8

The calculus of variations

The calculus of variations, in its present form, provides a powerful method for the treatment of variational principles in physics and has become increasingly important in the development of modern physics. It is originated as a study of certain extremum (maximum and minimum) problems not treatable by elementary calculus. To see this more precisely let us consider the following integral whose integrand is a function of x, y, and of the first derivative $y'(x) = dy/dx$:

$$I = \int_{x_1}^{x_2} f\{y(x), y'(x); x\} dx, \tag{8.1}$$

where the semicolon in f separates the independent variable x from the dependent variable $y(x)$ and its derivative $y'(x)$. For what function $y(x)$ is the value of the integral I a maximum or a minimum? This is the basic problem of the calculus of variations.

The quantity f depends on the functional form of the dependent variable $y(x)$ and is called the functional which is considered as given, the limits of integration are also given. It is also understood that $y = y_1$ at $x = x_1$, $y = y_2$ at $x = x_2$. In contrast with the simple extreme-value problem of differential calculus, the function $y(x)$ is not known here, but is to be varied until an extreme value of the integral I is found. By this we mean that if $y(x)$ is a curve which gives to I a minimum value, then any neighboring curve will make I increase.

We can make the definition of a neighboring curve clear by giving $y(x)$ a parametric representation:

$$y(\varepsilon, x) = y(0, x) + \varepsilon\eta(x), \tag{8.2}$$

where $\eta(x)$ is an arbitrary function which has a continuous first derivative and ε is a small arbitrary parameter. In order for the curve (8.2) to pass through (x_1, y_1) and (x_2, y_2), we require that $\eta(x_1) = \eta(x_2) = 0$ (see Fig. 8.1). Now the integral I

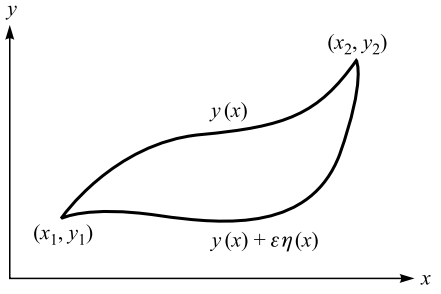

Figure 8.1.

also becomes a function of the parameter ε

$$I(\varepsilon) = \int_{x_1}^{x_2} f\{y(\varepsilon, x), y'(\varepsilon, x); x\}dx. \tag{8.3}$$

We then require that $y(x) = y(0, x)$ makes the integral I an extreme, that is, the integral $I(\varepsilon)$ has an extreme value for $\varepsilon = 0$:

$$I(\varepsilon) = \int_{x_1}^{x_2} f\{y(\varepsilon, x), y'(\varepsilon, x); x\}dx = \text{extremum for } \varepsilon = 0.$$

This gives us a very simple method of determining the extreme value of the integral I. The necessary condition is

$$\left.\frac{dI}{d\varepsilon}\right|_{\varepsilon=0} = 0 \tag{8.4}$$

for all functions $\eta(x)$. The sufficient conditions are quite involved and we shall not pursue them. The interested reader is referred to mathematical texts on the calculus of variations.

The problem of the extreme-value of an integral occurs very often in geometry and physics. The simplest example is provided by the problem of determining the shortest curve (or distance) between two given points. In a plane, this is the straight line. But if the two given points lie on a given arbitrary surface, then the analytic equation of this curve, which is called a geodesic, is found by solution of the above extreme-value problem.

The Euler–Lagrange equation

In order to find the required curve $y(x)$ we carry out the indicated differentiation in the extremum condition (8.4):

$$\frac{\partial I}{\partial \varepsilon} = \frac{\partial}{\partial \varepsilon}\int_{x_1}^{x_2} f\{y(\varepsilon, x), y'(\varepsilon, x); x\}dx$$
$$= \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \varepsilon} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \varepsilon}\right)dx, \tag{8.5}$$

where we have employed the fact that the limits of integration are fixed, so the differential operation affects only the integrand. From Eq. (8.2) we have

$$\frac{\partial y}{\partial \varepsilon} = \eta(x) \quad \text{and} \quad \frac{\partial y'}{\partial \varepsilon} = \frac{d\eta}{dx}.$$

Substituting these into Eq. (8.5) we obtain

$$\frac{\partial I}{\partial \varepsilon} = \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}\eta(x) + \frac{\partial f}{\partial y'}\frac{d\eta}{dx}\right)dx. \tag{8.6}$$

Using integration by parts, the second term on the right hand side becomes

$$\int_{x_1}^{x_2}\frac{\partial f}{\partial y'}\frac{d\eta}{dx}dx = \frac{\partial f}{\partial y'}\eta(x)\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2}\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\eta(x)dx.$$

The integrated term on the right hand side vanishes because $\eta(x_1) = \eta(x_2) = 0$ and Eq. (8.6) becomes

$$\frac{\partial I}{\partial \varepsilon} = \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \varepsilon} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\frac{\partial y}{\partial \varepsilon}\right)dx$$
$$= \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)\eta(x)dx. \tag{8.7}$$

Note that $\partial f/\partial y$ and $\partial f/\partial y'$ are still functions of ε. However, when $\varepsilon = 0, y(\varepsilon, x) = y(x)$ and the dependence on ε disappears.

Then $(\partial I/\partial \varepsilon)|_{\varepsilon=0}$ vanishes, and since $\eta(x)$ is an arbitrary function, the integrand in Eq. (8.7) must vanish for $\varepsilon = 0$:

$$\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} = 0. \tag{8.8}$$

Eq. (8.8) is known as the Euler–Lagrange equation; it is a necessary but not sufficient condition that the integral I have an extreme value. Thus, the solution of the Euler–Lagrange equation may not yield the minimizing curve. Ordinarily we must verify whether or not this solution yields the curve that actually minimizes the integral, but frequently physical or geometrical considerations enable us to tell whether the curve so obtained makes the integral a minimum or a maximum. The Euler–Lagrange equation can be written in the form (Problem 8.2)

$$\frac{d}{dx}\left(f - y'\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial x} = 0. \tag{8.8a}$$

This is often called the second form of the Euler–Lagrange equation. If f does not involve x explicitly, it can be integrated to yield

$$f - y' \frac{\partial f}{\partial y'} = c, \tag{8.8b}$$

where c is an integration constant.

The Euler–Lagrange equation can be extended to the case in which f is a functional of several dependent variables:

$$f = f\{y_1(x), y_1'(x), y_2(x), y_2'(x), \ldots; x\}.$$

Then, in analogy with Eq. (8.2), we now have

$$y_i(\varepsilon, x) = y_i(0, x) + \varepsilon \eta_i(x), \qquad i = 1, 2, \ldots, n.$$

The development proceeds in an exactly analogous manner, with the result

$$\frac{\partial I}{\partial \varepsilon} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y_i} - \frac{d}{dx} \left(\frac{\partial f}{\partial y_i'} \right) \right) \eta_i(x) dx.$$

Since the individual variations, that is, the $\eta_i(x)$, are all independent, the vanishing of the above equation when evaluated at $\varepsilon = 0$ requires the separate vanishing of each expression in the brackets:

$$\frac{d}{dx} \frac{\partial f}{\partial y_i'} - \frac{\partial f}{\partial y_i} = 0, \qquad i = 1, 2, \ldots, n. \tag{8.9}$$

Example 8.1

The brachistochrone problem: Historically, the brachistochrone problem was the first to be treated by the method of the calculus of variations (first solved by Johann Bernoulli in 1696). As shown in Fig. 8.2, a particle is constrained to move in a gravitational field starting at rest from some point P_1 to some lower

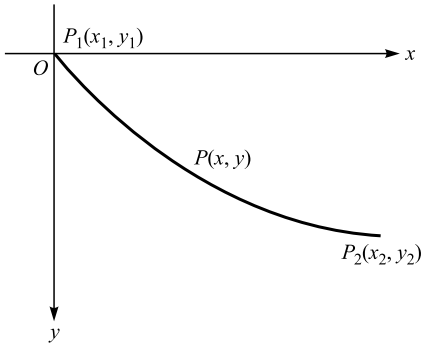

Figure 8.2

https://doi.org/10.1017/CBO9780511755781.009 Published online by Cambridge University Press

point P_2. Find the shape of the path such that the particle goes from P_1 to P_2 in the least time. (The word brachistochrone was derived from the Greek *brachistos* (shortest) and *chronos* (time).)

Solution: If O and *P* are not very far apart, the gravitational field is constant, and if we ignore the possibility of friction, then the total energy of the particle is conserved:

$$0 + mgy_1 = \frac{1}{2}m\left(\frac{ds}{dt}\right)^2 + mg(y_1 - y),$$

where the left hand side is the sum of the kinetic energy and the potential energy of the particle at point P_1, and the right hand side refers to point $P(x, y)$. Solving for ds/dt:

$$ds/dt = \sqrt{2gy}.$$

Thus the time required for the particle to move from P_1 to P_2 is

$$t = \int_{P_1}^{P_2} dt = \int_{P_1}^{P_2} \frac{ds}{\sqrt{2gy}}\,.$$

The line element ds can be expressed as

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + y'^2}dx, \qquad y' = dy/dx;$$

thus, we have

$$t = \int_{P_1}^{P_2} dt = \int_{P_1}^{P_2} \frac{ds}{\sqrt{2gy}} = \frac{1}{\sqrt{2g}}\int_0^{x_2} \frac{\sqrt{1 + y'^2}}{\sqrt{y}}dx.$$

We now apply the Euler–Lagrange equation to find the shape of the path for the particle to go from P_1 to P_2 in the least time. The constant does not affect the final equation and the functional f may be identified as

$$f = \sqrt{1 + y'^2}/\sqrt{y},$$

which does not involve x explicitly. Using Problem 8.2(b), we find

$$f - y'\frac{\partial f}{\partial y'} = \frac{\sqrt{1 + y'^2}}{\sqrt{y}} - y'\left[\frac{y'}{\sqrt{1 + y'^2}\sqrt{y}}\right] = c,$$

which simplifies to

$$\sqrt{1 + y'^2}\sqrt{y} = 1/c.$$

https://doi.org/10.1017/CBO9780511755781.009 Published online by Cambridge University Press

Letting $1/c = \sqrt{a}$ and solving for y' gives

$$y' = \frac{dy}{dx} = \sqrt{\frac{a-y}{y}},$$

and solving for dx and integrating we obtain

$$\int dx = \int \sqrt{\frac{y}{a-y}}\, dy.$$

We then let

$$y = a\sin^2\theta = \frac{a}{2}(1-\cos 2\theta)$$

which leads to

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

Thus the parametric equation of the path is given by

$$x = b(1-\cos\phi), \qquad y = b(\phi - \sin\phi) + k,$$

where $b = a/2, \phi = 2\theta$. The path passes through the origin so we have $k = 0$ and

$$x = b(1-\cos\phi), \qquad y = b(\phi - \sin\phi).$$

The constant b is determined from the condition that the particle passes through $P_2(x_2, y_2)$.

The required path is a cycloid and is the path of a fixed point P' on a circle of radius b as it rolls along the x-axis (Fig. 8.3).

A line that represents the shortest path between any two points on some surface is called a geodesic. On a flat surface, the geodesic is a straight line. It is easy to show that, on a sphere, the geodesic is a great circle; we leave this as an exercise for the reader (Problem 8.3).

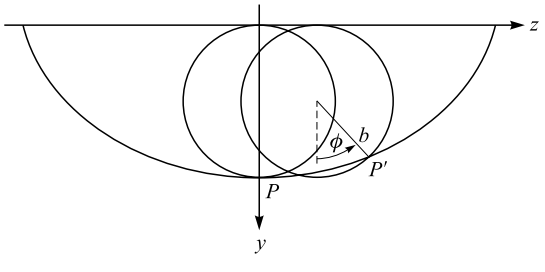

Figure 8.3.

Variational problems with constraints

In certain problems we seek a minimum or maximum value of the integral (8.1)

$$I = \int_{x_1}^{x_2} f\{y(x), y'(x); x\}dx \tag{8.1}$$

subject to the condition that another integral

$$J = \int_{x_1}^{x_2} g\{y(x), y'(x); x\}dx \tag{8.10}$$

has a known constant value. A simple problem of this sort is the problem of determining the curve of a given perimeter which encloses the largest area, or finding the shape of a chain of fixed length which minimizes the potential energy.

In this case we can use the method of Lagrange multipliers which is based on the following theorem:

> *The problem of the stationary value of $F(x, y)$ subject to the condition $G(x, y) = const.$ is equivalent to the problem of stationary values, without constraint, of $F + \lambda G$ for some constant λ, provided either $\partial G/\partial x$ or $\partial G/\partial y$ does not vanish at the critical point.*

The constant λ is called a Lagrange multiplier and the method is known as the method of Lagrange multipliers. To see the ideas behind this theorem, let us assume that $G(x, y) = 0$ defines y as a unique function of x, say, $y = g(x)$, having a continuous derivative $g'(x)$. Then

$$F(x, y) = F[x, g(x)]$$

and its maximum or minimum can be found by setting the derivative with respect to x equal to zero:

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0 \qquad \text{or} \qquad F_x + F_y g'(x) = 0. \tag{8.11}$$

We also have

$$G[x, g(x)] = 0,$$

from which we find

$$\frac{\partial G}{\partial x} + \frac{\partial G}{\partial y}\frac{dy}{dx} = 0 \qquad \text{or} \qquad G_x + G_y g'(x) = 0. \tag{8.12}$$

Eliminating $g'(x)$ between Eq. (8.11) and Eq. (8.12) we obtain

$$F_x - (F_y/G_y)G_x = 0, \tag{8.13}$$

provided $G_y = \partial G/\partial y \neq 0$. Defining $\lambda = -F_y/G_y$ or

$$F_y + \lambda G_y = \frac{\partial F}{\partial y} + \lambda \frac{\partial G}{\partial y} = 0, \tag{8.14}$$

Eq. (8.13) becomes

$$F_x + \lambda G_x = \frac{\partial F}{\partial x} + \lambda \frac{\partial G}{\partial x} = 0. \tag{8.15}$$

If we define

$$H(x, y) = F(x, y) + \lambda G(x, y),$$

then Eqs. (8.14) and (8.15) become

$$\partial H(x, y)/\partial x = 0, \qquad H(x, y)/\partial y = 0,$$

and this is the basic idea behind the method of Lagrange multipliers.

It is natural to attempt to solve the problem $I =$ minimum subject to the condition $J =$ constant by the method of Lagrange multipliers. We construct the integral

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

and consider its free extremum. This implies that the function $y(x)$ that makes the value of the integral an extremum must satisfy the equation

$$\frac{d}{dx}\frac{\partial(F + \lambda G)}{\partial y'} \frac{\partial(F + \lambda G)}{\partial y} = 0 \tag{8.16}$$

or

$$\left[\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y}\right] + \lambda\left[\frac{d}{dx}\left(\frac{\partial G}{\partial y'}\right) - \frac{\partial G}{\partial y}\right] = 0. \tag{8.16a}$$

Example 8.2

Isoperimetric problem: Find that curve C having the given perimeter l that encloses the largest area.

Solution: The area bounded by C can be expressed as

$$A = \tfrac{1}{2}\int_C (xdy - ydx) = \tfrac{1}{2}\int_C (xy' - y)dx$$

and the length of the curve C is

$$s = \int_C \sqrt{1 + y'^2}dx = l.$$

https://doi.org/10.1017/CBO9780511755781.009 Published online by Cambridge University Press

Then the function H is

$$H = \int_C [\tfrac{1}{2}(xy' - y) + \lambda\sqrt{1 + y'^2}]dx$$

and the Euler–Lagrange equation gives

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

or

$$\frac{\lambda y'}{\sqrt{1 + y'^2}} = -x + c_1.$$

Solving for y', we get

$$y' = \frac{dy}{dx} = \pm\frac{x - c_1}{\sqrt{\lambda^2 - (x - c_1)^2}},$$

which on integrating gives

$$y - c_2 = \pm\sqrt{\lambda^2 - (x - c_1)^2}$$

or

$$(x - c_1)^2 + (y - c_2)^2 = \lambda^2, \quad \text{a circle.}$$

Hamilton's principle and Lagrange's equation of motion

One of the most important applications of the calculus of variations is in classical mechanics. In this case, the functional f in Eq. (8.1) is taken to be the Lagrangian L of a dynamical system. For a conservative system, the Lagrangian L is defined as the difference of kinetic and potential energies of the system:

$$L = T - V,$$

where time t is the independent variable and the generalized coordinates $q_i(t)$ are the dependent variables. What do we mean by generalized coordinates? Any convenient set of parameters or quantities that can be used to specify the configuration (or state) of the system can be assumed to be generalized coordinates; therefore they need not be geometrical quantities, such as distances or angles. In suitable circumstances, for example, they could be electric currents.

Eq. (8.1) now takes the form that is known as the action (or the action integral)

$$I = \int_{t_1}^{t_2} L(q_i(t), \dot{q}_i(t); t)dt, \qquad \dot{q} = dq/dt \tag{8.17}$$

and Eq. (8.4) becomes

$$\delta I = \frac{\partial I}{\partial \varepsilon}\bigg|_{\varepsilon=0} d\varepsilon = \delta \int_{t_1}^{t_2} L(q_i(t), \qquad \dot{q}_i(t); t)dt = 0, \tag{8.18}$$

where $q_i(t)$, and hence $\dot{q}_i(t)$, is to be varied subject to $\delta q_i(t_1) = \delta q_i(t_2) = 0$. Equation (8.18) is a mathematical statement of Hamilton's principle of classical mechanics. In this variational approach to mechanics, the Lagrangian L is given, and $q_i(t)$ taken on the prescribed values at t_1 and t_2, but may be arbitrarily varied for values of t between t_1 and t_2.

In words, Hamilton's principle states that for a conservative dynamical system, the motion of the system from its position in configuration space at time t_1 to its position at time t_2 follows a path for which the action integral (8.17) has a stationary value. The resulting Euler–Lagrange equations are known as the Lagrange equations of motion:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0. \tag{8.19}$$

These Lagrange equations can be derived from Newton's equations of motion (that is, the second law written in differential equation form) and Newton's equations can be derived from Lagrange's equations. Thus they are 'equivalent.' However, Hamilton's principle can be applied to a wide range of physical phenomena, particularly those involving fields, with which Newton's equations are not usually associated. Therefore, Hamilton's principle is considered to be more fundamental than Newton's equations and is often introduced as a basic postulate from which various formulations of classical dynamics are derived.

Example 8.3

Electric oscillations: As an illustration of the generality of Lagrangian dynamics, we consider its application to an LC circuit (inductive–capacitive circuit) as shown in Fig. 8.4. At some instant of time the charge on the capacitor C is $Q(t)$ and the current flowing through the inductor is $I(t) = \dot{Q}(t)$. The voltage drop around the

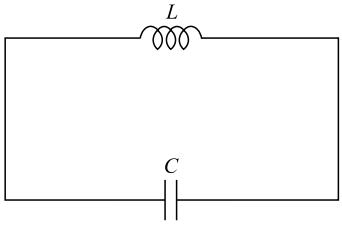

Figure 8.4. LC circuit.

circuit is, according to Kirchhoff's law

$$L\frac{dI}{dt}+\frac{1}{C}\int I(t)dt=0$$

or in terms of Q

$$L\ddot{Q}+\frac{1}{C}Q=0.$$

This equation is of exactly the same form as that for a simple mechanical oscillator:

$$m\ddot{x}+kx=0.$$

If the electric circuit also contains a resistor R, Kirchhoff's law then gives

$$L\ddot{Q}+R\dot{Q}+\frac{1}{C}Q=0,$$

which is of exactly the same form as that for a damped oscillator

$$m\ddot{x}+b\dot{x}+kx=0,$$

where b is the damping constant.

By comparing the corresponding terms in these equations, an analogy between mechanical and electric quantities can be established:

x	displacement	Q	charge (generalized coordinate)
$\dot{x}$	velocity	$\dot{Q}=I$	electric current
m	mass	L	inductance
$1/k$	$k=$ spring constant	C	capacitance
b	damping constant	R	electric resistance
$\frac{1}{2}m\dot{x}^2$	kinetic energy	$\frac{1}{2}L\dot{Q}^2$	energy stored in inductance
$\frac{1}{2}mx^2$	potential energy	$\frac{1}{2}Q^2/C$	energy stored in capacitance

If we recognize in the beginning that the charge Q in the circuit plays the role of a generalized coordinate, and $T=\frac{1}{2}L\dot{Q}^2$ and $V=\frac{1}{2}Q^2/C$, then the Langrangian L of the system is

$$L=T-V=\tfrac{1}{2}L\dot{Q}^2-\tfrac{1}{2}Q^2/C$$

and the Lagrange equation gives

$$L\ddot{Q}+\frac{1}{C}Q=0,$$

the same equation as given by Kirchhoff's law.

Example 8.4
A bead of mass m slides freely on a frictionless wire of radius b that rotates in a horizontal plane about a point on the circular wire with a constant angular velocity ω. Show that the bead oscillates as a pendulum of length $l = g/\omega^2$.

Solution: The circular wire rotates in the xy plane about the point O, as shown in Fig. 8.5. The rotation is in the counterclockwise direction, C is the center of the circular wire, and the angles θ and ϕ are as indicated. The wire rotates with an angular velocity ω, so $\phi = \omega t$. Now the coordinates x and y of the bead are given by

$$x = b\cos\omega t + b\cos(\theta + \omega t),$$
$$y = b\sin\omega t + b\sin(\theta + \omega t),$$

and the generalized coordinate is θ. The potential energy of the bead (in a horizontal plane) can be taken to be zero, while its kinetic energy is

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \tfrac{1}{2}mb^2[\omega^2 + (\dot{\theta} + \omega)^2 + 2\omega(\dot{\theta} + \omega)\cos\theta],$$

which is also the Lagrangian of the bead. Inserting this into Lagrange's equation

$$\frac{d}{d\theta}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = 0$$

we obtain, after some simplifications,

$$\ddot{\theta} + \omega^2 \sin\theta = 0.$$

Comparing this equation with Lagrange's equation for a simple pendulum of length l

$$\ddot{\theta} + (g/l)\sin\theta = 0$$

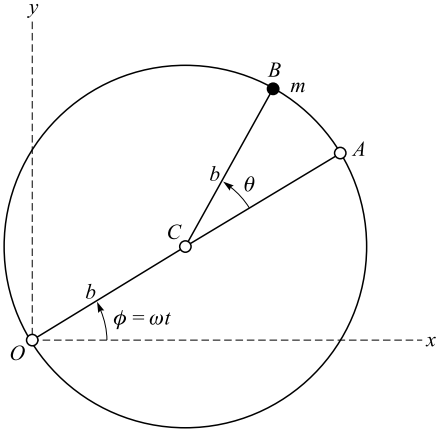

Figure 8.5.

https://doi.org/10.1017/CBO9780511755781.009 Published online by Cambridge University Press

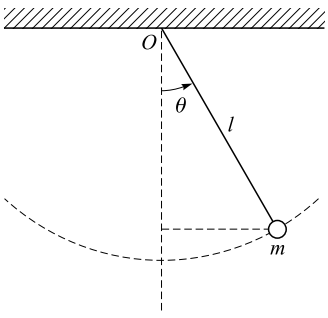

Figure 8.6.

(Fig. 8.6) we see that the bead oscillates about the line OA like a pendulum of length $l = g/\omega^2$.

Rayleigh–Ritz method

Hamilton's principle views the motion of a dynamical system as a whole and involves a search for the path in configuration space that yields a stationary value for the action integral (8.17):

$$\delta I = \delta \int_{t_1}^{t_2} L(q_i(t), \dot{q}_i(t); t)dt = 0, \tag{8.18}$$

with $\delta q_i(t_1) = \delta q_i(t_2) = 0$. Ordinarily it is used as a variational method to obtain Lagrange's and Hamilton's equations of motion, so we do not often think of it as a computational tool. But in other areas of physics variational formulations are used in a much more active way. For example, the variational method for determining the approximate ground-state energies in quantum mechanics is very well known. We now use the Rayleigh–Ritz method to illustrate that Hamilton's principle can be used as computational device in classical mechanics. The Rayleigh–Ritz method is a procedure for obtaining approximate solutions of problems expressed in variational form directly from the variational equation.

The Lagrangian is a function of the generalized coordinates qs and their time derivatives $\dot{q}$s. The basic idea of the approximation method is to guess a solution for the qs that depends on time and a number of parameters. The parameters are then adjusted so that Hamilton's principle is satisfied. The Rayleigh–Ritz method takes a special form for the trial solution. A complete set of functions $\{f_i(t)\}$ is chosen and the solution is assumed to be a linear combination of a finite number of these functions. The coefficients in this linear combination are the parameters that are chosen to satisfy Hamilton's principle (8.18). Since the variations of the qs

must vanish at the endpoints of the integral, the variations of the parameter must be so chosen that this condition is satisfied.

To summarize, suppose a given system can be described by the action integral

$$I = \int_{t_1}^{t_2} L(q_i(t), \dot{q}_i(t); t)dt, \qquad \dot{q} = dq/dt.$$

The Rayleigh–Ritz method requires the selection of a trial solution, ideally in the form

$$q = \sum_{i=1}^{n} a_i f_i(t), \tag{8.20}$$

which satisfies the appropriate conditions at both the initial and final times, and where *a*s are undetermined constant coefficients and the *f*s are arbitrarily chosen functions. This trial solution is substituted into the action integral I and integration is performed so that we obtain an expression for the integral I in terms of the coefficients. The integral I is then made 'stationary' with respect to the assumed solution by requiring that

$$\frac{\partial I}{\partial a_i} = 0 \tag{8.21}$$

after which the resulting set of n simultaneous equations is solved for the values of the coefficients a_i. To illustrate this method, we apply it to two simple examples.

Example 8.5

A simple harmonic oscillator consists of a mass M attached to a spring of force constant k. As a trial function we take the displacement x as a function t in the form

$$x(t) = \sum_{n=1}^{\infty} A_n \sin n\omega t.$$

For the boundary conditions we have $x = 0, t = 0$, and $x = 0, t = 2\pi/\omega$. Then the potential energy and the kinetic energy are given by, respectively,

$$V = \tfrac{1}{2}kx^2 = \tfrac{1}{2}k \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_n A_m \sin n\omega t \sin m\omega t,$$

$$T = \tfrac{1}{2}M\dot{x}^2 = \tfrac{1}{2}M\omega^2 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_n A_m nm \cos n\omega t \cos m\omega t.$$

The action I has the form

$$I = \int_0^{2\pi/\omega} Ldt = \int_0^{2\pi/\omega} (T-V)dt = \frac{\pi}{2\omega} \sum_{n=1}^{\infty} (kA_n^2 - Mn^2 A_n^2 \omega^2).$$

In order to satisfy Hamilton's principle we must choose the values of A_n so as to make I an extremum:

$$\frac{dI}{dA_n} = (k - n^2\omega^2 M)A_n = 0.$$

The solution that meets the physics of the problem is

$$A_1 = 0, \quad \omega^2 = k/M; \qquad \text{or} \qquad \tau = (2\pi/\omega)^{1/2} = 2\pi(M/k)^{1/2},$$

$$A_n = 0, \quad \text{for} \quad n = 2, 3, \text{etc.}$$

Example 8.6

As a second example, we consider a bead of mass M sliding freely along a wire shaped in the form of a parabola along the vertical axis and of the form $y = ax^2$. In this case, we have

$$L = T - V = \tfrac{1}{2}M(\dot{x}^2 + \dot{y}^2) - Mgy = \tfrac{1}{2}M(1 + 4a^2x^2)\dot{x}^2 - Mgy.$$

We assume

$$x = A \sin \omega t$$

to be an approximate value for the displacement x, and then the action integral becomes

$$I = \int_0^{2\pi/\omega} L dt = \int_0^{2\pi/\omega} (T - V) dt = A^2 \left\{ \frac{\omega^2(1 + a^2A^2)}{2} - ga \right\} \frac{M\pi}{\omega}.$$

The extremum condition, $dI/dA = 0$, gives an approximate ω:

$$\omega = \frac{\sqrt{2ga}}{1 + a^2A^2},$$

and the approximate period is

$$\tau = \frac{2\pi(1 + a^2A^2)}{\sqrt{2ga}}.$$

The Rayleigh–Ritz method discussed in this section is a special case of the general Rayleigh–Ritz methods that are designed for finding approximate solutions of boundary-value problems by use of varitional principles, for example, the eigenvalues and eigenfunctions of the Sturm–Liouville systems.

Hamilton's principle and canonical equations of motion

Newton first formulated classical mechanics in the seventeenth century and it is known as Newtonian mechanics. The essential physics involved in Newtonian

mechanics is contained in Newton's three laws of motion, with the second law serving as the equation of motion. Classical mechanics has since been reformulated in a few different forms: the Lagrange, the Hamilton, and the Hamilton–Jacobi formalisms, to name just a few.

The essential physics of Lagrangian dynamics is contained in the Lagrange function L of the dynamical system and Lagrange's equations (the equations of motion). The Lagrangian L is defined in terms of independent generalized coordinates q_i and the corresponding generalized velocity $\dot{q}_i$. In Hamiltonian dynamics, we describe the state of a system by Hamilton's function (or the Hamiltonian) H defined in terms of the generalized coordinates q_i and the corresponding generalized momenta p_i, and the equations of motion are given by Hamilton's equations or canonical equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \qquad i = 1, 2, \ldots, n. \tag{8.22}$$

Hamilton's equations of motion can be derived from Hamilton's principle. Before doing so, we have to define the generalized momentum and the Hamiltonian. The generalized momentum p_i corresponding to q_i is defined as

$$p_i = \frac{\partial L}{\partial q_i} \tag{8.23}$$

and the Hamiltonian of the system is defined by

$$H = \sum_i p_i \dot{q}_i - L. \tag{8.24}$$

Even though $\dot{q}_i$ explicitly appears in the defining expression (8.24), H is a function of the generalized coordinates q_i, the generalized momenta p_i, and the time t, because the defining expression (8.23) can be solved explicitly for the $\dot{q}_i$s in terms of p_i, q_i, and t. The qs and ps are now treated the same: $H = H(q_i, p_i, t)$. Just as with the configuration space spanned by the n independent qs, we can imagine a space of $2n$ dimensions spanned by the $2n$ variables $q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n$. Such a space is called phase space, and is particularly useful in both statistical mechanics and the study of non-linear oscillations. The evolution of a representative point in this space is determined by Hamilton's equations.

We are ready to deduce Hamilton's equation from Hamilton's principle. The original Hamilton's principle refers to paths in configuration space, so in order to extend the principle to phase space, we must modify it such that the integrand of the action I is a function of both the generalized coordinates and momenta and their derivatives. The action I can then be evaluated over the paths of the system

point in phase space. To do this, first we solve Eq. (8.24) for L

$$L = \sum_i p_i \dot{q}_i - H$$

and then substitute L into Eq. (8.18) and we obtain

$$\delta I = \delta \int_{t_1}^{t_2} \left(\sum_i p_i \dot{q}_i - H(p, q, t) \right) dt = 0, \tag{8.25}$$

where $q_I(t)$ is still varied subject to $\delta q_i(t_1) = \delta q_i(t_2) = 0$, but p_i is varied without such end-point restrictions.

Carrying out the variation, we obtain

$$\int_{t_1}^{t_2} \sum_i \left(p_i \delta \dot{q}_i + \dot{q}_i \delta p_i - \frac{\partial H}{\partial q_i} \delta q_i - \frac{\partial H}{\partial p_i} \delta p_i \right) dt = 0, \tag{8.26}$$

where the $\delta \dot{q}$s are related to the δqs by the relation

$$\delta \dot{q}_i = \frac{d}{dt} \delta q_i. \tag{8.27}$$

Now we integrate the term $p_i \delta \dot{q}_i dt$ by parts. Using Eq. (8.27) and the endpoint conditions on δq_i, we find that

$$\begin{aligned} \int_{t_1}^{t_2} \sum_i p_i \delta \dot{q}_i dt &= \int_{t_1}^{t_2} \sum_i p_i \frac{d}{dt} \delta q_i dt \\ &= \int_{t_1}^{t_2} \sum_i \frac{d}{dt} p_i \delta q_i dt - \int_{t_1}^{t_2} \sum_i \dot{p}_i \delta q_i dt \\ &= p_i \delta q_i \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \sum_i \dot{p}_i \delta q_i dt \\ &= - \int_{t_1}^{t_2} \sum_i \dot{p}_i \delta q_i dt. \end{aligned}$$

Substituting this back into Eq. (8.26), we obtain

$$\int_{t_1}^{t_2} \sum_i \left[\left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i - \left(\dot{p}_i + \frac{\partial H}{\partial q_i} \right) \delta q_i \right] dt = 0. \tag{8.28}$$

Since we view Hamilton's principle as a variational principle in phase space, both the δqs and the δps are arbitrary, the coefficients of δq_i and δp_i in Eq. (8.28) must vanish separately, which results in the $2n$ Hamilton's equations (8.22).

Example 8.7

Obtain Hamilton's equations of motion for a one-dimensional harmonic oscillator.

Solution: We have

$$T = \tfrac{1}{2}m\dot{x}^2, \qquad V = \tfrac{1}{2}Kx^2,$$

$$p = \frac{\partial L}{\partial \dot{x}} = \frac{\partial T}{\partial \dot{x}} = m\dot{x}, \qquad \dot{x} = \frac{p}{m}.$$

Hence

$$H = p\dot{x} - L = T + V = \frac{1}{2m}p^2 + \frac{1}{2}Kx^2.$$

Hamilton's equations

$$\dot{x} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial x}$$

then read

$$\dot{x} = \frac{p}{m}, \qquad \dot{p} = -Kx.$$

Using the first equation, the second can be written

$$\frac{d}{dt}(m\dot{x}) = -Kx \quad \text{or} \quad m\ddot{x} + Kx = 0$$

which is the familiar equation of the harmonic oscillator.

The modified Hamilton's principle and the Hamilton–Jacobi equation

The Hamilton–Jacobi equation is the cornerstone of a general method of integrating equations of motion. Before the advent of modern quantum theory, Bohr's atomic theory was treated in terms of Hamilton–Jacobi theory. It also plays an important role in optics as well as in canonical perturbation theory. In classical mechanics books, the Hamilton–Jacobi equation is often obtained via canonical transformations. We want to show that the Hamilton–Jacobi equation can also be obtained directly from Hamilton's principle, or, a modified Hamilton's principle.

In formulating Hamilton's principle, we have considered the action

$$I = \int_{t_1}^{t_2} L(q_i(t), \dot{q}_i(t); t)dt, \qquad \dot{q} = dq/dt,$$

taken along a path between two given positions $q_i(t_1)$ and $q_i(t_2)$ which the dynamical system occupies at given instants t_1 and t_2. In varying the action, we compare the values of the action for neighboring paths with fixed ends, that is, with $\delta q_i(t_1) = \delta q_i(t_2) = 0$. Only one of these paths corresponds to the true dynamical path for which the action has its extremum value.

We now consider another aspect of the concept of action, by regarding I as a quantity characterizing the motion along the true path, and comparing the value

of I for paths having a common beginning at $q_i(t_1)$, but passing through different points at time t_2. In other words we consider the action I for the true path as a function of the coordinates at the upper limit of integration:

$$I = I(q_i, t),$$

where q_i are the coordinates of the final position of the system, and t is the instant when this position is reached.

If $q_i(t_2)$ are the coordinates of the final position of the system reached at time t_2, the coordinates of a point near the point $q_i(t_2)$ can be written as $q_i(t_1) + \delta q_i$, where δq_i is a small quantity. The action for the trajectory bringing the system to the point $q_i(t_1) + \delta q_i$ differs from the action for the trajectory bringing the system to the point $q_i(t_2)$ by the quantity

$$\delta I = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right] dt, \tag{8.29}$$

where δq_i is the difference between the values of q_i taken for both paths at the same instant t; similarly, $\delta \dot{q}_i$ is the difference between the values of $\dot{q}_i$ at the instant t.

We now integrate the second term on the right hand side of Eq. (8.25) by parts:

$$\begin{aligned} \int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i dt &= \frac{\partial L}{\partial \dot{q}_i} \delta q_i - \int_{t_1}^{t_2} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i dt \\ &= p_i \delta q_i - \int_{t_1}^{t_2} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i dt, \end{aligned} \tag{8.30}$$

where we have used the fact that the starting points of both paths coincide, hence $\delta q_i(t_1) = 0$; the quantity $\delta q_i(t_2)$ is now written as just δq_i. Substituting Eq. (8.30) into Eq. (8.29), we obtain

$$\delta I = \sum_i p_i \delta q_i + \int_{t_1}^{t_2} \sum_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta q_i dt. \tag{8.31}$$

Since the true path satisfies Lagrange's equations of motion, the integrand and, consequently, the integral itself vanish. We have thus obtained the following value for the increment of the action I due to the change in the coordinates of the final position of the system by δq_i (at a constant time of motion):

$$\delta I = \sum_i p_i \delta q_i, \tag{8.32}$$

from which it follows that

$$\frac{\partial I}{\partial q_i} = p_i, \tag{8.33}$$

that is, the partial derivatives of the action with respect to the generalized coordinates equal the corresponding generalized momenta.

The action I may similarly be regarded as an explicit function of time, by considering paths starting from a given point $q_i(1)$ at a given instant t_1, ending at a given point $q_i(2)$ at various times $t_2 = t$:

$$I = I(q_i, t).$$

Then the total time derivative of I is

$$\frac{dI}{dt} = \frac{\partial I}{\partial t} + \sum_i \frac{\partial I}{\partial q_i}\dot{q}_i = \frac{\partial I}{\partial t} + \sum_i p_i \dot{q}_i. \tag{8.34}$$

From the definition of the action, we have $dI/dt = L$. Substituting this into Eq. (8.34), we obtain

$$\frac{\partial I}{\partial t} = L - \sum_i p_i \dot{q}_i = -H$$

or

$$\frac{\partial I}{\partial t} + H(q_i, p_i, t) = 0. \tag{8.35}$$

Replacing the momenta p_i in the Hamiltonian H by $\partial I/\partial q_i$ as given by Eq. (8.33), we obtain the Hamilton–Jacobi equation

$$H(q_i, \partial I/\partial q_i, t) + \frac{\partial I}{\partial t} = 0. \tag{8.36}$$

For a conservative system with stationary constraints, the time is not contained explicitly in Hamiltonian H, and $H = E$ (the total energy of the system). Consequently, according to Eq. (8.35), the dependence of action I on time t is expressed by the term $-Et$. Therefore, the action breaks up into two terms, one of which depends only on q_i, and the other only on t:

$$I(q_i, t) = I_o(q_i) - Et. \tag{8.37}$$

The function $I_o(q_i)$ is sometimes called the contracted action, and the Hamilton–Jacobi equation (8.36) reduces to

$$H(q_i, \partial I_o/\partial q_i) = E. \tag{8.38}$$

Example 8.8

To illustrate the method of Hamilton–Jacobi, let us consider the motion of an electron of charge $-e$ revolving about an atomic nucleus of charge Ze (Fig. 8.7). As the mass M of the nucleus is much greater than the mass m of the electron, we may consider the nucleus to remain stationary without making any very appreciable error. This is a central force motion and so its motion lies entirely in one plane (see *Classical Mechanics*, by Tai L. Chow, John Wiley, 1995). Employing

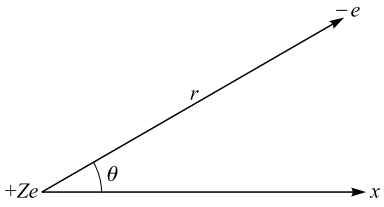

Figure 8.7.

polar coordinates r and θ in the plane of motion to specify the position of the electron relative to the nucleus, the kinetic and potential energies are, respectively,

$$T = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2), \qquad V = -\frac{Ze^2}{r}.$$

Then

$$L = T - V = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{Ze^2}{r}$$

and

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \qquad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}.$$

The Hamiltonian H is

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) - \frac{Ze^2}{r}.$$

Replacing p_r and p_θ in the Hamiltonian by $\partial I/\partial r$ and $\partial I/\partial\theta$, respectively, we obtain, by Eq. (8.36), the Hamilton–Jacobi equation

$$\frac{1}{2m}\left[\left(\frac{\partial I}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial I}{\partial \theta}\right)^2\right] - \frac{Ze^2}{r} + \frac{\partial I}{\partial t} = 0.$$

Variational problems with several independent variables

The functional f in Eq. (8.1) contains only one independent variable, but very often f may contain several independent variables. Let us now extend the theory to this case of several independent variables:

$$I = \iiint_V f\{u, u_x, u_y, u_z; x, y, z)\, dxdydz, \tag{8.39}$$

where V is assumed to be a bounded volume in space with prescribed values of $u(x, y, z)$ at its boundary S; $u_x = \partial u/\partial x$, and so on. Now, the variational problem

is to find the function $u(x, y, z)$ for which I is stationary with respect to small changes in the functional form $u(x, y, z)$.

Generalizing Eq. (8.2), we now let

$$u(x, y, z, \varepsilon) = u(x, y, z, 0) + \varepsilon\eta(x, y, z), \tag{8.40}$$

where $\eta(x, y, z)$ is an arbitrary well-behaved (that is, differentiable) function which vanishes at the boundary S. Then we have, from Eq. (8.40),

$$u_x(x, y, z, \varepsilon) = u_x(x, y, z, 0) + \varepsilon\eta_x,$$

and similar expressions for u_y, u_z; and

$$\left.\frac{\partial I}{\partial \varepsilon}\right|_{\varepsilon=0} = \iiint_V \left(\frac{\partial f}{\partial u}\eta + \frac{\partial f}{\partial u_x}\eta_x + \frac{\partial f}{\partial u_y}\eta_y + \frac{\partial f}{\partial u_z}\eta_z\right) dxdydz = 0.$$

We next integrate each of the terms $(\partial f/\partial u_i)\eta_i$ using 'integration by parts' and the integrated terms vanish at the boundary as required. After some simplifications, we finally obtain

$$\iiint_V \left\{\frac{\partial f}{\partial u} - \frac{\partial}{\partial x}\frac{\partial f}{\partial u_x} - \frac{\partial}{\partial y}\frac{\partial f}{\partial u_y} - \frac{\partial}{\partial z}\frac{\partial f}{\partial u_z}\right\}\eta(x, y, z)dxdydz = 0.$$

Again, since $\eta(x, y, z)$ is arbitrary, the term in the braces may be set equal to zero, and we obtain the Euler–Lagrange equation:

$$\frac{\partial f}{\partial u} - \frac{\partial}{\partial x}\frac{\partial f}{\partial u_x} - \frac{\partial}{\partial y}\frac{\partial f}{\partial u_y} - \frac{\partial}{\partial z}\frac{\partial f}{\partial u_z} = 0. \tag{8.41}$$

Note that in Eq. (8.41) $\partial/\partial x$ is a partial derivative, in that y and z are constant. *But* $\partial/\partial x$ is also a total derivative in that it acts on implicit x dependence and on explicit x dependence:

$$\frac{\partial}{\partial x}\frac{\partial f}{\partial u_x} = \frac{\partial^2 f}{\partial x \partial u_x} + \frac{\partial^2 f}{\partial u \partial u_x}u_x + \frac{\partial^2 f}{\partial u_x^2} + \frac{\partial^2 f}{\partial u_y \partial u_x}u_{xy} + \frac{\partial^2 f}{\partial u_z \partial u_x}u_{xz}. \tag{8.42}$$

Example 8.9

The Schrödinger wave equation. The equations of motion of classical mechanics are the Euler–Lagrange differential equations of Hamilton's principle. Similarly, the Schrödinger equation, the basic equation of quantum mechanics, is also a Euler–Lagrange differential equation of a variational principle the form of which is, in the case of a system of N particles, the following

$$\delta \int L d\tau = 0, \tag{8.43}$$

with

$$L = \sum_{i=1}^{N} \frac{\hbar^2}{2m_i} \left(\frac{\partial \psi^*}{\partial x_i} \frac{\partial \psi}{\partial x_i} + \frac{\partial \psi^*}{\partial y_i} \frac{\partial \psi}{\partial y_i} + \frac{\partial \psi^*}{\partial z_i} \frac{\partial \psi}{\partial z_i} \right) + V \psi^* \psi \tag{8.44}$$

and the constraint

$$\int \psi^* \psi d\tau = 1, \tag{8.45}$$

where m_i is the mass of particle I, V is the potential energy of the system, and $d\tau$ is a volume element of the $3N$-dimensional space.

Condition (8.45) can be taken into consideration by introducing a Lagrangian multiplier $-E$:

$$\delta \int (L - E\psi^* \psi) d\tau = 0. \tag{8.46}$$

Performing the variation we obtain the Schrödinger equation for a system of N particles

$$\sum_{i=1}^{N} \frac{\hbar^2}{2m_i} \nabla_i^2 \psi + (E - V)\psi = 0, \tag{8.47}$$

where ∇_i^2 is the Laplace operator relating to particle i. Can you see that E is the energy parameter of the system? If we use the Hamiltonian operator $\hat{H}$, Eq. (8.47) can be written as

$$\hat{H}\psi = E\psi. \tag{8.48}$$

From this we obtain for E

$$E = \frac{\int \psi^* H \psi d\tau}{\int \psi^* \psi d\tau}. \tag{8.49}$$

Through partial integration we obtain

$$\int L d\tau = \int \psi^* H \psi d\tau$$

and thus the variational principle can be formulated in another way: $\delta \int \psi^* (H - E)\psi d\tau = 0$.

Problems

8.1 As a simple practice of using varied paths and the extremum condition, we consider the simple function $y(x) = x$ and the neighboring paths

https://doi.org/10.1017/CBO9780511755781.009 Published online by Cambridge University Press

$y(\varepsilon, x) = x + \varepsilon \sin x$. Draw these paths in the xy plane between the limits $x = 0$ and $x = 2\pi$ for $\varepsilon = 0$ for two different non-vanishing values of ε. If the integral $I(\varepsilon)$ is given by

$$I(\varepsilon) = \int_0^{2\pi} (dy/dx)^2 dx,$$

show that the value of $I(\varepsilon)$ is always greater than $I(0)$, no matter what value of ε (positive or negative) is chosen. This is just condition (8.4).

8.2 (*a*) Show that the Euler–Lagrange equation can be written in the form

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

This is often called the second form of the Euler–Lagrange equation.

(*b*) If f does not involve x explicitly, show that the Euler–Lagrange equation can be integrated to yield

$$f - y' \frac{\partial f}{\partial y'} = c,$$

where c is an integration constant.

8.3 As shown in Fig. 8.8, a curve C joining points (x_1, y_1) and (x_2, y_2) is revolved about the x-axis. Find the shape of the curve such that the surface thus generated is a minimum.

8.4 A geodesic is a line that represents the shortest distance between two points. Find the geodesic on the surface of a sphere.

8.5 Show that the geodesic on the surface of a right circular cylinder is a helix.

8.6 Find the shape of a heavy chain which minimizes the potential energy while the length of the chain is constant.

8.7 A wedge of mass M and angle α slides freely on a horizontal plane. A particle of mass m moves freely on the wedge. Determine the motion of the particle as well as that of the wedge (Fig. 8.9).

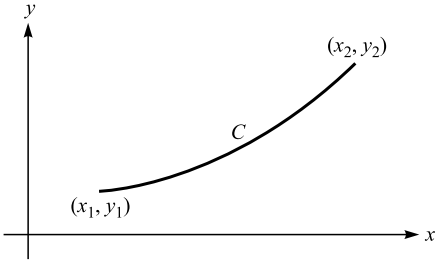

Figure 8.8.

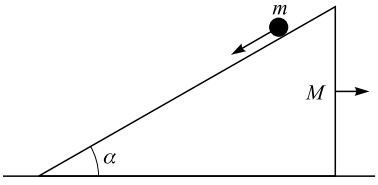

Figure 8.9.

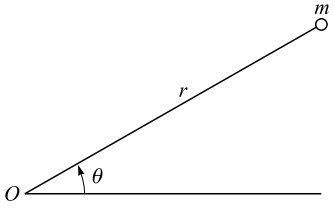

Figure 8.10.

8.8 Use the Rayleigh–Ritz method to analyze the forced oscillations of a harmonic oscillation:

$$m\ddot{x} + kx = F_0 \sin \omega t.$$

8.9 A particle of mass m is attracted to a fixed point O by an inverse square force $F_r = -k/r^2$ (Fig. 8.10). Find the canonical equations of motion.

8.10 Set up the Hamilton–Jacobi equation for the simple harmonic oscillator.