y'} \right) = y'' \frac{\partial F}{\partial y'} + y' \frac{d}{dx}\left(\frac{\partial F}{\partial y'} \right) \ .$$

To obtain an expression with no derivatives higher than the first, subtract to eliminate the y'' term and obtain

$$\frac{d}{dx}\left(F - y'\,\frac{\partial F}{\partial y'}\right) = \frac{\partial F}{\partial x} + \left(\frac{\partial F}{\partial y} - \frac{d}{dx}\,F_{y'}\right) y'$$

By Euler's equation the last term vanishes and we are left with the "second form" of Euler's equation

$$\boxed{\frac{d}{dx}\left(F - y'\;F_{y'}\right) = \frac{\partial F}{\partial x}}$$

Remark:

1. If the variational integrand F is independent of x, i.e. if $\frac{\partial F}{\partial x} = 0$, then one obtains an *integral of motion*,

   $$F - y'F_{y'} = \text{constant}. \tag{1.10}$$

   This expression involves only the first derivative of the to-be-determined function $y(x)$. It is evident that we have *reduced* the problem from solving a second order differential equation to solving one of first order.

2. By switching independent variables from x to y:

   $$F\left(x, y,\,\frac{dy}{dx}\right)\frac{dx}{dy}\;dy = G\left(y, x, \frac{dx}{dy}\right)\;dy$$

   the Euler equation becomes

   $$0 = \frac{d}{dy}\left(\frac{\partial G}{\partial\left(\frac{dx}{dy}\right)}\right) - \frac{\partial G}{\partial x}\;,$$

   which implies that

   $$\frac{\partial G}{\partial\left(\frac{\partial x}{\partial y}\right)} = const.$$

   is an integral of motion. Thus there is an alternate way of arriving at the integral of motion (1.10).

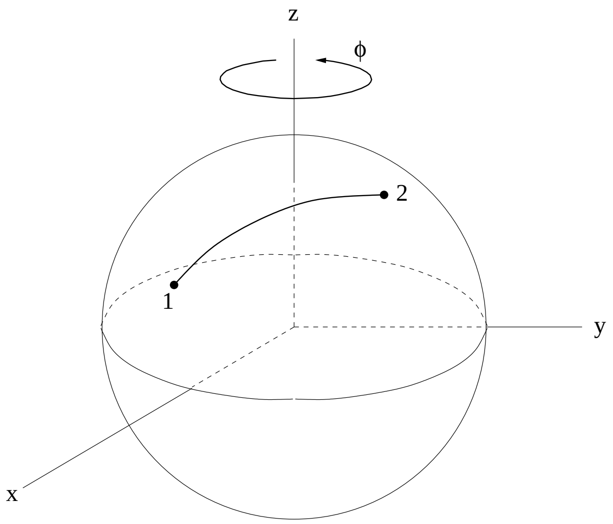


Figure 1.10: Curve on a two-sphere of radius ρ.

Example. Geodesic on a Sphere.

A geodesic is a line which represents the shortest path between two given points when the path is restricted to lie on some surface. On a sphere of radius ρ the square element of length is

$$(ds)^2 = \rho^2 d\theta^2 + \rho^2 \sin^2 \theta d\varphi^2$$

Hence the length of a path $\theta(\varphi)$ between two points, say 1 and 2, is

$$s = \int_1^2 ds = \rho \int_1^2 \sqrt{\left(\frac{d\theta}{d\varphi}\right)^2 + \sin^2 \theta(\varphi)}\; d\varphi$$

Here the variational integrand is

$$F = \sqrt{\theta_\varphi^2 + \sin^2 \theta}$$

Note that F is independent of φ. Consequently, one can apply the "second form" of Euler equation. One finds that

$$F - \theta_\varphi \frac{\partial F}{\partial \theta_\varphi} = \text{constant} \equiv a.$$

Explicitly one has

$$\sqrt{\theta_\varphi^2 + \sin^2\theta} - \theta_\varphi \frac{\theta_\varphi}{\sqrt{\theta_\varphi^2 + \sin^2\theta}} = a$$

This is a differential equation which one readily solves in the usual way: Multiply by $\sqrt{}$ and cancel θ_φ^2 to obtain

$$\sin^2\theta = a\sqrt{\left(\frac{d\theta}{d\varphi}\right)^2 + \sin^2\theta} \tag{1.11}$$

Isolate the derivative by solving for it. One obtains

$$\begin{aligned}
\frac{d\varphi}{d\theta} &= \left(\frac{\sin^4\theta}{a^2} - \sin^2\theta\right)^{-1/2} \\
&= \frac{\csc^2\theta}{\sqrt{\frac{1}{a^2} - \csc^2\theta}} = \frac{\csc^2\theta}{\sqrt{\beta^2 - \cot^2\theta}}.
\end{aligned}$$

Here $\beta^2 = \frac{1}{a^2} - 1$. The antiderivative is

$$\varphi(\theta) = -\sin^{-1}\frac{\cot\theta}{\beta} + \alpha \ .$$

Consequently, one obtains

$$-\cot\theta = \beta\sin(\varphi - \alpha)$$

This is the equation of a plane through the origin: Multiply by $\rho\sin\theta$ and expand the sine function,

$$-\underbrace{\rho\cos\theta}_{z} = \beta\underbrace{\rho\sin\theta\sin\varphi}_{y}\cos\alpha - \beta\underbrace{\rho\sin\theta\cos\varphi}_{x}\sin\alpha$$

$$\boxed{z + y\,\beta\cos\alpha - x\,\beta\sin\alpha = 0}$$

Thus the geodesic is the locus of the sphere intersected by a plane. In other words, the geodesic is a segment of a great circle.
Remark: The variational integral is extremized by *many* solutions to the Euler equation. The shortest geodesic miniimizes s. Other solutions are geodesics which consist of great circles that wind around the sphere. They constitute *saddle points* of s in the space of functions!

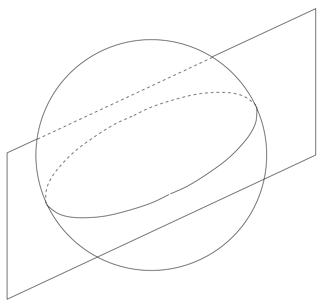


Figure 1.11: A great circle is the intersection of a sphere with a plane through the origin

Chapter 2

GENERALIZATIONS

2.1 Functional with Several Unknown Functions

Consider a variational integral which is a functional of several functions with the value of the functions fixed at the endpoints. The extremization of such an integral, namely

$$J[y_i] = \int_{x_1}^{x_2} F(x; y_1, y_1', y_2, y_2', \ldots, y_n, y_n')dx,$$

where $y_i(x)$ $i = 1, \ldots, n$ satisfy the prescribed boundary conditions at $x = x$ and $x = x_2$, is achieved by a method analogous to the one described before. Thus we consider a system of variants

$$y_j = y_j(x, 0) + \sum_{i=1} \alpha_i^j h_i(x) = f_j(x, 0) + \sum_{i=1} \alpha_i^j h_i(x) j = 1, \ldots, n$$

of the system of functions

$$y_j = y_j(x, 0) = f_j(x, 0) \quad j = 1, \ldots, n$$

which we assume extremizes the functional $J[y_i]$.

The fact that these functions and their variants have prescribed values at the endpoints x_1 and x_2 can be stated by saying that each of the α-parametrized curves,

$$\{y_j = f_j(x, \alpha_i^j), j = 1, \ldots n\} \ ,$$

passes through the same two end points. This means that the system of

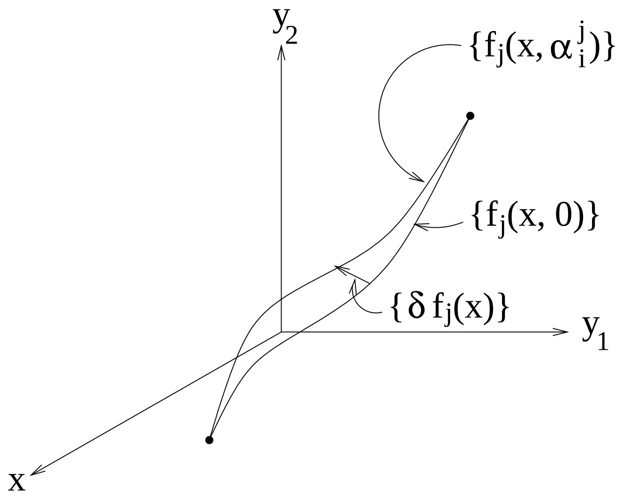


Figure 2.1: A variant and its system of variations $\delta\{f_j(x)\}$

variations

$$\delta f_j(x) \quad j = 1, \ldots, n \ ,$$

which can be pictured as an n-component vector, vanishes at the endpoints:

$$\delta f_j(x_1) = \delta f_j(x_2) = 0 \quad j = 1, \ldots, n \ .$$

The procedure for obtaining the necessary condition for J to be an extremum follows the same pattern as before. The first variation of the functional J is

$$\delta J = \sum_{j=1}^{n} \int_{x_1}^{x_2} \left\{ \frac{\partial F}{\partial y_j} \ \delta f_j + \frac{\partial F}{\partial y_j'} \ \delta \left(\frac{df_j}{dx} \right) \right\} dx$$

Using the fact that

$$\begin{aligned} \delta \left(\frac{df_j}{dx} \right) &= \frac{df_j(x, \alpha_i^j)}{dx} - \frac{df_j(x, 0)}{dx} \\ &= \frac{d}{dx} \ \delta f_j(x) \quad j = 1, \ldots, n \end{aligned}$$

and integrating by parts yields

$$\delta J = \sum_{j=1}^{n} \int_{x_1}^{x_2} \left\{ \frac{\partial F}{\partial y_j} - \frac{d}{dx} \left(\frac{\partial F}{\partial y_j'} \right) \right\} \delta f_j(x) dx + \frac{\partial F}{\partial y_j'} \ \delta f_j \Bigg|_{x_1}^{x_2} .$$

The boundary term vanishes. The fact that for J to be an extremum for arbitrary variations δf_j $j = 1, \ldots, n$, implies that

$$\frac{\delta J}{\delta y_j(x)} \equiv F_{y_j} - \frac{d}{dx}(F_{y_j'}) = 0 \quad j = 1, \ldots, n \tag{2.1}$$

the Euler equations for "a system with n degrees of freedom". This set of equations constitutes the necessary condition for the variational integral $J[y_1, \ldots, y_n]$ to be an extremum. It is evident that this condition generalizes Euler equation from a single degree of freedom, Eq.(1.7) on page 18, to many degrees of freedom.

There is, however, one qualitatively new feature. It may happen that the vanishing of the first variation of J is due to the fact that J is an *maximum* with respect to variations in, say, $y_1(x)$ but is *minimum* with respect to variations in, say, $y_2(x)$.

Whatever the case may be, we shall call a curve $\{y_1(x), \ldots, y_n(x)\}$ which solves the set of Euler equations (2.1) an "*extremal curve*", even though often an "*optimal curve*" may be more descriptive.

Example (Fermat's Principle in three dimensions).

The propagation of a beam of light or sound whose local propagation speed is

$$v = \frac{c}{n(x, y, z)}$$

is governed by the extremum principle

$$J[y, z] = c \int_{x_1}^{x_2} \frac{\sqrt{1 + y'^2 + z'^2}}{n(x, y, z)} dx = \text{Minimum.}$$

This principle assumes that the local propagation speed depends only on the position (x, y, z) of the propagating beam. In that case the propagation medium in saaaaid to be *inhomogeneous* but *isotropic*. If, however, the propagation also depends on the direction of the beam, i.e.

$$v = \frac{c}{n(x, y, z, y', z')}$$

then the medium is said to be *inhomogeneous* and *anisotropic*.

In that case the propagation law is given by

$$J[y,z] = c \int_{x_1}^{x_2} \frac{\sqrt{1+y'^2+z'^2}}{n(x,y,z,y',z')}\, dx = \text{Minimum}$$

Example (Geodesic on a Sphere)

Let us find the λ-parametrized path $(\theta(\lambda), \varphi(\lambda))$ on the two sphere which has the property that its length

$$J = \int \sqrt{\left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2\theta \left(\frac{d\varphi}{d\lambda}\right)^2}\, d\lambda = \text{Extremum}$$

There are two functions and hence two differential equations.

(i) $\quad \dfrac{d}{d\lambda}\left(\dfrac{\partial F}{\partial \theta'}\right) = \dfrac{\partial F}{\partial \theta}$

or

$$\frac{d}{d\lambda} \frac{\frac{d\theta}{d\lambda}}{\sqrt{\theta'^2 + \sin^2\theta\varphi'^2}} = \frac{\sin\theta\cos\theta}{\sqrt{\theta'^2+\sin^2\varphi'^2}} \left(\frac{d\varphi}{d\lambda}\right)^2 \tag{2.2}$$

(ii) The second equation is

$$\frac{d}{d\lambda}\left(\frac{\partial F}{\partial \varphi'}\right) = \frac{\partial F}{\partial \varphi}$$

The fact that the *variational integrand* F is independent of φ (one says that F is "cyclic" in φ) means that

$$\frac{\partial F}{\partial \varphi} = 0$$

Consequently, we have immediately an integral of motion, namely

$$\frac{\partial F}{\partial \varphi'} = \frac{\varphi' \sin^2\theta}{\sqrt{\theta'^2 + \sin^2\theta\varphi'^2}} = const. \equiv P_\varphi\ . \tag{2.3}$$

Equation (2.2) and Eq. (2.3) constitute the system of ordinary differential equations which we must solve. If we let $\lambda = \varphi$ be the curve parameter, then Eq.(2.3) becomes Eq.(1.11) (on page 25) and we can solve the problem. However, if the variational integrand had been

$$F = \left[g_{11} \left(\frac{dy_1}{d\lambda} \right)^2 + g_{22} \left(\frac{dy_2}{d\lambda} \right)^2 + 2g_{12} \frac{dy_1}{d\lambda} \frac{dy_2}{d\lambda} \right]^{1/2} ,$$

where the g_{ij}'s are functions which depend on the coordinates y_1, and y_2, then we could not have solved the problem. Nevertheless, a crisp and simple conclusion is possible.

What one does is to introduce the arclength as the curve parameter:

$$s = \int_0^\lambda \sqrt{\left(\frac{d\theta}{d\lambda} \right)^2 + \sin^2\theta \left(\frac{d\varphi}{d\lambda} \right)^2} \, d\lambda$$

or

$$ds = \sqrt{\theta'^2 + \sin^2\theta\varphi'^2} \, d\lambda \ .$$

This brings about a tremendous simplification. The arclength parameter is introduced as follows

$$\frac{d}{ds} = \frac{d\lambda}{ds} \frac{d}{d\lambda} = \frac{1}{\sqrt{\theta'^2 + \sin^2\theta\varphi'^2}} \frac{d}{d\lambda} \ . \tag{2.4}$$

One applies this differential operator to θ and obtains

$$\frac{1}{\sqrt{}} \frac{d\theta}{d\lambda} = \frac{d\theta}{ds} \ .$$

Indicating differentiation with respect to s by a dot, $\frac{d\theta}{ds} = \dot\theta$, one obtains for Eq.(2.2)

$$\frac{1}{\sqrt{}} \frac{d\dot\theta}{d\lambda} = \sin\theta\cos\theta \left(\frac{1}{\sqrt{}} \frac{d\varphi}{d\lambda} \right)^2 \ .$$

Applying the above differential operator to $\dot\theta$ and to φ twice, one obtains

$$\ddot\theta = \dot\varphi^2 \sin\theta\cos\theta \ .$$

Similarly, the constant of motion, Eq.(2.3) becomes

$$P_\varphi = \dot\varphi \sin^2\theta \ .$$

If s is proportional to time, then in mechanics this is the "angular momentum" around the z-axis.

The introduction of the arclength s as the curve parameter always implies that

$$\left(\frac{d\theta}{ds}\right)^2 + \sin^2\left(\frac{d\varphi}{ds}\right)^2 = constant$$

along a curve. This expression is an integral of motion even for curves which are not extremals. The availability of this integral is guaranteed without having to integrate the Euler equations for the variational problem. Indeed, one always has

$$1 = \frac{ds}{ds} = \frac{\sqrt{d\theta^2 + \sin^2 d\varphi^2}}{ds} = \sqrt{\left(\frac{d\theta}{ds}\right)^2 + \sin^2\left(\frac{d\varphi}{ds}\right)^2}\,. \tag{2.5}$$

The constancy of this quantity will be discussed and used again on page 46, but this time in the context of a *constrained* variational problem.
Remarks:

1. The arclength parameter *always* simplifies the parametric representation in this way.

2. Equation (2.5) constitutes an integral of motion:

   $$1 = \left(\frac{d\theta}{ds}\right)^2 + \sin^2\theta\left(\frac{d\varphi}{ds}\right)^2.$$

   It is a constant of motion along *any curve*, geodesic or non-geodesic on S^2, provided s is the arclength parameter. In other words,

   $$\frac{d}{ds}\left[\sqrt{\left(\frac{d\theta}{ds}\right)^2 + \sin^2\theta\left(\frac{d\varphi}{ds}\right)^2}\,\right] = 0$$

   along *any* curve. Thus there always is an "energy type" integral

   $$\frac{1}{2}\left[\left(\frac{d\theta}{ds}\right)^2 + \sin^2\theta\left(\frac{d\varphi}{ds}\right)^2\right] = constant.$$

   The same conclusion holds in the general context as expressed by the variational integrand given by Eq. (2.4).

The conclusion is this: For a geodesic we have

$$\begin{aligned} &\ddot{\theta} = \dot{\varphi}^2 \sin\theta\cos\theta \\ &\dot{\varphi}\sin^2\theta = const. && \text{“angular momentum”} \\ &\frac{1}{2}(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2) = \frac{1}{2} && \text{“energy”} \end{aligned}$$

3. These three equations in two unknowns are not independent. Instead, any *two* equations imply the *third*. For example, $\frac{d}{ds}(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2) = 0$ because of the two Euler-Lagrange equations. The benefit is that one can grasp the trajectories on a two-sphere in terms of the two simple integrable equations,

$$\frac{1}{2}\left(\dot{\theta}^2 + \frac{(const.)^2}{\sin^2\theta}\right) = constant$$

and

$$\dot{\varphi}\sin^2\theta = const. \ ,$$

whose solutions are arclength-parametrized curves.

Lecture 4

2.2 Extremum Problem with Side Conditions.

We would like to consider a class of problems in the calculus of variations, the most famous of which is the iso-perimetric problem; namely, find a curve such that

$$\text{Area} = \iint dxdy = \oint y dx = J[y] = \text{extremum}$$

subject to the constraint that the bounding curve $y(x)$ has fixed length

$$\ell = \oint \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx = \text{constant.}$$

An interesting generalization of this problem was the one which confronted Dido. Her problem was to circumscribe by means of a curve of *fixed length* an area which yields the largest amount of fertility. In other words, for a

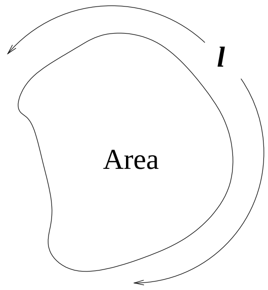


Figure 2.2: Area having perimeter of length ℓ.

given function $\rho(x,y)$ (say, yield per unit area), extremize the weighted area integral

$$\iint \rho(x,y)dxdy$$

subject to the condition that the bounding curve $y(x)$ has fixed length, $\int \sqrt{1+\left(\frac{dy}{dx}\right)^2} \equiv \ell = \text{constant}$.

More generally, the problem is to find in the space of functions that function $y(x)$ for which $J[y] \equiv \int_{x_1}^{x_2} F(x;y,y')dx = \text{extremum}$, but which at the same time satisfies the constraints

$$\begin{aligned} K_1[y] &\equiv \int_{x_1}^{x_2} G_1(x;y,y')dx = \text{constant} \equiv c_1 \\ &\vdots \\ K_k[y] &\equiv \int_{x_1}^{x_2} G_k(x;y,y')dx = \text{constant} \equiv c_k \end{aligned} \tag{2.6}$$

2.2.1 Heuristic Solution

We can solve this optimization ("extremum") problem with constraints by setting up the necessary conditions the desired function $y(x)$ must satisfy. As usual, we consider the variants,

$$y(x) + \delta y(x)$$

of $y(x)$. Each variant is uniquely characterized by an infinitesimal vector, the variation $\delta y(x)$ of $y(x)$. The resulting variations in J and K_i are

$$\begin{aligned} \delta J &= J[y+\delta y] - J[y] \equiv \int_{x_1}^{x_2} \frac{\delta J}{\delta y(x)}\, \delta y(x) dx \\ \delta K_i &= K_i[y+\delta y] - K[y] \equiv \int_{x_1}^{x_2} \frac{\delta K_i}{\delta y(x)}\, \delta y(x) dx \end{aligned}$$

There are two types of variants $y(x) + \delta y(x)$:

I. those which do and

II. those which do not

satisfy the given constraints.

I. Those which do satisfy the contraints have variations which are perpendicular to the gradient of the constraint surfaces $K_i = c_i (i = 1, \ldots, k)$, i.e. those for which

$$\begin{aligned} \int_{x_1}^{x_2} \frac{\delta K_i}{\delta y(x)}\, \delta y(x) dx &= 0 \quad \forall i = 1, \ldots, k \\ K_i[y] &= c_i \quad i = 1, \ldots, k_1 \end{aligned}$$

We shall call such vectors the "yes" vectors. They are tangent to all the constraint manifolds.

II. Those variants which do *not* satisfy the constraints have variations which are *not* perpendicular to the constraint surfaces $K_i = c_i$, i.e. those for which

$$\int_{x_1}^{x_2} \frac{\delta K_i}{\delta y(x)}\, \delta y(x) dx \neq 0 \quad \text{for some } i = 1, \ldots, k$$

We shall call such vectors the "no" vectors.

Thus a variation vector $\delta y(x)$ always belongs to class I ("yes") or to class II ("no"): If δy belongs to class I, then

$$\delta J = \int_{x_1}^{x_2} \frac{\delta J}{\delta y(x)}\, \delta y(x) dx = 0$$

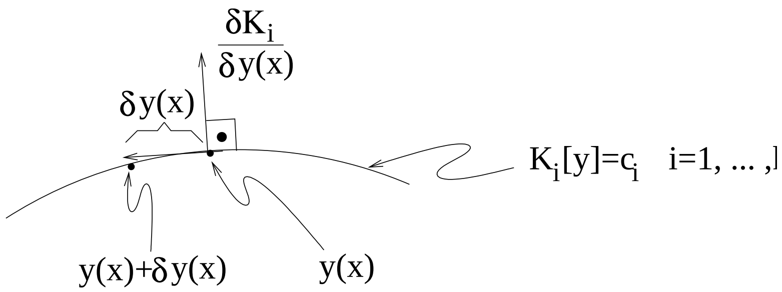


Figure 2.3: ith constraint manifold: its gradient $\frac{\delta K_i}{\delta y(x)}$ which is perpendicular to any class I ("yes") vector $\delta y(x)$ which is tangent to that manifold, $K_i[y] = c_i$.

because $y(x)$ was assumed to be an optimal function which satisfies the constraint, i.e. for which

$$\delta K_i = \int_{x_1}^{x_2} \frac{\delta K_i}{\delta y(x)} \, \delta y(x) dx = 0 \quad i = 1, \ldots, k \text{ (for class I variants).}$$

If, on the other hand, δy belongs to class II, i.e. such that

$$\delta K_i \equiv \int \frac{\delta K_i}{\delta y(x)} \, \delta y(x) dx \neq 0 \quad \text{for some } i = 1, \ldots, k,$$

then we can, and we will, choose constants $\lambda_1, \ldots, \lambda_k$ such that

$$\delta J = \lambda_1 \delta K_1 + \ldots \lambda_k \delta K_k.$$

The conclusion is therefore this: *all* variation vectors $\delta y(x)$ satisfy the equation

$$\int_{x_1}^{x_2} \left[\frac{\delta J}{\delta y(x)} - \lambda_1 \frac{\delta K_1}{\delta y(x)} - \ldots \lambda_k \frac{\delta K_k}{\delta y(x)} \right] \delta y(x) dx = 0$$

whenever

$$J - \lambda_1 K_1 - \cdots - \lambda_k K_k = \text{extremum}$$

where the constants $\lambda_1, \ldots, \lambda_k$ *must be chosen* appropriately. They are determined by the values $c_1, \ldots, c_k$ of the constraints.

Summary:

To solve the optimization problem with constraints, solve the following $k+1$ equations

$$\begin{aligned} \frac{\delta}{\delta y(x)}\left(J-\lambda_1 K_1-\ldots \lambda_k K_k\right) &= 0 \\ K_1 &= c_1 \\ &\vdots \\ K_k &= c_k \end{aligned}$$

for $y(x)$ and $\lambda_1, \ldots, \lambda_k$. The constants $\lambda_1, \ldots, \lambda_k$ are called "Lagrange multipliers".

2.2.2 Solution via Constraint Manifold

The "yes" and "no" vectors play a basic role in establishing necessary conditions for the existence of an extremum subject to constraints. It is therefore appropriate to highlight how these vectors tie into well-established knowledge. We shall do this by splitting the space of tangent vectors into two orthogonal vector spaces, whose only common vector is the zero vector.

Finite-dimensional Space

Our task is to extremize a given function $f(\vec{x})$ on the finite-dimensional space R^n subject to a finite number of constraints, say, $g_1(x) = c_1, \cdots, g_k(x) = c_k$, with $k < n$. More precisely, we wish to find the equations for those critical points $\vec{x}_0$ of $f(\vec{x})$ which lie on the constraint manifold, which is the intersection of the constraint hypersurfaces given by the finite number of constraints.

We shall solve this problem in two steps. First we shall develop the chain of reasoning when there are no constraints present. Then we shall extend this reasoning to the case where k constraints are present.

Consider the trajectory $\vec{x} = \vec{r}(t)$ of a particle in the conservative force field ∇f of a potential $f(\vec{x})$. Then

$$\nabla f \cdot \frac{d\vec{r}}{dt} = \text{"force} \times \text{velocity"} \tag{2.7}$$

is the rate at which the force field does work on the particle.

Suppose there exists a point $\vec{x}_0$ such that

$$\nabla f(\vec{x}_0) \cdot \frac{d\vec{r}}{dt} = 0$$

for the tangent vectors of *all* curves $\vec{x} = \vec{r}(t)$ passing through this point. These vectors form the *tangent space* at $\vec{x}_0$. One now makes the observation that

$$\nabla f(\vec{x}_0) \cdot \frac{d\vec{r}}{dt} = 0 \Rightarrow \nabla f(\vec{x}_0) \in \text{orthogonal complement of the tangent space at } \vec{x}_0 \,.$$

The orthogonal complement of the space of *all* tangent vectors is zero dimensional: it consists of the zero vector only. Thus

$$\nabla f(\vec{x}_0) = \vec{0} \,.$$

This is the necessary condition for $\vec{x}_0$ to be the location of an unconstrained extremum of f.

Next, suppose there exists a point $\vec{x}_0$ such that

$$\nabla f(\vec{x}_0) \cdot \frac{d\vec{r}}{dt} = 0 \tag{2.8}$$

for the tangent vectors of only those curves $\vec{x} = \vec{r}(t)$ passing through $\vec{x}_0$ which lie in the *constraint manifold*

$$\{\vec{x} : \; g_1(\vec{x}) = c_1, \cdots, g_k(\vec{x}) = c_k\} \,, \tag{2.9}$$

and hence satisfy

$$\nabla g_1(\vec{x}_0) \cdot \frac{d\vec{r}}{dt} = 0, \cdots, \nabla g_k(\vec{x}_0) \cdot \frac{d\vec{r}}{dt} = 0 \,. \tag{2.10}$$

The collection of these vectors $\frac{d\vec{r}}{dt}$ forms a *subspace* of the above tangent space: any one of their linear combinations also satisfies these k constraints. This subspace is called the *constraint subspace* at $\vec{x}_0$. One now makes the observation that
$\nabla f(\vec{x}_0) \cdot \frac{d\vec{r}}{dt} = 0 \;\Rightarrow\; \nabla f(\vec{x}_0) \in$ orthogonal complement of the constraint subspace of vectors, which satisfy $\nabla g_1(\vec{x}_0) \cdot \frac{d\vec{r}}{dt} = 0, \cdots, \nabla g_k(\vec{x}_0) \cdot \frac{d\vec{r}}{dt} = 0$.
This orthogonal complement consists precisely of those vectors which have the form $\lambda_1 \nabla g_1 + \cdots + \lambda_k \nabla g_k$. Hence

$$\nabla f(\vec{x}_0) = \lambda_1 \nabla g_1 + \cdots + \lambda_k \nabla g_k \tag{2.11}$$

for some set of coefficients $\lambda_1, \cdots, \lambda_k$. They are called the *Lagrange multipliers* of the constrained extremum problem. It is evident that their values and the location of the critical point of f on the constraint manifold are given by the following "Lagrange Multiplier" theorem

Theorem 1

Given: *Let f be a scalar function which has an extremum at $\vec{x}_0$ on the intersection of the k constraint hypersurfaces ("the constraint manifold"), Eq.(2.9).*

Conclusion: *The location of the constrained critical point is determined by*

$$\boxed{\nabla f(\vec{x}_0) - \lambda_1 \nabla g_1(\vec{x}_0) - \cdots - \lambda_k \nabla g_k(\vec{x}_0) = 0\}} \tag{2.12}$$

where the coefficients $\lambda_1, \cdots, \lambda_k$ are obtained with the help of

$$\begin{gathered} g_1(\vec{x}_0) = c_1 \\ \vdots \\ g_k(\vec{x}_0) = c_k \end{gathered}$$

Remark 1: In mechanics where $f(\vec{x})$ is a scalar potential and $\nabla f(\vec{x})$ its (conservative) force field, the force

$$\nabla f(\vec{x}_0) = \lambda_1 \nabla g_1(\vec{x}_0) + \cdots \lambda_k \nabla g_k(\vec{x}_0)$$

is the constraining force which guarantees that the point particle stays in equilibrium ("no net forces") on the constraint manifold.
Remark 2: This theorem is easily visualized in three dimensions, $R^n = R^3$, with two constraints, say $g_1(\vec{x}) = c_1$ and $g_2(\vec{x}) = c_2$. There the constraint manifold is one-dimensional, a single curve $\vec{r}(t)$, the intersection of the two surfaces $g_1(\vec{x}) = c_1$ and $g_2(\vec{x}) = c_2$. Consequently,

$$\begin{aligned} \vec{\nabla} g_1 \cdot \frac{d\vec{r}}{dt} &= 0 \\ \vec{\nabla} g_2 \cdot \frac{d\vec{r}}{dt} &= 0 \end{aligned}$$

The fact that f has an extremum on this curve implies

$$\vec{\nabla} f \cdot \frac{d\vec{r}}{dt} = 0$$

Consequently, $\exists\lambda_1$ and λ_2 such that

$$\nabla f = \lambda_1 \vec{\nabla} g_1 + \lambda_2 \vec{\nabla} g_2.$$

Example (Function with two finite constraints):
In multivariable calculus the problem

$$f(x,y,z) = \text{extremum}$$

subject to

$$\begin{aligned} g(x,y,z) &= c_1 \\ h(x,y,z) &= c_2 \end{aligned}$$

is solved by considering

$$f^* = f - \lambda_1 g - \lambda_2 h$$

and solving

$$\begin{aligned} \vec{\nabla}(f - \lambda_1 g - \lambda_2 h) &= 0 \\ g(x,y,z) &= c_1 \\ h(x,y,z) &= c_2 \end{aligned}$$

for x_0, y_0, z_0, λ_1 and λ_2. These are 3+2 equations in 3+2 unknowns.

Infinite-dimensional Space

The Lagrange Multiplier theorem readily extends to a functional which is defined on curves in function space, which is infinite-dimensional. The development runs parallel to the one preceeding Theorem 1.

First of all recall that a "curve in function space" is a 1-parameter set of functions

$$y(x,t) \qquad a \le t \le b\ . \tag{2.13}$$

Second, the rate of change of the functional $J[y]$ along this 1-parameter set is

$$\begin{aligned}
\frac{d}{dt}J[y] &= \frac{d}{dt}\int_{x_1}^{x_2} F(x;y(x,t),y'(x,t))dx \\
&= \int_{x_1}^{x_2}\left[\left.\frac{\partial F}{\partial y}\right|_{y(x,t)}\frac{\partial y(x,t)}{\partial t}+\left.\frac{\partial F}{\partial y'}\right|_{y(x,t)}\frac{\partial y'(x,t)}{\partial t}\right]dx \\
&= \int_{x_1}^{x_2}\left.\frac{\delta J}{\delta y(x)}\right|_{y(x,t)}\frac{\partial y(x,t)}{\partial t}dx\ ,
\end{aligned}$$

where $\delta J/\delta y(x)$ is the variational derivative of J as defined by Eq.(1.5) on page 18.

Third, if $J[y]$ has an extremal value on the parametrized set of functions $y(x,t)$ at $t = t_0$, which is to say that its rate of change $dJ[y]/dt$ vanishes at $t = t_0$, then

$$0 = \left.\frac{d}{dt}J[y]\right|_{t=t_0} = \int_{x_1}^{x_2}\left.\frac{\delta J}{\delta y(x)}\right|_{y(x,t_0)}\left.\frac{\partial y(x,t)}{\partial t}\right|_{t_0}dx\ . \tag{2.14}$$

replaces Eq.(2.8) as the extremum condition.

Fourth, the fact that the one-parameter family $y(x,t)$ lies in the intersection of the constraints, Eqs.(2.6) on page 34, implies that

$$\begin{aligned}
\frac{dK_1[y]}{dt} &= \int_{x_1}^{x_2}\left.\frac{\delta K_1}{\delta y(x)}\right|_{y(x,t)}\frac{\partial y(x,t)}{\partial t}dx = 0 \\
&\vdots \\
\frac{dK_k[y]}{dt} &= \int_{x_1}^{x_2}\left.\frac{\delta K_k}{\delta y(x)}\right|_{y(x,t)}\frac{\partial y(x,t)}{\partial t}dx = 0\ .
\end{aligned} \tag{2.15}$$

At $t = t_0$ these equations replace Eqs.(2.10).

Finally, Eq.(2.11) gets replaced by

$$\frac{\delta J}{\delta y(x)} = \lambda_1\frac{\delta K_1}{\delta y(x)}+\cdots+\lambda_k\frac{\delta K_k}{\delta y(x)}\ .$$

Nota bene: A class I ("yes") or a class II ("no") vector $\delta y(x)$ on page 35 is an infinitesimal vector ("a variation")

$$\delta y(x) = \frac{\partial y(x,t)}{\partial t}\delta t$$

tangent to one of the t-parametrized sets of functions ("curve in function space"), Eq.(2.13), on page 40. A curve which satisfies the integral constraints has tangent vectors which obey Eq.(2.15). These tangents are proportional to the "yes" vectors. The orthogonal complement of all such vectors make up the vector space of "no" vectors. Each of them is some linear combination of $\delta K_1/\delta y(x), \cdots, \delta K_k/\delta y(x)$.

In spite of the fact that the domain of the functional $J[y]$ is infinite-dimensional, the line of reasoning that leads to the infinite-dimensional version of the Lagrange Multiplier Theorem 1 remains unchanged. Thus one has

Theorem 2

Given: *Let*

$$J[y] = \int_{x_1}^{x_2} F(x,y,y')\,dx$$

be a scalar functional which has an optimal function $y(x)$ satisfying the k constraint conditions (2.6) on page 34.

Conclusion: *This optimal function satisfies the differeential equation*

$$F_y - \frac{d}{dx}F_{y'} - \lambda_1\left(G_{1y} - \frac{d}{dx}G_{1y'}\right) - \cdots - \lambda_k\left(G_{ky} - \frac{d}{dx}G_{ky'}\right) = 0\ , \tag{2.16}$$

where the coefficients $\lambda_1, \cdots, \lambda_k$ are obtained with the help of

$$\int_{x_1}^{x_2} G_1(x,y,y')\,dx = c_1$$

$$\vdots$$

$$\int_{x_1}^{x_2} G_k(x,y,y')\,dx = c_k\ .$$

Example (Functional with two integral constraints):

Apply Lagrange's method to solve the variational problem

$$
\begin{aligned}
J[y] &\equiv \int_{x_1}^{x_2} F(x,y,y')\,dx = \text{extremum} \\
K[y] &\equiv \int_{x_1}^{x_2} G(x,y,y')\,dx = c_1 = \text{fixed} \\
L[y] &\equiv \int_{x_1}^{x_2} H(x,y,y')\,dx = c_2 = \text{fixed}
\end{aligned}
$$

This problem is a generalization to infinite dimensions of the example on page 40. One proceeds by first constructing

$$J^* = J - \lambda_1 K - \lambda_2 L$$

and then considering

$$\delta J^* = \delta(J - \lambda_1 K - \lambda_2 L)\ .$$

Thus

$$
\begin{aligned}
\delta J^* = \quad & \int \left(\frac{\partial F}{\partial y} - \frac{d}{dx}\,\frac{\partial F}{\partial y'} \right) \delta f(x) dx - \lambda_1 \int \left(\frac{\partial G}{\partial y} - \frac{d}{dx}\,\frac{\partial G}{\partial y'} \right) \delta f(x) dx \\
& - \lambda_2 \int \left(\frac{\partial H}{\partial y} - \frac{d}{dx}\,\frac{\partial H}{\partial y'} \right) \delta f(x) dx\ .
\end{aligned}
$$

The problem is solved by setting

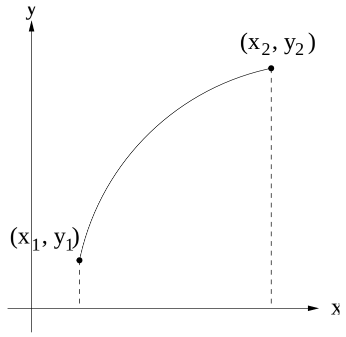
$$\delta J^* = \delta J - \lambda_1 \delta K - \lambda_2 \delta L = 0\ .$$

Thus one has the following set of equations

$$
\begin{aligned}
\frac{\partial}{\partial y}(F - \lambda_1 G - \lambda_2 H) - \frac{d}{dx}\,\frac{\partial}{\partial y'}(F - \lambda_1 G - \lambda_2 H) &= 0 \\
\int\limits_{x_1}^{x_2} G(x;y,y')dx &= c_1 \\
\int\limits_{x_1}^{x_2} H(x;y,y')dx &= c_2
\end{aligned}
$$

The solutions to these equations yield 4 constants:
(1) 2 integration constants
(2) 2 Lagrange multipliers
The 4 constants are determined by
(1) 2 boundary conditions on $y(x)$
(2) 2 constraints $K[y] = c_1$ and $L[y] = c_2$.

Example (Isoperimetric Problem):



Among all the curves $(x(\lambda),\, y(\lambda))$ of fixed length ℓ,

$$\int_{\lambda_1}^{\lambda_2} \sqrt{\left(\frac{dx}{d\lambda}\right)^2 + \left(\frac{dy}{d\lambda}\right)^2}\, d\lambda = \text{constant} = \ell,$$

through the points (x_1, y_1) and (x_2, y_2) find that curve for which

$$\int_{\lambda_1}^{\lambda_2} y\, \frac{dx}{d\lambda}\, d\lambda = \text{extremum}$$

i.e. extremize the area!
Remark. Other formulations of this problem are
(i) $\int\ y\ dx$=extremum
subject to

$$\int \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = \text{fixed}$$

(ii) $\int\ y\ \frac{dx}{ds}\ ds$=extremum subject to

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

Solution: Consider the variational integrand

$$F^* = yx' - \mu\sqrt{x'^2 + y'^2}$$

where μ is a Lagrange multiplier. The Euler equations are

$$\begin{aligned} \frac{d}{d\lambda}\frac{\partial F^*}{\partial x'} &= \frac{\partial F^*}{\partial x} \\ \frac{d}{d\lambda}\frac{\partial F^*}{\partial y'} &= \frac{\partial F^*}{\partial y} \end{aligned}$$

The fact that x is a cyclic coordinate, i.e. that $\frac{\partial F^*}{\partial x} = 0$ implies that $\frac{\partial F^*}{\partial x'}$ is constant:

$$y - \mu\frac{x'}{\sqrt{x'^2 + y'^2}} = \text{constant} \equiv C$$

The second equation gives us

$$\frac{d}{d\lambda}\left((-)\mu\frac{y'}{\sqrt{x'^2 + y'^2}}\right) - x' = 0$$

The form of these equations suggests that one introduce arclength,

$$ds = \sqrt{x'^2 + y'^2}d\lambda,$$

as a new curve parameter. In terms of this parameter we have

$$\frac{1}{\sqrt{\left(\frac{dx}{d\lambda}\right)^2 + \left(\frac{dy}{d\lambda}\right)^2}}\frac{d}{d\lambda} = \frac{d}{ds} = \text{“ . ”}$$

We also notice that for *any* curve extremal or nonextremal, the use of s as a curve parameter implies

$$\sqrt{\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2} = 1 \tag{2.17}$$

In terms of the derivative with respect to arclength the Euler equations become

$$y - \mu\dot{x} = c \tag{2.18}$$

and

$$\mu\ddot{y} + \dot{x} = 0 \tag{2.19}$$

> *Comments about consistency*: Any two of the equations (2.17), (2.18), and (2.19) imply the third. Thus, notice that the last two equations are consistent with the requirement that
>
> $$\boxed{\dot{x}^2 + \dot{y}^2 = \text{const}}$$
>
> This requirement was already highlighted on page 32, but we also validate it by direct differentiation: Differentiate Eq(2.18)
>
> $$\begin{aligned} \mu\ddot{x} - \dot{y} &= 0 \\ \mu\ddot{y} + x &= 0 \end{aligned} \tag{2.20}$$
>
> These are the unintegrated Euler equations. Insert them into the derivative of $\dot{x}^2 + \dot{y}^2$,
>
> $$\frac{d}{ds}(\dot{x}^2 + \dot{y}^2) = 2\ddot{x}\dot{x} + 2\ddot{y}\dot{y} = 2\ddot{x}(-)\lambda\ddot{y} + 2\ddot{y}\lambda\ddot{x} = 0$$
>
> Thus $\dot{x}^2 + \dot{y}^2$=constant along the curve. This is consistent with Eq.(2.17). This constancy is a general property and we shall return to it when we discuss the symmetries of the functional J.

Now return to the task of solving Eqs (2.18) and (2.19). Insert Eq.(2.18) into (2.19) and obtain

$$\ddot{y} + \frac{1}{\mu^2}\,y = \frac{c}{\mu^2}$$

The solution is

$$\boxed{y = A\cos\left(\frac{s}{\mu} + \delta\right) + c} \tag{2.21}$$

To obtain $x(s)$, use Eq.(2.18),

$$\dot{x} = \frac{y-c}{\mu} = \frac{A}{\mu}\cos\left(\frac{s}{\mu}+\delta\right) \ ,$$

and obtain

$$\boxed{x = A\sin\left(\frac{s}{\mu}+\delta\right) + x_0} \tag{2.22}$$

Equations (2.21) and (2.22) are parametrized equations for a *circular arc* because $(y-c)^2 + (x-x_0)^2 = A^2$.

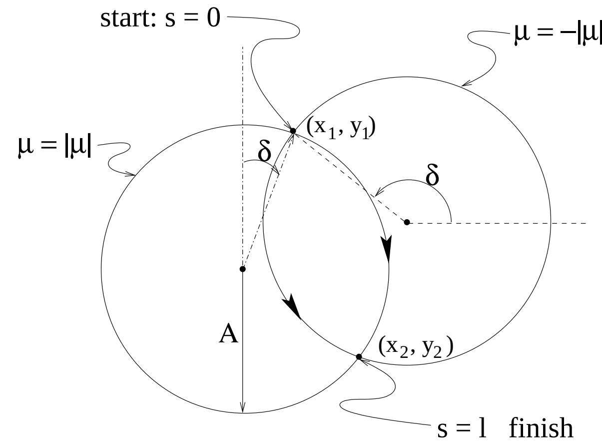


Figure 2.4: Two intersecting circles as solutions to the isoperimetric problem.

We have five constants A, δ, μ, c, and x_0 whose values are determined by five conditions, namely,

$$\text{At } s=0: \ x(0) = x_1 \tag{2.23}$$
$$y(0) = y_1 \tag{2.24}$$
$$\text{At } s=\ell: \ x(\ell) = x_2 \tag{2.25}$$
$$y(\ell) = y_2 \tag{2.26}$$
$$\int_0^\ell \sqrt{\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2} ds = \ell \ . \tag{2.27}$$

The last conditions is an expression of the fact that $\dot{x}^2 + \dot{y}^2 = 1$, which is to say that the squared length of the tangent to the curve has unit length. This condition demands that

$$\frac{A^2}{\mu^2} = 1 \ .$$

It follws that $\mu = \pm A$. Thus there are *two* solutions, one corresponding to $\mu = A$, which gives an arc whose length increases with angle in a clockwise fashion; the other corresponds to $\mu = -A$, which gives an arc whose length increases with angle in a counter-clockwise fashion as the curve proceeds from point 1 to point 2.

Furthermore, it is evident that the Langrange multiplier μ has a a definite geometrical interpretation:

$$\mu = \text{radius of the circular arc} \ .$$

Lecture 5

2.2.3 Variational Problems with Finite Constraints

Aside from isoperimetric variation problems, i.e. those having a set of integral constraints as subsidiary conditions, another important class is exemplified by the following:
Find the curve $\{y(x), z(x)\}$ through points $(x_1, y_1, z_1), (x_2, y_2, z_2)$ such that

$$J = \int F(x, y, z, y', z')dx = \text{extremum}$$

subject to the "finite" constraint

$$G(x, y, z) = 0 \ .$$

Given that this problem has a solution, we consider the first variation of J due to deviations $\delta y(x), \delta z(x)$ of the variants $y + \delta y$ and $z + \delta z$ away from the functions $y(x)$ and $z(x)$:

$$\begin{aligned} \delta J \ &= \int_{x_1}^{x_2} \left[\left(F_y - \frac{d}{dx} F_{y'} \right) \delta y(x) + \left(F_z - \frac{d}{dx} F_{z'} \right) \delta z(x) \right] dx \\ &\quad + F_{y'} \delta y|_{x_1}^{x_2} + F_{z'} \delta z|_{x_1}^{x_2} \end{aligned}$$

The fact that the variant $\{y(x) + \delta y(x), \ z(x) + \delta z(x)\}$ is required to lie on the constraint manifold $G = 0$, i.e., it must satisfy

$$G(x, y + \delta y, z + \delta z) = 0 \ ,$$

implies that the variations δy and δz are not independent. They must satisfy

$$\frac{\partial G}{\partial y}\delta y(x) + \frac{\partial G}{\partial z}\delta z(x) = 0$$

so that

$$\delta z = -\frac{G_y}{G_z}\delta y.$$

Insert this into the expression for the first variation and obtain

$$\delta J = \int_{x_1}^{x_2} \left\{ (F_y - \frac{d}{dx}F_{y'}) - (F_z - \frac{d}{dx}F_{z'})\frac{G_y}{G_z} \right\} \delta y dx$$
$$+ F_{y'}\delta y|_{x_1}^{x_2} - F_{z'}\frac{G_y}{G_z}\delta y\bigg|_{x_1}^{x_2}$$

The endpoint terms are zero because all variants pass through the same pair of end points (x_1, y_1, z_1) and (x_2, y_2, z_2) so that $\delta y(x_1) = \delta y(x_2) = 0$.
The variations

$$\delta y = \sum \alpha_i h_i(x)$$

are arbitrary. Consequently, $\delta J = 0$ implies that for each x one has

$$\frac{F_y - \frac{d}{dx}F_{y'}}{G_y} = \frac{F_z - \frac{d}{dx}F_{z'}}{G_z}$$

Let us designate this common function by $\lambda(x)$. Thus one obtains two Euler equations,

$$\boxed{F_y - \frac{d}{dx}F_{y'} - \lambda(x)G_y = 0 \text{ and } F_z - \frac{d}{dx}F_{z'} - \lambda(x)G_z = 0}.$$

These equations together with

$$\boxed{G(x, y(x), z(x)) = 0}$$

determine the three functions

$$y(x), z(x), \text{ and } \lambda(x).$$

Even though $\lambda = \lambda(x)$ not a constant, it still is referred to a "Lagrange multiplier."

2.3 Variable End Point Problem

It is not necessary that the endpoints of a variational problem be fixed. Consider as an example the Brachistochrone problem (Lecture 3) in which the second end point, instead of being fixed, is required to lie somewhere on the vertical line. Which of the curves will allow a sliding bead to reach its destination in the least amount of time?

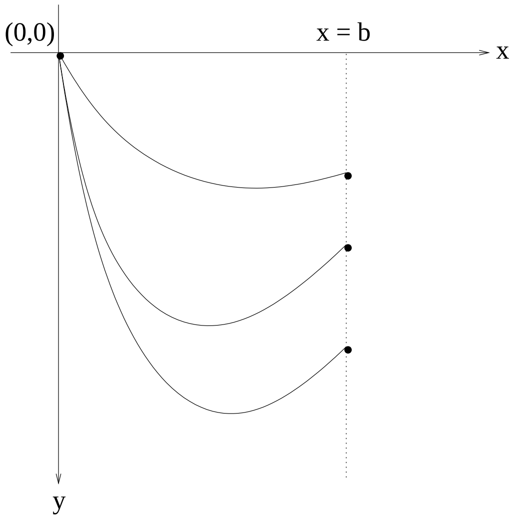


Figure 2.5: Brachistochrone trial curves with different end points.

The general problem is this: Find that curve $y(x)$, which starts at $x = a$ and ends at $x = b$, for which

$$J = \int_a^b F(x, y, y')dx + \psi(y(b)) - \varphi(y(a)) = \text{ extremum}, \tag{2.28}$$

where F, ψ, and φ are given.

The variational principles we have considered so far can be said to have had *Dirichelet* boundary conditions. By contrast, the present variational principle can be said to have just the opposite type, namely "natural" or "free" boundary conditions.

Having intergrated by parts, one finds that the first variation of the functional J is

$$\delta J = \int_a^b [F_y - \frac{d}{dx}F_{y'}]\delta f(x)dx + \left(\frac{\partial F}{\partial y'} + \frac{\partial \psi}{\partial y}\right)\delta f(x)\Big|^{x=b} - \left(\frac{\partial F}{\partial y'} + \frac{\partial \varphi}{\partial y}\right)\delta f(x)\Big|^{x=a}$$

The variations δf fall into two main classes:

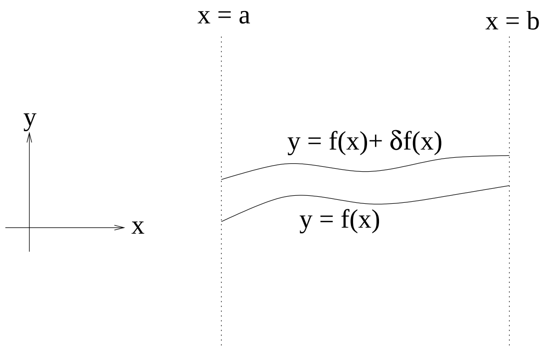


Figure 2.6: Trial curves with variable starting points and end points.

class 1: $\delta f : \delta f(a) = 0 \;\; \delta f(b) = 0$

class 2: $\delta f : \delta f(a) \neq 0 \;\; \delta f(b) \neq 0$

We know that $J =$ extremum $\Rightarrow \delta J = 0$ for all variations δf .
First we let δf be of class 1. In that case there are no endpoint terms in δJ.
Consequently, $\delta J = 0 \Rightarrow F_y - \frac{d}{dx}(F_{y'}) = 0$.
Thus we are left with

$$\delta J = \left(F_{y'} + \frac{\partial \psi}{\partial y}\right)\delta f(x)\Big|^{x=b} - \left(F_{y'} + \frac{\partial \varphi}{\partial y}\right)\delta f(x)\Big|^{x=a}.$$

Next we let δf be of class 2. Then

$$\delta J = 0 \Rightarrow \left(F_{y'} + \frac{\partial \psi}{\partial y}\right)\Big|^{x=b} = 0 \text{ and } \left(F_{y'} + \frac{\partial \varphi}{\partial y}\right)\Big|^{x=a} = 0.$$

These are the optimal boundary conditions associated for the given variational principle, Eq.(2.28), with variable endpoints as expressed by the boundary terms ψ and ϕ.

2.4 Extremum Principle at a Moment of Time Symmetry

The extremum principle with side conditions and with variable end points is illustrated by the following example.

Consider a vibrating string imbedded in an elastic medium and with a spring attached to each end. The tension of the string, the coefficient of restitution of the elastic medium, and the spring constants are properties which determine the temporal evolution, i.e. the vibrations, of the system. These properties we assume to be independent of time. Consequently, the system is said to be time-invariant. One can show that for such a time-invariant system its total energy,

$$\text{T.E.} = \text{Kinetic energy} + \text{Potential energy}$$

is independent of time, i.e. the total energy of a time invariant system is conserved.

In general the vibratory motion of the system is such that its kinetic energy is never zero. Moreover, there exist certain types of motion, namely those in which the vibrations are characterized by a single frequency. Such states of motions are called *normal modes.* They have the property that at some instant of time the system is momentarily stationary. At that moment the system's velocity is zero, its kinetic energy is zero and all of its energy is in the form of potential energy,

$$\text{T.E.} = 0 + \text{P.E. (at a moment of time symmetry)}$$

We inquire as to the amplitude profile of the system at this moment of time symmetry. The amplitude profile which nature allows is governed by a extremum principle.

Consider the total energy of the system at a moment of time symmetry ($\frac{\partial y}{\partial t} = 0$)

$$\begin{aligned} \text{T.E.} = J[y] &= \frac{1}{2}\int_a^b \left[T(x)\left(\frac{dy}{dx}\right)^2 + \kappa(x)y^2 \right] dx \\ &+ \frac{1}{2}k_1\left(y(a)-y_1\right)^2 + \frac{1}{2}k_2\left(y(b)-y_2\right)^2 . \end{aligned}$$

This is the potential energy of the system. It consists of three parts:

1. The "stretch" energy

$$\int_a^b T\sqrt{(dx)^2+(dy)^2}-\int_a^b T dx \approx \frac{1}{2}\int_a^b T\left(\frac{dy}{dx}\right)^2 dx$$

   Here $T = T(x)$ is the (position dependent) string tension.

2. The displacement energy due to the string being displaced in an elastic medium with a local restoration ("Hook's") coefficient

$$\kappa(x)dx$$

   for each incremental interval dx.

3. The potential energy stored in the two springs at $x = a$ and $x = b$. Their spring constants are k_1 and k_2.

The extremum principle governing the amplitude profile at a moment of time symmetry is given by

$$J[y] = \text{extremum} \tag{2.29}$$

subject to

$$K[y] = \frac{1}{2}\int_a^b \rho(x)y^2 dx = 1. \tag{2.30}$$

Here $\rho(x)$ is the linear mass density of the string.
With the help of the Lagrange multiplier λ, this principle reads

$$J[y] - \lambda K[y] = \text{extremum}.$$

Nota bene. It turns out that this principle is equivalent to

$$\frac{J[y]}{K[y]} = \text{extremum}.$$

This is Rayleigh's variational principle and we shall consider it later.

The Euler equation for either variational principle yield the same differential equation, namely

$$-\frac{d}{dx}T(x)\frac{dy}{dx} + (K(x) - \lambda\rho(x))y = 0 \tag{2.31}$$

This is the result of what we called “class 1” variations on page 51.

By contrast,“class 2” variations, which do not vanish at $x = a$ yield

$$T(x)y'(x)\Big|^{a} + k_1(y(b) - y_1) = 0$$

or

$$\frac{T(a)}{k_1}y'(a) + y(a) = y_1 \tag{2.32}$$

Similarly at $x = b$ one has

$$\frac{T(b)}{k_2}y'(b) + y(b) = y_2 \tag{2.33}$$

Equations 2.32 and 2.33 are known as the inhomogeneous mixed Dirichelet Neumann boundary conditions for the vibrating string. If we let the springs become infinitely stiff, i.e. let $k_1 \to \infty$ and $k_2 \to \infty$, then the problem reduces to a Dirichelet boundary value problem where the end point amplitude has the fixed values

$$y(a) = y_1$$
$$y(b) = y_2$$

We conclude that the Sturm-Liouville boundary value problem, Equations 2.31, 2.32 and 2.33 is a consequence of the variational principle,

$$\text{P.E.} - \lambda \int_a^b \frac{1}{2}\rho(x)y^2\,dx = \text{extremum.} \tag{2.34}$$

Remark 1. In vibrational problems the Lagrange multiplier λ, the eigenvalue of the S-L problem, refers to the squared vibrational angular frequency,

$$\lambda \equiv \omega^2 = (2\pi \text{ frequency})^2,$$

of the system, and the kinetic energy (K.E.) of the system vibrating at this frequency is

$$\frac{1}{2}\omega^2 \int_a^b \rho(x)y^2 dx.$$

Consequently, the amplitude profile of such a normal mode is governed by the variational principle

$$\text{P.E.} - \text{K.E.} = \text{extremum}.$$

We shall see later that the "least action principle"

$$\int_{t_1}^{t_2} (\text{P.E. - K.E.})\, dt = \text{extremum}$$

yields not only the amplitude profile at any instant of time, but also the evolution of this profile with time. Evidently, equation (2.34), the least energy problem with its amplitude constraint is a specialization of the least action principle.
Remark 2. This principle can be extended to two and three dimensions where the vibration can be those of a drum, solid, liquid sphere (vibrating sphere, vibrating neutron star, vibrating uranium nucleus), vibrating black hole, sound in cavity, vibrating bridge, etc.

Lecture 6

2.5 Generic Variable Endpoint Problem

There are variable end point problems in which the endpoint lies on a curve or a surface. For example,

1. what is the optimal curve from a point to a surface, or for that matter from one surface to another?

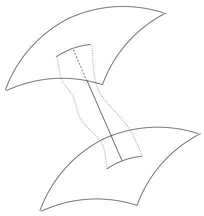


Figure 2.7: Optimal curve and its variants between two surfaces.

2. How does an extremal meet a terminating surface?

3. Given that a functional extends over two different media,

$$J = \int_a^c F(x, y, y')\, dx \ + \ \int_c^b F^*(x, y, y')\, dx \ ,$$

how must an extremal in $x < c$ be related to to its continuation in $c < x$, i.e. what are the junction conditions that prevail at the interface. For example, what path does light take as it passes from one medium to another?

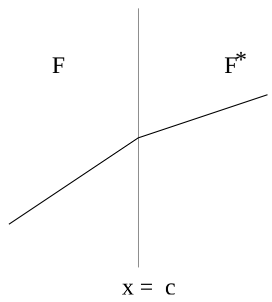


Figure 2.8: Curve crossing the boundary between two media characterized by F and F^*.

These questions and others can be answered within the framework of the following generic variable end point problem:
Find the curve $\{y_1(x), \ldots, y_n(x)\}$ in an $n + 1$ dimension space for which the variational integral

$$J[y_1, \ldots, y_n] = \int_{x_0}^{x_1} F(x, y, \ldots, y_n, y_n', \ldots, y_n')dx = \text{extremum}$$

subject to the constraint that the starting and end points $\{y_i(x_0)\}$ and $\{y_i(x_1)\}$ lie on two prescribed curves (or surfaces).
This variational principle asks us to consider changes in J not only due to variation in the curves,

$$y_i \to y_i + \delta y_i,$$

but also due to variations in the points

$$\begin{aligned} x_0 &\to x_0 + \delta x_0 \\ x_1 &\to x_1 + \delta x_0, \end{aligned}$$

where the curves start and finish. The determination of these changes in J is a task which one must do before the endpoint constraints are imposed.
To simplify discussion we consider the simplest case

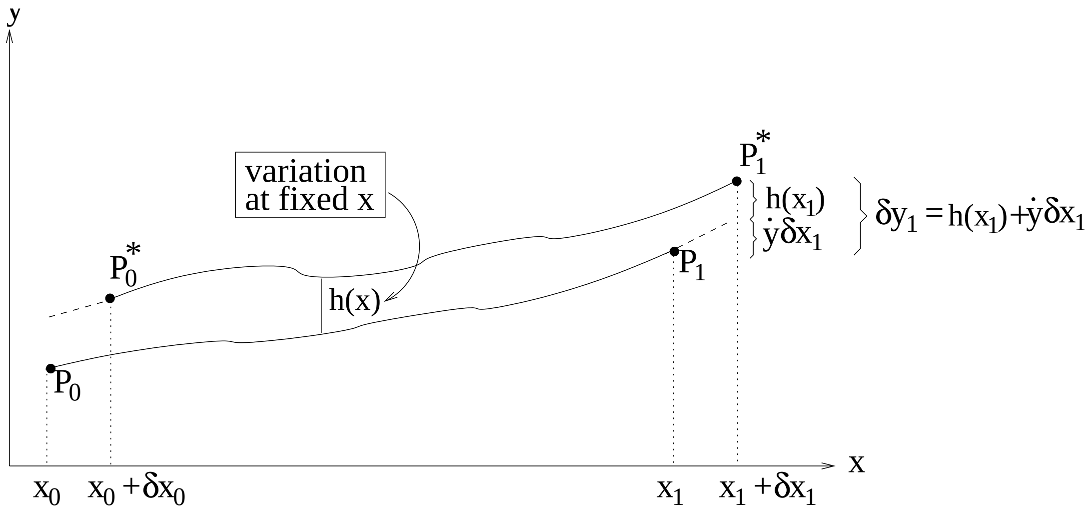


Figure 2.9: A curve and its generic variant.The total vertical variation δy_1 includes only the principal linear part. With that stipulation one has $h(x_1) = h(x_1 + \delta x_1)$.

$$J = \int_{x_0}^{x_1} F(x, y, y')dx = \text{extremum}$$

and calculate the first variation δJ allowing the endpoints to vary. Consider the curves

		starts	ends
curve:	$y(x)$	$P_0 = (x_0, y_0)$	$P_1 = (x_1, y_1)$
its variant:	$y^*(x)$	$P_0^* = (x_0 + \delta x_0, y_0 + \delta y_0)$	$P_1^* = (x_1 + \delta x_1, y_1 + \delta y_1)$

$$\boxed{h(x) = y^*(x) - y(x) = \text{difference at fixed } x}$$

2.5.1 General Variations in the Functional

The total variation in the functional to be extremized is

$$\begin{aligned}
\Delta J &= \int_{x_0+\delta x_0}^{x_1+\delta x_1} F(x, y+h, y'+h')dx - \int_{x_0}^{x_1} F(x, y, y')dx \\
&= \int_{x_0}^{x_1} [F(x, y+h, y'+h') - F(x, y, y')]dx \\
&+ \int_{x_0+\delta x_0}^{x_0} F(x, y+h, y'+h')dx + \int_{x_1}^{x_1+\delta x_1} F(x, y+h, y'+h')dx
\end{aligned}$$

whose principal linear part is

$$\Delta J = \int_{x_0}^{x_1} [F_y - \frac{d}{dx} F_{y'}] h(x) dx + F_{y'} h(x) \Big|_{x_0}^{x_1} + F(x, y, y') \delta x \Big|_{x_0}^{x_1} .$$

The last term is new. It is proportional to the shift in x :

$$\delta x \Big|^{x_0} = \delta x_0 \quad \text{and} \quad \delta x \Big|^{x_1} = \delta x_1 .$$

Referring to Figure 2.9, we express any type of end point variation in terms of the independent variations δy_i and δx_i :

$$\begin{aligned} h(x_0) = &\quad \delta y_0 - y'(x_0) \delta x_0 \\ h(x_1) = &\quad \delta y_1 - y'(x_1) \delta x_1 \end{aligned}$$

Thus one obtains for the total first variation

$$\begin{aligned} \Delta J \quad = &\quad \int_{x_0}^{x_1} \frac{\delta J}{\delta y} h(x) dx + F_{y'} \delta y \Big|_{x=x_0}^{x=x_1} \\ &\quad + (F - y' \frac{\partial F}{\partial y'}) \delta x \Big|_{x=x_0}^{x=x_1} \end{aligned} \tag{2.35}$$

We have two endpoint terms because in general both y and x are allowed to change at the boundary.
This is very useful formula; it has many consequences. It is used, among others,

(a) to establish boundary conditions

(b) to establish junction conditions

(c) as a starting point for the "canonical" formulation of mechanics

(d) as a starting point for the Hamilton-Jacobi formulation of mechanics.

2.5.2 Transversality Conditions

Let us now subject that first variation to the prescribed endpoint constraints. We require that the paths start and terminate on the two respective given curves

$$y = \varphi(x) \tag{2.36}$$

and

$$y = \psi(x) \tag{2.37}$$

or equivalently,

$$x = \Phi(y) \tag{2.38}$$

and

$$x = \Psi(y) \tag{2.39}$$

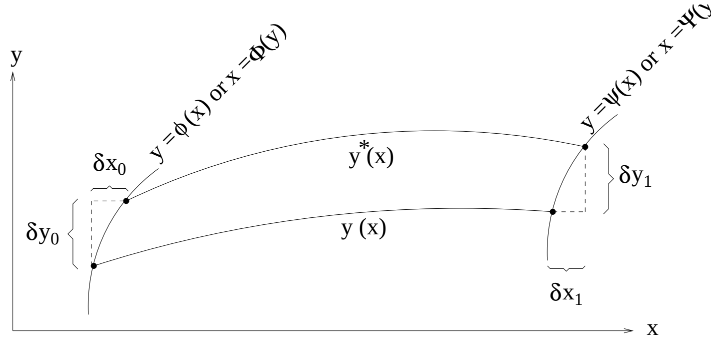


Figure 2.10: A curve and its variant with end point constraints.

In that case the vertical and horizontal variations in the end point are related by

$$\delta y_0 = \varphi'(x_0)\delta x_0 \quad \text{and} \quad \delta y_1 = \psi'(x_1)\delta x_1 \tag{2.40}$$

Consequently, the first variation in J is

$$\begin{aligned} \Delta J &= \int_{x_0}^{x_1} \frac{\delta J}{\delta y} h(x)dx + [F_{y'}\psi'(x) + (F - y'F_{y'})]_{x=x_1}\delta x_1 \\ &\quad -[F_{y'}\varphi'(x) + (F - y'F_{y'})]_{x=x_1}\delta x_0 \end{aligned}$$

If $y(x)$ is an optimal curve, then $\Delta J = 0$ for arbitrary variations $h(x)$ and arbitrary variations δx_i at the end points. The former implies that $y(x)$ satisfies the Euler equation. The latter implies

$$\begin{aligned} \Big[F + (\psi' - y')F_{y'}\Big]_{x=x_1} &= 0 \\ \Big[F + (\varphi' - y')F_{y'}\Big]_{x=x_0} &= 0, \end{aligned}$$

which are known as the transversality conditions.

2.6 Many Degrees of Freedom

One would like to generalize the variable end point problem from extremal paths in the plane to extremal path in a higher dimensional space. In that case the two curves (2.38) and (2.39) generalize to the two surfaces

$$x = \Phi(y_1, \ldots, y_n) \tag{2.41}$$

and

$$x = \Psi(y_1, \ldots, y_n) \tag{2.42}$$

for the loci of points where the extremal path starts and finishes.

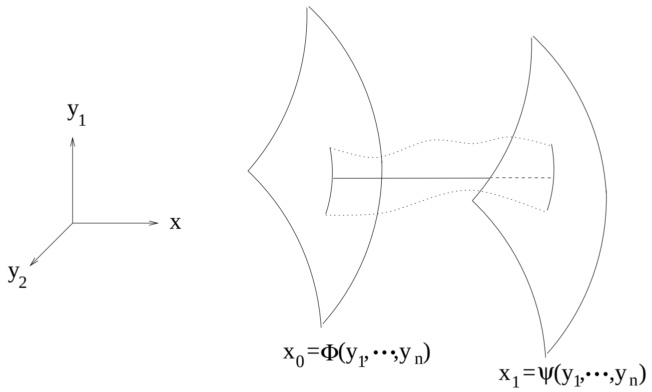


Figure 2.11: A curve and its variant between surfaces.

The corresponding generalization for the total first variation consists of replacing the equation (2.35) with

$$\begin{aligned} \Delta J &= \int_{x_0}^{x_1} \sum_{i=1}^{n} \frac{\delta J}{\delta y_i(x)} \delta y_i(x) dx + \sum_{i=1}^{n} F_{y_i'} \delta y_i \Big|_{x_0}^{x_1} \\ &\quad + \left(F - \sum_{i=1}^{n} y_j' \frac{\partial F}{y_j'} \right) \delta x \Big|_{x_0}^{x_1} \end{aligned}$$

Conceptual unit economy demands that it is best not to single out one of the y_i's over all the others. Consequently, we express the variation in δx at in terms of the variations of all the δy_i's in the surfaces (2.41) and (2.42),

$$\begin{aligned} \delta x_0 &= \sum_{i=1}^{n} \tfrac{\partial \Phi}{\partial y_i} \delta y_i \\ \delta x_1 &= \sum_{i=1}^{n} \tfrac{\partial \Psi}{\partial y_i} \delta y_i . \end{aligned}$$

With the help of these expressions the transversality conditions at these starting and terminating surface are given by

$$\left[F_{y'_i} + \frac{\partial \Phi}{\partial y_i}\left(F - \sum_{j=1}^{n} y'_j \frac{\partial F}{\partial y'_j}\right)\right]_{x=x_0} = 0, \qquad i = 1, \ldots, n$$

and analogous condition at $x = x_1$, where the gradient of the surface is $\frac{\partial \Psi}{\partial y_i}$.

Example. (Fermat's Principle of Least Time).
Consider Fermat's principle of least time for a light beam propagating through a medium whose refractive index is $n(x, y, z)$. Let the light beam start at (x_1, y_1, z_1) and finish at some point x_2, y_2, z_2 on the surface

$$x = \Psi(y, z)$$

such that the path minimizes the time of travel. The quantity to be extemized is

$$\begin{aligned} J[y, z] &= \frac{1}{c}\int_1^2 n(x, y, z)\sqrt{dx^2 + dy^2 + dz^2} \\ &= \frac{1}{c}\int_{x_1}^{x_2} n(x, y, z)\sqrt{1 + y'^2 + z'^2}\,dx \end{aligned}$$

One applies the transversality conditions to the variational integrand. We have

$$\begin{aligned} F &= n(x, y, z)\sqrt{1 + y'^2 + z'^2} \\ \frac{\partial F}{\partial y'} &= \frac{ny'}{\sqrt{1 + y'^2 + z'^2}}; \qquad \frac{\partial F}{\partial z'} = \frac{nz'}{\sqrt{1 + y'^2 + z'^2}} \\ F - y'F_{y'} - z'F_{z'} &= n\sqrt{1 + y'^2 + z'^2} - \frac{y'^2 + z'^2}{\sqrt{1 + y'^2 + z'^2}} \\ &= \frac{n}{\sqrt{1 + y'^2 + z'^2}} \end{aligned}$$

There are two transversality conditions at $x = x_2$. They are

$$\frac{ny'}{\sqrt{1 + y'^2 + z'^2}} + \frac{\partial \Psi}{\partial y}\frac{n}{\sqrt{1 + y'^2 + z'^2}} = 0$$

and

$$\frac{nz'}{\sqrt{1+y'^2+z'^2}}+\frac{\partial\Psi}{\partial z}\frac{n}{\sqrt{1+y'^2+z'^2}}=0\ .$$

Thus after cancellation we obtain

$$(y',z')=(-\frac{\partial\Psi}{\partial y},-\frac{\partial\Psi}{\partial z})$$

or

$$(x',y',z')=(1,-\frac{\partial\Psi}{\partial y},-\frac{\partial\Psi}{\partial z})=\vec{\nabla}g$$

The left hand side is $\frac{d\vec{x}}{dx}$, the tangent to the light beam. the right hand side is the gradient of the function $g(x,y,z)=x-\Psi(y,z)$. This gradient is $\perp$ to the surface. Thus the transversality condition says that the optimal light beam is the one which strikes the surface perpendicularly.

2.6.1 Junction Conditions

Suppose one wished to find an optimal curve between two points in adjoining regions characterized by different variational integrands. Thus we consider an optimal curve which leaves one region characterized by F and enters another characterized by F^* so that

$$J=\int_a^c F(x,y,y')dx+\int_c^b F^*(x,y,y')dx=\text{extremum.}$$

The boundary between the two adjoining regions is $x=c.$, and it is depicted in Figure 2.8 on page 56. The extremum principle establishes a relationship between the properties of the curve in the two regions. To determine this relation we, as usual, consider the total first variation

$$\begin{aligned}\Delta J &= \int_a^c \frac{\delta J}{\delta y}\delta y(x)dx+\int_c^b \frac{\delta J}{\delta y}\delta y(x)dx\\ &\quad +F_{y'}\delta y\Big|_a^{c^-}+F^*_{y'}\delta y\Big|_{c^+}^b\\ &\quad +(F-y'\frac{\partial F}{\partial y'})\delta x\Big|_a^{c^-}+(F^*-y'\frac{\partial F^*}{\partial y'})\delta x\Big|_{c^+}^b\ .\end{aligned}$$

The fact that the curve is optimal implies

$$\frac{\delta J}{\delta y}=0.$$

The fact that the boundary between the two regions is $x = c$, implies that $\delta x(c) = 0$, i.e. it is fixed and given.
Consequently the extremal condition implies

$$0 = \frac{\Delta J}{\delta y(c)} = F_{y'}\Big|_{c^+}^{c^-}$$

$$\boxed{F_{y'}\Big|_{c^+} = F^*_{y'}\Big|_{c^-}} \text{ The Junction Condition}$$

Example.(Snell's Law).
Consider a light beam propagating from one medium where the refractive index is $n(x, y, z)$ to another where it is $n^*(x, y, z,)$. The variational principle is

$$J = \int_a^c n(x, y)\sqrt{1 + y'^2}dx + \int_c^b n^*(x, y)\sqrt{1 + y'^2}dx = \text{extremum!}$$

n=refraction index on the left, n*=refraction index on the right

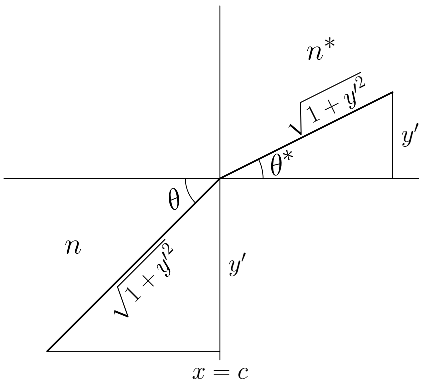


Figure 2.12: Ray refraction at the interface between two media having constant refractive indeces n and n^*.

The junction condition is

$$n(x, y)\frac{y'}{\sqrt{1 + y'^2}}\Bigg|_{c^-} = n^*(x, y)\frac{y'}{\sqrt{1 + y'^2}}\Bigg|_{c^+},$$

which has the geometrical interpretation

$$n \sin\theta = n^* \sin\theta^*,$$

which is Snell's law of refraction.