CHAPTER 2

Variational Principles and Lagrange's Equations

2.1 ■ HAMILTON'S PRINCIPLE

The derivation of Lagrange's equations presented in Chapter 1 started from a consideration of the instantaneous state of the system and small virtual displacements about the instantaneous state, i.e., from a "differential principle" such as D'Alembert's principle. It is also possible to obtain Lagrange's equations from a principle that considers the entire motion of the system between times t_1 and t_2, and small virtual variations of this motion from the actual motion. A principle of this nature is known as an "integral principle."

Before presenting the integral principle, the meaning attached to the phrase "motion of the system between times t_1 and t_2" must first be stated in more precise language. The instantaneous configuration of a system is described by the values of the n generalized coordinates $q_1, \ldots, q_n$, and corresponds to a particular point in a Cartesian hyperspace where the q's form the n coordinate axes. This n-dimensional space is therefore known as configuration space. As time goes on, the state of the system changes and the system point moves in configuration space tracing out a curve, described as "the path of motion of the system." The "motion of the system," as used above, then refers to the motion of the system point along this path in *configuration space*. Time can be considered formally as a parameter of the curve; to each point on the path there is associated one or more values of the time. Note that configuration space has no necessary connection with the physical three-dimensional space, just as the generalized coordinates are not necessarily position coordinates. The path of motion in configuration space has no resemblance to the path in space of any actual particle; each point on the path represents the *entire* system configuration at some given instant of time.

The integral *Hamilton's principle* describes the motion of those mechanical systems for which all forces (except the forces of constraint) are derivable from a generalized scalar potential that may be a function of the coordinates, velocities, and time. Such systems will be denoted as *monogenic*. Where the potential is an explicit function of position coordinates only, then a monogenic system is also conservative (cf. Section 1.2).

For monogenic systems, Hamilton's principle can be stated as

> *The motion of the system from time t_1 to time t_2 is such that the line integral (called the* action *or the* action integral*),*

$$I = \int_{t_1}^{t_2} L\,dt, \tag{2.1}$$

where $L = T - V$, has a stationary value for the actual path of the motion.

That is, out of all possible paths by which the system point could travel from its position at time t_1 to its position at time t_2, it will actually travel along that path for which the value of the integral (2.1) is stationary. By the term "stationary value" for a line integral, we mean that the integral along the given path has the same value to within first-order infinitesimals as that along all neighboring paths (i.e., those that differ from it by infinitesimal displacements). (Cf. Fig. 2.1.) The notion of a stationary value for a line integral thus corresponds in ordinary function theory to the vanishing of the first derivative.

We can summarize Hamilton's principle by saying that the motion is such that the *variation* of the line integral I for fixed t_1 and t_2 is zero:

$$\delta I = \delta \int_{t_1}^{t_2} L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)\,dt = 0. \tag{2.2}$$

Where the system constraints are holonomic, Hamilton's principle, Eq. (2.2), is both a necessary and sufficient condition for Lagrange's equations, Eqs. (1.57). Thus, it can be shown that Hamilton's principle follows directly from Lagrange's equations. Instead, however, we shall prove the converse, namely, that Lagrange's equations follow from Hamilton's principle, as being the more important theorem. That Hamilton's principle is a sufficient condition for deriving the equations of motion enables us to construct the mechanics of monogenic systems from Hamilton's principle as the basic postulate rather than Newton's laws of motion. Such a formulation has advantages; e.g., since the integral I is obviously invariant to the system of generalized coordinates used to express L, the equations of motion must always have the Lagrangian form no matter how the generalized coordinates

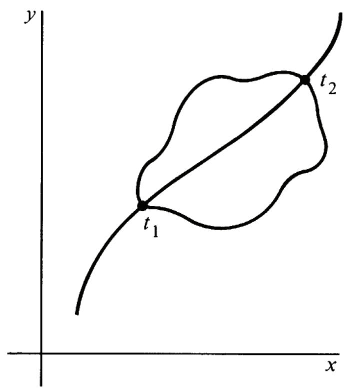

FIGURE 2.1 Path of the system point in configuration space.

are transformed. More important, the formulation in terms of a variational principle is the route that is generally followed when we try to describe apparently nonmechanical systems in the mathematical clothes of classical mechanics, as in the theory of fields.

2.2 ■ SOME TECHNIQUES OF THE CALCULUS OF VARIATIONS

Before demonstrating that Lagrange's equations do follow from (2.2), we must first examine the methods of the calculus of variations, for a chief problem of this calculus is to find the curve for which some given line integral has a stationary value. (See website for necessary comments.)

Consider first the problem in an essentially one-dimensional form: We have a function $f(y, \dot{y}, x)$ defined on a path $y = y(x)$ between two values x_1 and x_2, where $\dot{y}$ is the derivative of y with respect to x. We wish to find a particular path $y(x)$ such that the line integral J of the function f between x_1 and x_2,

$$\dot{y} \equiv \frac{dy}{dx},$$

$$J = \int_{x_1}^{x_2} f(y, \dot{y}, x)\, dx, \tag{2.3}$$

has a stationary value relative to paths differing infinitesimally from the correct function $y(x)$. The variable x here plays the role of the parameter t, and we consider only such varied paths for which $y(x_1) = y_1$, $y(x_2) = y_2$. (Cf. Fig. 2.2.) Note that Fig. 2.2 does *not* represent configuration space. In the one-dimensional configuration space, both the correct and varied paths are the segment of the straight line connecting y_1 and y_2; the paths differ only in the functional relation between y and x. The problem is one-dimensional, y is not a coordinate, it is a function of x.

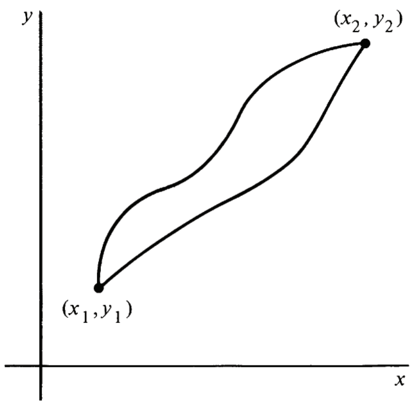

FIGURE 2.2 Varied paths of the function of $y(x)$ in the one-dimensional extremum problem.

We put the problem in a form that enables us to use the familiar apparatus of the differential calculus for finding the stationary points of a function. Since J must have a stationary value for the correct path relative to *any* neighboring path, the variation must be zero relative to *some* particular set of neighboring paths labeled by an infinitesimal parameter α. Such a set of paths might be denoted by $y(x, \alpha)$, with $y(x, 0)$ representing the correct path. For example, if we select any function $\eta(x)$ that vanishes at $x = x_1$ and $x = x_2$, then a possible set of varied paths is given by

$$y(x, \alpha) = y(x, 0) + \alpha\eta(x). \tag{2.4}$$

For simplicity, it is assumed that both the correct path $y(x)$ and the auxiliary function $\eta(x)$ are well-behaved functions—continuous and nonsingular between x_1 and x_2, with continuous first and second derivatives in the same interval. For any such parametric family of curves, J in Eq. (2.3) is also a function of α:

$$J(\alpha) = \int_{x_1}^{x_2} f\,(y(x, \alpha), \dot{y}(x, \alpha), x)\; dx, \tag{2.5}$$

and the condition for obtaining a stationary point is the familiar one that

$$\left(\frac{dJ}{d\alpha}\right)_{\alpha=0} = 0. \tag{2.6}$$

By the usual methods of differentiating under the integral sign, we find that

$$\frac{dJ}{d\alpha} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha}\right) dx. \tag{2.7}$$

Consider the second of these integrals:

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha}\, dx = \int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}}\frac{\partial^2 y}{\partial x\, \partial \alpha}\, dx.$$

Integrating by parts, the integral becomes

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}}\frac{\partial^2 y}{\partial x\, \partial \alpha}\, dx = \left.\frac{\partial f}{\partial \dot{y}}\frac{\partial y}{\partial \alpha}\right|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right)\frac{\partial y}{\partial \alpha}\, dx. \tag{2.8}$$

The conditions on all the varied curves are that they pass through the points (x_1, y_1), (x_2, y_2), and hence the partial derivative of y with respect to α at x_1 and x_2 must vanish. Therefore, the first term of (2.8) vanishes and Eq. (2.7) reduces to

$$\frac{dJ}{d\alpha} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}}\right)\frac{\partial y}{\partial \alpha}\, dx.$$

The condition for a stationary value, Eq. (2.6), is therefore equivalent to the equation

$$\int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial \dot{y}} \right) \left(\frac{\partial y}{\partial \alpha} \right)_0 dx = 0. \tag{2.9}$$

Now, the partial derivative of y with respect to α occurring in Eq. (2.9) is a function of x that is arbitrary except for continuity and end point conditions. For example, for the particular parametric family of varied paths given by Eq. (2.4), it is the arbitrary function $\eta(x)$. We can therefore apply to Eq (2.9) the so-called "fundamental lemma" of the calculus of variations, which says if

$$\int_{x_1}^{x_2} M(x)\eta(x)\, dx = 0 \tag{2.10}$$

for all arbitrary functions $\eta(x)$ continuous through the second derivative, then $M(x)$ must identically vanish in the interval (x_1, x_2). While a formal mathematical proof of the lemma can be found in texts on the calculus of variations, the validity of the lemma is easily seen intuitively. We can imagine constructing a function η that is positive in the immediate vicinity of any chosen point in the interval and zero everywhere else. Equation (2.10) can then hold only if $M(x)$ vanishes at that (arbitrarily) chosen point, which shows M must be zero throughout the interval. From Eq. (2.9) and the fundamental lemma, it therefore follows that J can have a stationary value only if

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial \dot{y}} \right) = 0. \tag{2.11}$$

The differential quantity,

$$\left(\frac{\partial y}{\partial \alpha} \right)_0 d\alpha \equiv \delta y, \tag{2.12}$$

represents the infinitesimal departure of the varied path from the correct path $y(x)$ at the point x and thus corresponds to the virtual displacement introduced in Chapter 1 (hence the notation δy). Similarly, the infinitesimal variation of J about the correct path can be designated

$$\left(\frac{dJ}{d\alpha} \right)_0 d\alpha \equiv \delta J. \tag{2.13}$$

The assertion that J is stationary for the correct path can thus be written

$$\delta J = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial \dot{y}} \right) \delta y\, dx = 0,$$

requiring that $y(x)$ satisfy the differential equation (2.11). The δ-notation, introduced through Eqs. (2.12) and (2.13), may be used as a convenient shorthand for treating the variation of integrals, remembering always that it stands for the manipulation of parametric families of varied paths such as Eq. (2.4).

Some simple examples of the application of Eq. (2.11) (which clearly resembles a Lagrange equation) may now be considered:

1. *Shortest distance between two points in a plane.* An element of length in a plane is

$$ds = \sqrt{dx^2 + dy^2}$$

and the total length of any curve going between points 1 and 2 is

$$I = \int_1^2 ds = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx.$$

The condition that the curve be the shortest path is that I be a minimum. This is an example of the extremum problem as expressed by Eq. (2.3), with

$$f = \sqrt{1 + \dot{y}^2}.$$

Substituting in (2.11) with

$$\frac{\partial f}{\partial y} = 0, \qquad \frac{\partial f}{\partial \dot{y}} = \frac{\dot{y}}{\sqrt{1 + \dot{y}^2}},$$

we have

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

or

$$\frac{\dot{y}}{\sqrt{1 + \dot{y}^2}} = c,$$

where c is constant. This solution can be valid only if

$$\dot{y} = a,$$

where a is a constant related to c by

$$a = \frac{c}{\sqrt{1 - c^2}}.$$

But this is clearly the equation of a straight line,

$$y = ax + b,$$

where b is another constant of integration. Strictly speaking, the straight line has only been proved to be an extremum path, but for this problem it is obviously also a minimum. The constants of integration, a and b, are determined by the condition that the curve pass through the two end points, (x_1, y_1), (x_2, y_2).

In a similar fashion we can obtain the shortest distance between two points on a sphere, by setting up the arc length on the surface of the sphere in terms of the angle coordinates of position on the sphere. In general, curves that give the shortest distance between two points on a given surface are called the *geodesics* of the surface. (See website about part 2 below.)

2. *Minimum surface of revolution.* Suppose we form a surface of revolution by taking some curve passing between two fixed end points (x_1, y_1) and (x_2, y_2) defining the xy plane, and revolving it about the y axis (cf. Fig. 2.3a). The problem then is to find that curve for which the surface area is a minimum. The area of a strip of the surface is $2\pi x\, ds = 2\pi x\sqrt{1+\dot{y}^2}\, dx$, and the total area is

$$2\pi \int_1^2 x\sqrt{1+\dot{y}^2}\, dx.$$

The extremum of this integral is again given by (2.11) where

$$f = x\sqrt{1+\dot{y}^2}$$

and

$$\frac{\partial f}{\partial y} = 0, \qquad \frac{\partial f}{\partial \dot{y}} = \frac{x\dot{y}}{\sqrt{1+\dot{y}^2}}.$$

Equation (2.11) becomes in this case

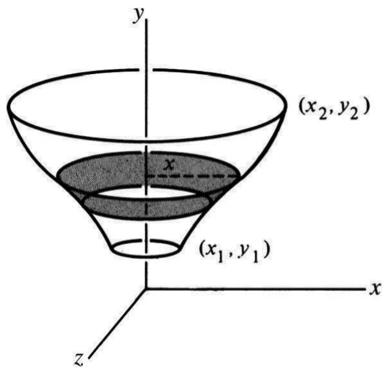

FIGURE 2.3a Minimum surface of revolution. Note that this figure is drawn for y_1 and y_2 having the same sign relative to the rotation axis. This is not assumed in the general solution.

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

or

$$\frac{x\dot{y}}{\sqrt{1+\dot{y}^2}}=a,$$

where a is some constant of integration clearly smaller than the minimum value of x. Squaring the above equation and factoring terms, we have

$$\dot{y}^2(x^2-a^2)=a^2,$$

or solving,

$$\frac{dy}{dx}=\frac{a}{\sqrt{x^2-a^2}}.$$

The general solution of this differential equation, in light of the nature of a, is

$$y=a\int\frac{dx}{\sqrt{x^2-a^2}}+b=a\operatorname{arc\,cosh}\frac{x}{a}+b$$

or

$$x=a\cosh\frac{y-b}{a},$$

which is the equation of a catenary. Again the two constants of integration, a and b, are determined in principle by the requirements that the curve pass through the two given end points, as shown in Fig. 2.3b.

Curves satisfying the preceding equation all scale as x/a and y/a with one independent parameter b/a. This suggests that when the solutions are examined in detail they turn out to be a great deal more complicated than these considerations

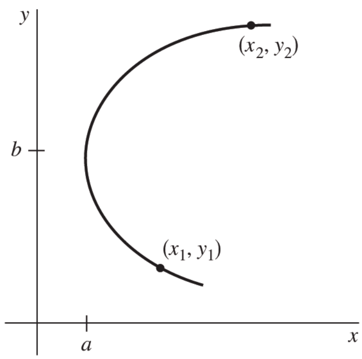

FIGURE 2.3b General catenary solution for minimum surface of revolution.

suggest. For some pairs of end points, unique constants of integration a and b can be found. But for other end points, it is possible to draw two different catenary curves through the end points, while for additional cases no possible values can be found for a and b. Further, recall that Eq. (2.11) represents a condition for finding curves $y(x)$ continuous through the second derivative that render the integral stationary. The catenary solutions therefore do not always represent minimum values, but may represent "inflection points" where the length of the curve is stationary but not minimum.

For certain combinations of end points (an example is x_1 and x_2 both positive and both much smaller than $y_2 - y_1$), the absolute minimum in the surface of revolution is provided (cf. Exercise 8) by a curve composed of straight line segments—from the first end point parallel to the x axis until the y axis is reached, then along the y axis until the point $(0, y_2)$ and then out in a straight line to the second end point corresponding to the area $\pi(x_1^2 + x_2^2)$. This curve results when $a = 0$, forcing either $x = 0$ or $y =$ constant. Since this curve has discontinuous first derivatives, we should not expect to find it as a solution to Eq. (2.11).

This example is valuable in emphasizing the restrictions that surround the derivation and the meaning of the stationary condition. Exercises 7 and 8 examine the conditions for the pathological behavior for a symmetric example. More information can be found in many texts on the calculus of variations.

3. *The brachistochrone problem.* (See Fig. 2.4a.) This well-known problem is to find the curve joining two points, along which a particle falling from rest under the influence of gravity travels from the higher to the lower point in the least time.

If v is the speed along the curve, then the time required to fall an arc length ds is ds/v, and the problem is to find a minimum of the integral

$$t_{12} = \int_1^2 \frac{ds}{v}.$$

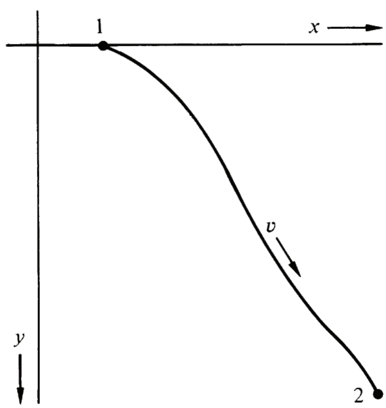

FIGURE 2.4a The brachistochrone problem.

If y is measured down from the initial point of release, the conservation theorem for the energy of the particle can be written as

$$\tfrac{1}{2}mv^2 = mgy$$

or

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

Then the expression for t_{12} becomes

$$t_{12} = \int_1^2 \frac{\sqrt{1+\dot{y}^2}}{\sqrt{2gy}}\,dx,$$

and f is identified as

$$f = \sqrt{\frac{1+\dot{y}^2}{2gy}}.$$

The integration of Eq. (2.11) with this form for f is straightforward and is left as an exercise.

The parametric solution in terms of its one parameter, a, given by

$$x = a(\phi - \sin\phi), \;\; y = a(1 - \cos\phi),$$

is sketched in Fig. 2.4b for the first cycle ($0 \le x \le 2\pi a$) and the beginning of the second cycle. Three cases of solutions are indicated. A power-series expansion of the solution for the limit $y \ll a$ gives

$$y = a\sqrt[3]{\frac{9}{2}(x/a)^2}.$$

The brachistochrone problem is famous in the history of mathematics, for it was the analysis of this problem by John Bernoulli that led to the formal foundation of the calculus of variations.

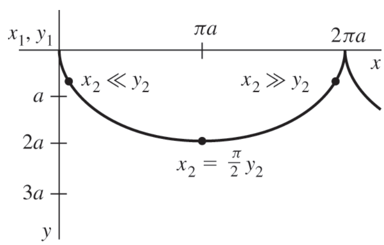

FIGURE 2.4b Cycloid solution to the brachistochrone problem showing positions on the curve for the three cases $x_2 \ll y_2$, $x_2 = \frac{\pi}{2}y_2$, and $x_2 \gg y_2$.

2.3 ■ DERIVATION OF LAGRANGE'S EQUATIONS FROM HAMILTON'S PRINCIPLE

The fundamental problem of the calculus of variations is easily generalized to the case where f is a function of many independent variables y_i, and their derivatives $\dot{y}_i$. (Of course, all these quantities are considered as functions of the parametric variable x.) Then a variation of the integral J,

$$\delta J = \delta \int_1^2 f(y_1(x); y_2(x), \ldots, \dot{y}_1(x); \dot{y}_2(x), \ldots, x)\, dx, \tag{2.14}$$

is obtained, as before, by considering J as a function of parameter α that labels a possible set of curves $y_1(x, \alpha)$. Thus, we may introduce α by setting

$$\begin{array}{ccc} y_1(x,\alpha) = & y_1(x,0) + & \alpha\eta_1(x), \\ y_2(x,\alpha) = & y_2(x,0) + & \alpha\eta_2(x), \\ \vdots & \vdots & \vdots \end{array} \tag{2.15}$$

where $y_1(x, 0)$, $y_2(x, 0)$, etc., are the solutions of the extremum problem (to be obtained) and η_1, η_2, etc., are independent functions of x that vanish at the end points and that are continuous through the second derivative, but otherwise are completely arbitrary.

The calculation proceeds as before. The variation of J is given in terms of

$$\frac{\partial J}{\partial \alpha} d\alpha = \int_1^2 \sum_i \left(\frac{\partial f}{\partial y_i} \frac{\partial y_i}{\partial \alpha} d\alpha + \frac{\partial f}{\partial \dot{y}_i} \frac{\partial \dot{y}_i}{\partial \alpha} d\alpha \right) dx. \tag{2.16}$$

Again we integrate by parts the integral involved in the second sum of Eq. (2.16):

$$\int_1^2 \frac{\delta f}{\partial \dot{y}_i} \frac{\partial^2 y_i}{\partial \alpha\, \partial x} dx = \left.\frac{\partial f}{\partial \dot{y}_i} \frac{\partial y_i}{\partial \alpha}\right|_1^2 - \int_1^2 \frac{\partial y_i}{\partial \alpha} \frac{d}{dx} \left(\frac{\partial f}{\partial \dot{y}_i} \right) dx,$$

where the first term vanishes because all curves pass through the fixed end points. Substituting in (2.16), δJ becomes

$$\delta J = \int_1^2 \sum_i \left(\frac{\partial f}{\partial y_i} - \frac{d}{dx} \frac{\partial f}{\partial \dot{y}_i} \right) \delta y_i \, dx, \tag{2.17}$$

where, in analogy with (2.12), the variation δy_i is

$$\delta y_i = \left(\frac{\partial y_i}{\partial \alpha} \right)_0 d\alpha.$$

Since the y variables are independent, the variations δy_i are independent (e.g., the functions $\eta_i(x)$ will be independent of each other). Hence, by an obvious extension of the fundamental lemma (cf. Eq. (2.10)), the condition that δJ is zero requires that the coefficients of the δy_i separately vanish:

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}_i} = 0, \qquad i = 1, 2, \ldots, n. \tag{2.18}$$

Equations (2.18) represent the appropriate generalization of (2.11) to several variables and are known as the *Euler-Lagrange differential equations*. Their solutions represent curves for which the variation of an integral of the form given in (2.14) vanishes. Further generalizations of the fundamental variational problem are easily possible. Thus, we can take f as a function of higher derivatives $\ddot{y}$, $\dddot{y}$, etc., leading to equations different from (2.18). Or we can extend it to cases where there are several parameters x_j and the integral is then multiple, with f also involving as variables derivatives of y_i with respect to each of the parameters x_j. Finally, it is possible to consider variations in which the end points are *not* held fixed.

For present purposes, what we have derived here suffices, for the integral in Hamilton's principle,

$$I = \int_1^2 L(q_i, \dot{q}_i, t)\, dt, \tag{2.19}$$

has just the form stipulated in (2.14) with the transformation

$$x \to t$$
$$y_i \to q_i$$
$$f(y_i, \dot{y}_i, x) \to L(q_i, \dot{q}_i, t).$$

In deriving Eqs. (2.18), we assumed that the y_i variables are independent. The corresponding condition in connection with Hamilton's principle is that the generalized coordinates q_i be independent, which requires that the constraints be holonomic. The Euler-Lagrange equations corresponding to the integral I then become the Lagrange equations of motion,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \qquad i = 1, 2, \ldots, n,$$

and we have accomplished our original aim, to show that Lagrange's equations follow from Hamilton's principle—for monogenic systems with holonomic constraints.

2.4 ■ EXTENDING HAMILTON'S PRINCIPLE TO SYSTEMS WITH CONSTRAINTS

In Section 1.3 we solved problems with holonomic constraints by choosing coordinates such that the constraint equations (1.37) become a trivial $0 = 0$ set of equations. In this section we show that Hamilton's principle can be used to solve systems with holonomic constraints as well as certain types of non-holonomic systems.

First consider holonomic constraints. When we derive Lagrange's equations from either Hamilton's or D'Alembert's principle, the holonomic constraints appear in the last step when the variations in the q_i were considered independent of each other. However, the virtual displacements in the δq_I's may not be consistent with the constraints. If there are n variables and m constraint equations f_α of the form of Eq. (1.37), the extra virtual displacements are eliminated by the method of *Lagrange undetermined multipliers*.

We modify the integral in Eq. (2.19) to be

$$I = \int_1^2 \left(L + \sum_{\alpha=1}^{m} \lambda_\alpha f_a \right) dt, \tag{2.20}$$

and allow the q_α and the λ_α to vary independently to obtain $n+m$ equations. The variations of the λ_α's give the m constraint equations. The variations of the q_i's give

$$\delta I = \int_1^2 dt \left(\sum_{i=1}^{n} \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} + \sum_{\alpha=1}^{m} \lambda_a \frac{\partial f_a}{\partial q_i} \right) \delta q_i \right) = 0. \tag{2.21}$$

However, the δq_i's are not independent. We choose the λ_α's so that m of the equations are satisfied for arbitrary δq_i, and then choose the variations of the δq_i in the remaining $n-m$ equations independently. Thus we obtain m equations of the form

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} + \sum_{\alpha=1}^{m} \lambda_a \frac{\partial f_a}{\partial q_k} = 0, \tag{2.22}$$

for $k = 1, \ldots, m$. The equality follows from the choice of the λ_α's. We also have the same expressions as Eq. (2.22) for $k = m+1, \ldots, n$, where the equality follows from the virtual variations of the δq_i's.

This solves the system at the expense of introducing m functions λ_α. We can understand this by considering that Eqs. (2.22), for $k = 1, \ldots, n$, can be written as

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = -\sum_{\alpha=1}^{m} \lambda_a \frac{\partial f_a}{\partial q_k} = Q_k, \tag{2.23}$$

where the Q_k are generalized forces. The functions, Q_κ, have the magnitudes of the forces needed to produce the individual constraints; however, since the choice of the "+" in the third term of Eq. (2.22) is arbitrary, we can mathematically determine only the magnitudes of these generalized forces. You need to understand the physics to determine their directions.

As an example, consider a smooth solid hemisphere of radius a placed with its flat side down and fastened to the Earth whose gravitational acceleration is g. Place a small mass M at the top of the hemisphere with an infinitesimal displacement off center so the mass slides down without friction. Choose coordinates x, y, z centered on the base of the hemisphere with z vertical and the x-z plane containing the initial motion of the mass.

Let θ be the angle from the top of the sphere to the mass. The Lagrangian is $L = \frac{1}{2}M(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz$. The initial conditions allow us to ignore the y coordinate, so the constraint equation is $a - \sqrt{x^2 + z^2} = 0$. Expressing the problem in terms of $r^2 = x^2 + z^2$ and $x/z = \cos\theta$, Lagrange's equations are $Ma\dot{\theta}^2 - Mg\cos\theta + \lambda = 0$, and $Ma^2\ddot{\theta} + Mga\sin\theta = 0$. Solve the second equation and then the first to obtain

$$\dot{\theta}^2 = -\frac{2g}{a}\cos\theta + \frac{2g}{a} \quad \text{and} \quad \lambda = Mg(3\cos\theta - 2).$$

So λ is the magnitude of the force keeping the particle on the sphere and since $\lambda = 0$ when $\theta = \cos^{-1}(\frac{2}{3})$, the mass leaves the sphere at that angle.

In general, nonholonomic constraints cannot be expressed by a variational principle. One of the exceptions is semi-holonomic constraints where the constraints can be written as a set of functions of the form

$$f_\alpha(q_1, \ldots, q_n; \dot{q}_1, \ldots, \dot{q}_n; t) = 0, \tag{2.24}$$

where $\alpha = 1, 2, \ldots, m$. Equation (2.24) commonly appears in the restricted form

$$f_\alpha = \sum_{k=1}^{n} a_{ak}\dot{q}_k + a_0 = 0, \tag{2.25}$$

where the f_α are a set of nonintegrable differential expressions and the $a_{\alpha k}$ and a_0 are functions of the q_ι and t. In these cases, since we cannot integrate the constraints, there are more variables than equations. However, we can treat the variations in the same fashion as before by writing*

$$\delta \int_{t_1}^{t_2} \left(L + \sum_{\alpha=1}^{m} \mu_a f_a \right) dt = 0, \tag{2.26}$$

where the symbol μ is used to distinguish these multipliers from the holonomic Lagrange multipliers. If we assume that $\mu_\alpha = \mu_\alpha(t)$, the equations resulting from the virtual displacements are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = Q_k = -\sum_{\alpha=1}^{m} \mu_a \frac{\partial f_a}{\partial \dot{q}_k}, \tag{2.27}$$

*J. Ray, *Amer. J. Phys.* **34** (1202), 1969; E. J. Saletan & A. H. Comer, *Amer. J. Phys.* **38** (892–897), 1970.

and the $\delta\mu_\alpha$ give the equations of constraint (2.23). These two sets (Eq. (2.26) and (2.27)) together constitute $n+m$ equations for the $n+m$ unknowns. Hence they can be interpreted as equivalent to an $n+m$ holonomic system with generalized forces Q_k. The generalization to $\mu_\alpha = \mu_\alpha(q_1, \ldots, q_n; \dot{q}_1, \ldots, \dot{q}_n; t)$ is straightforward.

As an example, consider a particle with the Lagrangian

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(x, y, z) \tag{2.28}$$

subject to the nonholonomic constraint

$$f(x, \dot{x}, y, \dot{y}, z) = \dot{x}y^2 + x\dot{y} + kz = 0, \tag{2.29}$$

with k a constant. The resulting equations of motion are

$$m\ddot{x} + \mu y^2 - \frac{\partial V}{\partial x} = 0, \tag{2.30a}$$

$$m\ddot{y} + \mu x - \frac{\partial V}{\partial y} = 0, \tag{2.30b}$$

and

$$m\ddot{z} - \frac{\partial V}{\partial z} = 0. \tag{2.30c}$$

We now solve the four equations ((2.29) and (3.30)) to find $x(t)$, $y(t)$, $z(t)$, and the multiplier $\mu(t)$.

In this process we have obtained more information than was originally sought. Not only do we get the q_k's we set out to find, but we also get $m\lambda_l$'s. What is the physical significance of the λ_l's? Suppose we remove the constraints on the system, but instead apply external forces Q'_k in such a manner as to keep the motion of the system unchanged. The equations of motion likewise remain the same. Clearly these extra applied forces must be equal to the forces of constraint, for they are the forces applied to the system so as to satisfy the condition of constraint. Under the influence of these forces Q'_k, the equations of motion are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = Q'_k. \tag{2.31}$$

But these must be identical with Eqs. (2.24). Hence, we can identify (2.25) with Q'_k, the generalized forces of constraint. In this type of problem we really do not eliminate the forces of constraint from the formulation. They are supplied as part of the answer.

Although it is not obvious, the version of Hamilton's principle adopted here for semiholonomic systems also requires that the constraints do no work in virtual displacements. This can be most easily seen by rewriting Hamilton's principle in the form

$$\delta \int_{t_1}^{t_2} L\,dt = \delta \int_{t_1}^{t_2} T\,dt - \delta \int_{t_1}^{t_2} U\,dt = 0. \tag{2.32}$$

If the variation of the integral over the generalized potential is carried out by the procedures of Section 2.3, the principle takes the form

$$\delta \int_{t_1}^{t_2} T\,dt = \int_{t_1}^{t_2} \sum_k \left[\frac{\partial U}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_k} \right) \right] \delta q_k dt; \tag{2.33}$$

or, by Eq. (1.58),

$$\delta \int_{t_1}^{t_2} T\,dt = - \int_{t_1}^{t_2} \sum_k Q_k \delta q_k dt. \tag{2.34}$$

In this dress, Hamilton's principle says that the difference in the time integral of the kinetic energy between two neighboring paths is equal to the negative of the time integral of the work done in the virtual displacements between the paths. The work involved is that done only by the forces derivable from the generalized potential. The same Hamilton's principle holds for both holonomic and semiholonomic systems, it must be required that the additional forces of semiholonomic constraints do no work in the displacements δq_k. This restriction parallels the earlier condition that the virtual work of the forces of holonomic constraint also be zero (cf. Section 1.4). In practice, the restriction presents little handicap to the applications, as many problems in which the semiholonomic formalism is used relate to rolling without slipping, where the constraints are obviously workless.

Note that Eq. (2.20) is not the most general type of nonholonomic constraint; e.g., it does not include equations of constraint in the form of inequalities. On the other hand, it does include holonomic constraints. A holonomic equation of constraint,

$$f(q_1, q_2, q_3, \ldots, q_n, t) = 0, \tag{2.35}$$

is equivalent to (2.20) with no dependence on $\dot{q}_k$. Thus, the Lagrange multiplier method can be used also for holonomic constraints when (1) it is inconvenient to reduce all the q's to independent coordinates or (2) we might wish to obtain the forces of constraint.

As another example of the method, let us consider the following somewhat trivial illustration—a hoop rolling, without slipping, down an inclined plane. In this example, the constraint of "rolling" is actually holonomic, but this fact will be immaterial to our discussion. On the other hand, the holonomic constraint that the hoop be on the inclined plane will be contained implicitly in our choice of generalized coordinates.

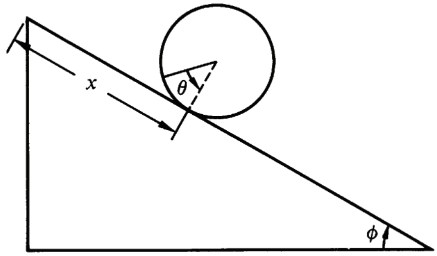

FIGURE 2.5 A hoop rolling down an inclined plane.

The two generalized coordinates are x, θ, as in Fig. 2.5, and the equation of rolling constraint is

$$r\,d\theta = dx.$$

The kinetic energy can be resolved into kinetic energy of motion of the center of mass plus the kinetic energy of motion about the center of mass:

$$T = \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}Mr^2\dot{\theta}^2.$$

The potential energy is

$$V = Mg(l - x)\sin\phi,$$

where l is the length of the inclined plane and the Lagrangian is

$$\begin{aligned} L &= T - V \\ &= \frac{M\dot{x}^2}{2} + \frac{Mr^2\dot{\theta}^2}{2} - Mg(l - x)\sin\phi. \end{aligned} \tag{2.36}$$

Since there is one equation of constraint, only one Lagrange multiplier λ is needed. The coefficients appearing in the constraint equation are:

$$a_\theta = r, \quad a_x = -1.$$

The two Lagrange equations therefore are

$$M\ddot{x} - Mg\sin\phi + \lambda = 0, \tag{2.37}$$

$$Mr^2\ddot{\theta} - \lambda r = 0, \tag{2.38}$$

which along with the equation of constraint,

$$r\dot{\theta} = \dot{x}, \tag{2.39}$$

constitutes three equations for three unknowns, θ, x, λ.

Differentiating (2.39) with respect to time, we have

$$r\ddot{\theta} = \ddot{x}.$$

Hence, from (2.38)

$$M\ddot{x} = \lambda,$$

and (2.37) becomes

$$\ddot{x} = \frac{g\sin\phi}{2}$$

along with

$$\lambda = \frac{Mg\sin\phi}{2}$$

and

$$\ddot{\theta} = \frac{g\sin\phi}{2r}.$$

Thus, the hoop rolls down the incline with only one-half the acceleration it would have slipping down a frictionless plane, and the friction force of constraint is $\lambda = Mg\sin\phi/2$.

2.5 ■ ADVANTAGES OF A VARIATIONAL PRINCIPLE FORMULATION

Although we can extend the original formulation of Hamilton's principle (2.2) to include some nonholonomic constraints, in practice this formulation of mechanics is most useful when a Lagrangian of independent coordinates can be set up for the system. The variational principle formulation has been justly described as "elegant," for in the compact Hamilton's principle is contained all of the mechanics of holonomic systems with forces derivable from potentials. The principle has the further merit that it involves only physical quantities that can be defined without reference to a particular set of generalized coordinates, namely, the kinetic and potential energies. The formulation is therefore automatically invariant with respect to the choice of coordinates for the system.

From the variational Hamilton's principle, it is also obvious why the Lagrangian is always uncertain to a total time derivative of any function of the coordinates and time, as mentioned at the end of Section 1.4. The time integral of such a total derivative between points 1 and 2 depends only on the values of the arbitrary function at the end points. As the variation at the end points is zero, the addition of the arbitrary time derivative to the Lagrangian does not affect the variational behavior of the integral.

Another advantage is that the Lagrangian formulation can be easily extended to describe systems that are not normally considered in dynamics—such as the elastic field, the electromagnetic field, and field properties of elementary particles. Some of these generalizations will be considered later, but as three simple examples of its application outside the usual framework of mechanics, let us consider the cases of an RL circuit, an LC circuit, and coupled circuits.

We consider the physical system of a battery of voltage V in series with an inductance L and a resistance of value R and choose the electric charge q as the dynamical variable. The inductor acts as the kinetic energy term since the inductive effect depends upon the time rate of change of the charge. The resistor provides a dissipative term and the potential energy is qV. The dynamic terms in Lagrange's equation with dissipation (1.70) are

$$T = \tfrac{1}{2}L\dot{q}^2, \quad \mathcal{F} = \tfrac{1}{2}R\dot{q}^2,$$

and potential energy $= qV$. The equation of motion is

$$V = L\ddot{q} + R\dot{q} = L\dot{I} + RI, \tag{2.40}$$

where $I = \dot{q}$ is the electric current. A solution for a battery connected to the circuit at time $t = 0$ is

$$I = I_0(1 - e^{-Rt/L}),$$

where $I_0 = V/R$ is the final steady-state current flow.

The mechanical analog for this is a sphere of radius a and effective mass m' falling in a viscous fluid of constant density and viscosity η under the force of gravity. The effective mass is the difference between the actual mass and the mass of the displaced fluid, and the direction of motion is along the y axis. For this system,

$$T = \tfrac{1}{2}m'\dot{y}^2, \quad \mathcal{F} = 3\pi\eta a\dot{y}^2,$$

and potential energy $= m'gy$, where the frictional drag force, $F_f = 6\pi\eta a\dot{y}$, called Stokes' law, was given at the end of Section 1.5.

The equation of motion is given by Lagrange's equations (1.70) as

$$m'g = m'\ddot{y} + 6\pi\eta a\dot{y}.$$

Using $v = \dot{y}$, the solution (if the motion starts from rest at $t = 0$), is

$$v = v_o(1 - e^{-t/\tau})$$

where $\tau = m'/(6\pi\eta a)$ is a measure of the time it takes for the sphere to reach $1/e$ of its terminal speed of $v_0 = m'g/6\pi\eta a$.

Another example from electrical circuits is an inductance, L, in series with a capacitance, C. The capacitor acts as a source of potential energy given by

q^2/C where q is the electric charge. The Lagrangian produces the equation of motion,

$$L\ddot{q} + \frac{q}{C} = 0, \tag{2.41}$$

which has the solution

$$q = q_0 \cos \omega_0 t,$$

where q_0 is the charge stored in the capacitor at $t = 0$, and the assumption is that no charge is flowing at $t = 0$. The quantity

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

is the resonant frequency of the system.

The mechanical analog of this system is the simple harmonic oscillator described by the Lagrangian $L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2$, which gives an equation of motion,

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

whose solution for the same boundary conditions is

$$x = x_0 \cos \omega_0 t \qquad \text{with} \quad \omega_0 = \sqrt{k/m}.$$

These two examples show that an inductance is an inertial term, the electrical analog of mass. Resistance is the analog of Stokes' law type of frictional drag, and the capacitance term $1/C$ represents a Hooke's law spring constant. With this background, a system of coupled electrical circuits of the type shown in Fig. 2.6 has a Lagrangian of the form

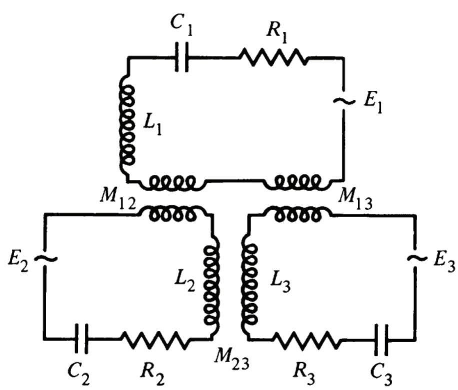

FIGURE 2.6 A system of coupled circuits to which the Lagrangian formulation can be applied.

$$L = \frac{1}{2}\sum_{j} L_j \dot{q}_j^2 + \frac{1}{2}\sum_{\substack{jk \\ j\neq k}} M_{jk}\dot{q}_j\dot{q}_k - \sum_j \frac{q_j^2}{2C_j} + \sum_j e_j(t) q_j,$$

and a dissipation function

$$\mathcal{F} = \frac{1}{2}\sum_j R_j \dot{q}_j^2.$$

where the mutual inductance terms, $M_{jk}\dot{q}_j\dot{q}_k$, are added to take into account the coupling between inductors. The Lagrange equations are

$$L_j \frac{d^2 q_j}{dt^2} + \sum_{\substack{k \\ j\neq k}} M_{jk}\frac{d^2 q_k}{dt^2} + R_j \frac{dq_j}{dt} + \frac{q_j}{C_j} = E_j(t). \tag{2.42}$$

where the $E_j(t)$ terms are the external emf's.

This description of two different physical systems by Lagrangians of the same form means that all the results and techniques devised for investigating one of the systems can be taken over immediately and applied to the other. In this particular case, the study of the behavior of electrical circuits has been pursued intensely and some special techniques have been developed; these can be directly applied to the corresponding mechanical systems. Much work has been done in formulating equivalent electrical problems for mechanical or acoustical systems, and vice versa. Terms hitherto reserved for electrical circuits (reactance, susceptance, etc.) are now commonly found in treatises on the theory of vibrations of mechanical systems.

Additionally, one type of generalization of mechanics is due to a subtler form of equivalence. We have seen that the Lagrangian and Hamilton's principle together form a compact invariant way of obtaining the mechanical equations of motion. This possibility is not reserved for mechanics only; in almost every field of physics variational principles can be used to express the "equations of motion," whether they be Newton's equations, Maxwell's equations, or the Schrödinger equation. Consequently, when a variational principle is used as the basis of the formulation, all such fields will exhibit, at least to some degree, a *structural analogy*. When the results of experiments show the need for alterating the physical content in the theory of one field, this degree of analogy has often indicated how similar alterations may be carried out in other fields. Thus, the experiments performed early in this century showed the need for quantization of both electromagnetic radiation and elementary particles. The methods of quantization, however, were first developed for particle mechanics, starting essentially from the Lagrangian formulation of classical mechanics. By describing the electromagnetic field by a Lagrangian and corresponding Hamilton's variational principle, it is possible to carry over the methods of particle quantization to construct a quantum electrodynamics (cf. Sections 13.5 and 13.6).

2.6 ■ CONSERVATION THEOREMS AND SYMMETRY PROPERTIES

Thus far, we have been concerned primarily with obtaining the equations of motion, but little has been said about how to solve them for a particular problem once they are obtained. In general, this is a question of mathematics. A system of n degrees of freedom will have n differential equations that are second order in time. The solution of each equation will require two integrations resulting, all told, in $2n$ constants of integration. In a specific problem these constants will be determined by the initial conditions, i.e., the initial values of the nq_j's and the $n\dot{q}_j$'s. Sometimes the equations of motion will be integrable in terms of known functions, but not always. In fact, the majority of problems are not completely integrable. However, even when complete solutions cannot be obtained, it is often possible to extract a large amount of information about the physical nature of the system motion. Indeed, such information may be of greater interest to the physicist than the complete solution for the generalized coordinates as a function of time. It is important, therefore, to see how much can be stated about the motion of a given system without requiring a complete integration of the problem.*

In many problems a number of first integrals of the equations of motion can be obtained immediately; by this we mean relations of the type

$$f(q_1, q_2, \ldots, \dot{q}_1, \dot{q}_2, \ldots, t) = \text{constant}, \tag{2.43}$$

which are first-order differential equations. These first integrals are of interest because they tell us something physically about the system. They include, in fact, the conservation laws obtained in Chapter 1.

Let us consider as an example a system of mass points under the influence of forces derived from potentials dependent on position only. Then

$$\begin{aligned}\frac{\partial L}{\partial \dot{x}_i} &\equiv \frac{\partial T}{\partial \dot{x}_i} - \frac{\partial V}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} = \frac{\partial}{\partial \dot{x}_i} \sum \frac{1}{2} m_i \left(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2\right) \\ &= m_i \dot{x}_i = p_{ix},\end{aligned}$$

which is the x component of the linear momentum associated with the ith particle. This result suggests an obvious extension to the concept of momentum. The generalized momentum associated with the coordinate q_j shall be defined as

$$p_j = \frac{\partial L}{\partial \dot{q}_j}. \tag{2.44}$$

The terms *canonical momentum* and *conjugate momentum* are often also used for p_j. Notice that if q_j is not a Cartesian coordinate, p_j does not necessarily have the dimensions of a linear momentum. Further, if there is a velocity-dependent potential, then even with a Cartesian coordinate q_j the associated *generalized*

*In this and succeeding sections it will be assumed, unless otherwise specified, the system is such that its motion is completely described by a Hamilton's principle of the form (2.2).

momentum will not be identical with the usual *mechanical* momentum. Thus, in the case of a group of particles in an electromagnetic field, the Lagrangian is (cf. 1.63)

$$L = \sum_i \frac{1}{2} m_i \dot{r}_i^2 - \sum_i q_i \phi(x_i) + \sum_i q_i \mathbf{A}(x_i) \cdot \dot{\mathbf{r}}_i$$

(q_i here denotes charge) and the generalized momentum conjugate to x_i is

$$p_{ix} = \frac{\partial L}{\partial \dot{x}_i} = m_i \dot{x}_i + q_i A_x, \tag{2.45}$$

i.e., mechanical momentum plus an additional term.

If the Lagrangian of a system does not contain a given coordinate q_j (although it may contain the corresponding velocity $\dot{q}_j$), then the coordinate is said to be *cyclic* or *ignorable*. This definition is not universal, but it is the customary one and will be used here. The Lagrange equation of motion,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0,$$

reduces, for a cyclic coordinate, to

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} = 0$$

or

$$\frac{dp_j}{dt} = 0,$$

which mean that

$$p_j = \text{constant}. \tag{2.46}$$

Hence, we can state as a general conservation theorem that *the generalized momentum conjugate to a cyclic coordinate is conserved.*

Note that the derivation of Eq. (2.46) assumes that q_j is a generalized coordinate; one that is linearly independent of all the other coordinates. When equations of constraint exist, all the coordinates are not linearly independent. For example, the angular coordinate θ is not present in the Lagrangian of a hoop rolling without slipping in a horizontal plane that was previously discussed, but the angle appears in the constraint equations $r d\theta = dx$. As a result, the angular momentum, $p_\theta = mr^2\dot{\theta}$, is not a constant of the motion.

Equation (2.46) constitutes a first integral of the form (2.43) for the equations of motion. It can be used formally to eliminate the cyclic coordinate from the problem, which can then be solved entirely in terms of the remaining generalized coordinates. Briefly, the procedure, originated by Routh, consists in modifying the Lagrangian so that it is no longer a function of the generalized velocity corresponding to the cyclic coordinate, but instead involves only its conjugate momentum. The advantage in so doing is that p_j can then be considered one of the constants of integration, and the remaining integrations involve only the noncyclic coordinates. We shall defer a detailed discussion of Routh's method until the Hamiltonian formulation (to which it is closely related) is treated.

Note that the conditions for the conservation of generalized momenta are more general than the two momentum conservation theorems previously derived. For example, they furnish a conservation theorem for a case in which the law of action and reaction is violated, namely, when electromagnetic forces are present. Suppose we have a single particle in a field in which neither ϕ nor $\mathbf{A}$ depends on x. Then x nowhere appears in L and is therefore cyclic. The corresponding canonical momentum p_x must therefore be conserved. From (1.63) this momentum now has the form

$$p_x = m\dot{x} + qA_x = \text{constant}. \tag{2.47}$$

In this case, it is not the mechanical linear momentum $m\dot{x}$ that is conserved but rather its sum with qA_x.* Nevertheless, it should still be true that the conservation theorems of Chapter 1 are contained within the general rule for cyclic coordinates; with proper restrictions (2.46) should reduce to the theorems of Section 1.2.

We first consider a generalized coordinate q_j, for which a change dq_j represents a translation of the system as a whole in some given direction. An example would be one of the Cartesian coordinates of the center of mass of the system. Then clearly q_j cannot appear in T, for velocities are not affected by a shift in the origin, and therefore the partial derivative of T with respect to q_j must be zero. Further, we will assume conservative systems for which V is not a function of the velocities, so as to eliminate such complications as electromagnetic forces. The Lagrange equation of motion for a coordinate so defined then reduces to

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} \equiv \dot{p}_j = -\frac{\partial V}{\partial q_j} \equiv Q_j. \tag{2.48}$$

We will now show that (2.48) is the equation of motion for the total linear momentum, i.e., that Q_j represents the component of the total force along the direction of translation of q_j, and p_j is the component of the total linear momentum along this direction. In general, the generalized force Q_j is given by Eq. (1.49):

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}.$$

Since dq_j corresponds to a translation of the system along some axis, the vectors $\mathbf{r}_i(q_j)$ and $\mathbf{r}_i(q_j + dq_j)$ are related as shown in Fig. 2.7. By the definition of a

*It can be shown from classical electrodynamics that under these conditions, i.e., neither $\mathbf{A}$ nor ϕ depending on x, that qA_x is exactly the x-component of the electromagnetic linear momentum of the field associated with the charge q.

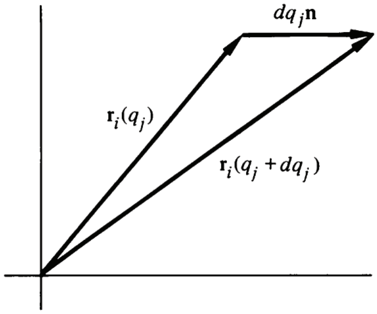

FIGURE 2.7 Change in a position vector under translation of the system.

derivative, we have

$$\frac{\partial \mathbf{r}_i}{\partial q_j} = \lim_{dq_j \to 0} \frac{\mathbf{r}_i\left(q_j + dq_j\right) - \mathbf{r}_i\left(q_j\right)}{dq_j} = \frac{dq_j}{dq_j}\mathbf{n} = \mathbf{n}, \tag{2.49}$$

where **n** is the unit vector along the direction of the translation. Hence,

$$Q_j = \sum \mathbf{F}_i \cdot \mathbf{n} = \mathbf{n} \cdot \mathbf{F},$$

which (as was stated) is the component of the total force in the direction of **n**. To prove the other half of the statement, note that with the kinetic energy in the form

$$T = \frac{1}{2}\sum m_i \dot{\mathbf{r}}_i^2,$$

the conjugate momentum is

$$\begin{aligned} p_j = \frac{\partial T}{\partial \dot{q}_j} &= \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} \\ &= \sum_i m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}, \end{aligned}$$

using Eq. (1.51). Then from Eq. (2.49)

$$p_j = \mathbf{n} \cdot \sum_i m_i \mathbf{v}_i,$$

which again, as predicted, is the component of the total system linear momentum along **n**.

Suppose now that the translation coordinate q_j that we have been discussing is cyclic. Then q_j cannot appear in V and therefore

$$-\frac{\partial V}{\partial q_j} \equiv Q_j = 0.$$

But this is simply the familiar conservation theorem for linear momentum—that if a given component of the total applied force vanishes, the corresponding component of the linear momentum is conserved.

In a similar fashion, it can be shown that if a cyclic coordinate q_j is such that dq_j corresponds to a rotation of the system of particles around some axis, then the conservation of its conjugate momentum corresponds to conservation of an angular momentum. By the same argument used above, T cannot contain q_j, for a rotation of the coordinate system cannot affect the magnitude of the velocities. Hence, the partial derivative of T with respect to q_j must again be zero, and since V is independent of $\dot{q}_j$, we once more get Eq. (2.48). But now we wish to show that with q_j a rotation coordinate the generalized force is the component of the total applied torque about the axis of rotation, and p_j is the component of the total angular momentum along the same axis.

The generalized force Q_j is again given by

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j},$$

only the derivative now has a different meaning. Here the change in q_j must correspond to an infinitesimal rotation of the vector $\mathbf{r}_i$, keeping the magnitude of the vector constant. From Fig. 2.8, the magnitude of the derivative can easily be obtained:

$$|d\mathbf{r}_i| = r_i \sin\theta \, dq_j \quad \text{and} \quad \left| \frac{\partial \mathbf{r}_i}{\partial q_j} \right| = r_i \sin\theta,$$

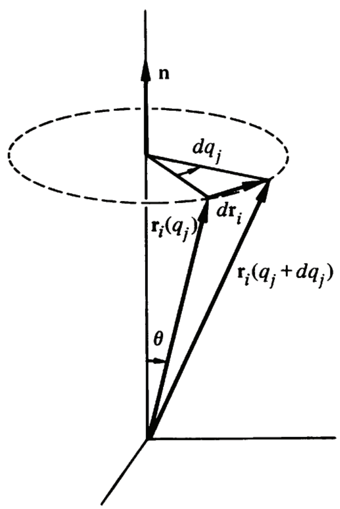

FIGURE 2.8 Change of a position vector under rotation of the system.

and its direction is perpendicular to both $\mathbf{r}_i$ and $\mathbf{n}$. Clearly, the derivative can be written in vector form as

$$\frac{\partial \mathbf{r}_i}{\partial q_j} = \mathbf{n} \times \mathbf{r}_i. \tag{2.50}$$

With this result, the generalized force becomes

$$\begin{aligned} Q_j &= \sum_i \mathbf{F}_i \cdot \mathbf{n} \times \mathbf{r}_i \\ &= \sum_i \mathbf{n} \cdot \mathbf{r}_i \times \mathbf{F}_i, \end{aligned}$$

reducing to

$$Q_j = \mathbf{n} \cdot \sum_i \mathbf{N}_i = \mathbf{n} \cdot \mathbf{N},$$

which proves the first part. A similar manipulation of p_j with the aid of Eq. (2.50) provides proof of the second part of the statement:

$$p_j = \frac{\partial T}{\partial \dot{q}_j} = \sum_i m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \sum_i \mathbf{n} \cdot \mathbf{r}_i \times m_i \mathbf{v}_i = \mathbf{n} \cdot \sum_i \mathbf{L}_i = \mathbf{n} \cdot \mathbf{L}.$$

Summarizing these results, we see that if the rotation coordinate q_j is cyclic, then Q_j, which is the component of the applied torque along $\mathbf{n}$, vanishes, and the component of $\mathbf{L}$ along $\mathbf{n}$ is constant. Here we have recovered the angular momentum conservation theorem out of the general conservation theorem relating to cyclic coordinates.

The significance of cyclic translation or rotation coordinates in relation to the properties of the system deserves some comment at this point. If a generalized coordinate corresponding to a displacement is cyclic, it means that a translation of the system, as if rigid, has no effect on the problem. In other words, if the system is *invariant* under translation along a given direction, the corresponding linear momentum is conserved. Similarly, the fact that a generalized rotation coordinate is cyclic (and therefore the conjugate angular momentum conserved) indicates that the system is invariant under rotation about the given axis. Thus, the momentum conservation theorems are closely connected with the *symmetry properties* of the system. If the system is spherically symmetric, we can say without further ado that all components of angular momentum are conserved. Or, if the system is symmetric only about the z axis, then only L_z will be conserved, and so on for the other axes. These symmetry considerations can often be used with relatively complicated problems to determine by inspection whether certain constants of the motion exist. (cf. Noether's theorem—Sec. 13.7.)

Suppose, for example, the system consists of a set of mass points moving in a potential field generated by fixed sources uniformly distributed on an infinite plane, say, the $z = 0$ plane. (The sources might be a mass distribution if the forces were gravitational, or a charge distribution for electrostatic forces.) Then the symmetry of the problem is such that the Lagrangian is invariant under a translation of the system of particles in the x- or y-directions (but not in the z-direction) and also under a rotation about the z axis. It immediately follows that the x- and y-components of the total linear momentum, P_x and P_y, are constants of the motion along with L_z, the z-component of the total angular momentum. However, if the sources were restricted only to the half plane, $x \geq 0$, then the symmetry for translation along the x axis and for rotation about the z axis would be destroyed. In that case, P_x and L_z could not be conserved, but P_y would remain a constant of the motion. We will encounter the connections between the constants of motion and the symmetry properties of the system several times in the following chapters.

2.7 ■ ENERGY FUNCTION AND THE CONSERVATION OF ENERGY

Another conservation theorem we should expect to obtain in the Lagrangian formulation is the conservation of total energy for systems where the forces are derivable from potentials dependent only upon position. Indeed, it is possible to demonstrate a conservation theorem for which conservation of total energy represents only a special case. Consider a general Lagrangian, which will be a function of the coordinates q_j and the velocities $\dot{q}_j$ and may also depend explicitly on the time. (The explicit time dependence may arise from the time variation of external potentials, or from time-dependent constraints.) Then the total time derivative of L is

$$\frac{dL}{dt} = \sum_j \frac{\partial L}{\partial q_j}\frac{dq_j}{dt} + \sum_j \frac{\partial L}{\partial \dot{q}_j}\frac{d\dot{q}_j}{dt} + \frac{\partial L}{\partial t}. \tag{2.51}$$

From Lagrange's equations,

$$\frac{\partial L}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right),$$

and (2.51) can be rewritten as

$$\frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\dot{q}_j + \sum_j \frac{\partial L}{\partial \dot{q}_j}\frac{d\dot{q}_j}{dt} + \frac{\partial L}{\partial t}$$

or

$$\frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j}\right) + \frac{\partial L}{\partial t}.$$

It therefore follows that

$$\frac{d}{dt}\left(\sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L\right) + \frac{\partial L}{\partial t} = 0. \tag{2.52}$$

The quantity in parentheses is oftentimes called the *energy function** and will be denoted by h:

$$h(q_1, \ldots, q_n; \quad \dot{q}_1, \ldots, \dot{q}_n; \quad t) = \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L, \tag{2.53}$$

and Eq. (2.52) can be looked on as giving the total time derivative of h:

$$\frac{dh}{dt} = -\frac{\partial L}{\partial t}. \tag{2.54}$$

If the Lagrangian is not an explicit function of time, i.e., if t does not appear in L explicitly but only implicitly through the time variation of q and $\dot{q}$, then Eq. (2.54) says that h is conserved. It is one of the first integrals of the motion and is sometimes referred to as Jacobi's integral.†

Under certain circumstances, the function h is the total energy of the system. To determine what these circumstances are, we recall that the total kinetic energy of a system can always be written as

$$T = T_0 + T_1 + T_2, \tag{1.73}$$

where T_0 is a function of the generalized coordinates only, $T_1(q, \dot{q})$ is linear in the generalized velocities, and $T_2(q, \dot{q})$ is a quadratic function of the $\dot{q}$'s. For a very wide range of systems and sets of generalized coordinates, the Lagrangian can be similarly decomposed as regards its functional behavior in the $\dot{q}$ variables:

$$L(q, \dot{q}, t) = L_0(q, t) + L_1(q, \dot{q}, t) + L_2(q, \dot{q}, t). \tag{2.55}$$

Here L_2 is a homogeneous function of the second degree (not merely quadratic) in $\dot{q}$, while L_1 is homogeneous of the first degree in $\dot{q}$. There is no reason intrinsic to mechanics that requires the Lagrangian to conform to Eq. (2.55), but in fact it does for most problems of interest. The Lagrangian clearly has this form when the forces are derivable from a potential not involving the velocities. Even with the velocity-dependent potentials, we note that the Lagrangian for a charged particle in an electromagnetic field, Eq. (1.63), satisfies Eq. (2.55). Now, recall that Euler's theorem states that if f is a homogeneous function of degree n in the variables x_i, then

$$\sum_i x_i \frac{\partial f}{\partial x_i} = nf. \tag{2.56}$$

*The energy function h is identical in value with the Hamiltonian H (See Chapter 8). It is given a different name and symbol here to emphasize that h is considered a function of n independent variables q_j and their time derivatives $\dot{q}_j$ (along with the time), whereas the Hamiltonian will be treated as a function of $2n$ independent variables, q_j, p_j (and possibly the time).

†This designation is most often confined to a first integral in the restricted three-body problem. However, the integral there is merely a special case of the energy function h, and there is some historical precedent to apply the name Jacobi integral to the more general situation.

Applied to the function h, Eq. (2.53), for the Lagrangians of the form (2.55), this theorem implies that

$$h = 2L_2 + L_1 - L = L_2 - L_0. \tag{2.57}$$

If the transformation equations defining the generalized coordinates, Eqs. (1.38), do not involve the time explicitly, then by Eqs. (1.73) $T = T_2$. If, further, the potential does not depend on the generalized velocities, then $L_2 = T$ and $L_0 = -V$, so that

$$h = T + V = E, \tag{2.58}$$

and the energy function is indeed the total energy. Under these circumstances, if V does not involve the time explicitly, neither will L. Thus, by Eq. (2.54), h (which is here the total energy), will be conserved.

Note that the conditions for conservation of h are in principle quite distinct from those that identify h as the total energy. We can have a set of generalized coordinates such that in a particular problem h is conserved but is not the total energy. On the other hand, h can be the total energy, in the form $T + V$, but not be conserved. Also note that whereas the Lagrangian is uniquely fixed for each system by the prescription $L = T - U$ independent of the choice of generalized coordinates, the energy function h depends in magnitude and functional form on the specific set of generalized coordinates. For one and the same system, various energy functions h of different physical content can be generated depending on how the generalized coordinates are chosen.

The most common case that occurs in classical mechanics is one in which the kinetic energy terms are all of the form $m\dot{q}_i^2/2$ or $p_i^2/2m$ and the potential energy depends only upon the coordinates. For these conditions, the energy function is both conserved and is also the total energy.

Finally, note that where the system is not conservative, but there are frictional forces derivable from a dissipation function $\mathcal{F}$, it can be easily shown that $\mathcal{F}$ is related to the decay rate of h. When the equations of motion are given by Eq. (1.70), including dissipation, then Eq. (2.52) has the form

$$\frac{dh}{dt} + \frac{\partial L}{\partial t} = \sum_j \frac{\partial \mathcal{F}}{\partial \dot{q}_j}\dot{q}_j.$$

By the definition of $\mathcal{F}$, Eq. (1.67), it is a homogeneous function of the $\dot{q}$'s of degree 2. Hence, applying Euler's theorem again, we have

$$\frac{dh}{dt} = -2\mathcal{F} - \frac{\partial L}{\partial t}. \tag{2.59}$$

If L is not an explicit function of time, *and* the system is such that h is the same as the energy, then Eq. (2.59) says that $2\mathcal{F}$ is the rate of energy dissipation,

$$\frac{dE}{dt} = -2\mathcal{F}, \tag{2.60}$$

a statement proved above (cf. Sec. 1.5) in less general circumstances.