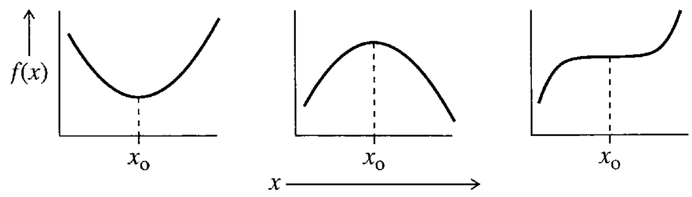

Figure 6.2 If $df/dx = 0$ at x_o, there are three possibilities: **(a)** If the second derivative is positive, then $f(x)$ has a minimum at x_o. **(b)** If the second derivative is negative, then $f(x)$ has a maximum. **(c)** If the second derivative is zero, then there may be a minimum, a maximum, or neither (as shown).

(6.3) **stationary**, in the sense that an infinitesimal variation of the path from its correct course doesn't change the value of the integral concerned. If you need to know that the integral is definitely minimum (or definitely maximum, or perhaps neither), you have to check this separately. Incidentally, we are now ready to explain the name of this chapter: Since our concern is how infinitesimal variations of a path change an integral, the subject is called the **calculus of variations**. For the same reason, the methods we shall develop are called variational methods, and a principle like Fermat's principle is a variational principle.

6.2 The Euler–Lagrange Equation

The two examples of the last section illustrate the general form of the so-called variational problem. We have an integral of the form

$$S = \int_{x_1}^{x_2} f[y(x), y'(x), x]\,dx \tag{6.4}$$

where $y(x)$ is an as-yet unknown curve joining two points (x_1, y_1) and (x_2, y_2) as in Figure 6.1; that is,

$$y(x_1) = y_1 \qquad \text{and} \qquad y(x_2) = y_2. \tag{6.5}$$

Among all the possible curves satisfying (6.5) (that is, joining the points 1 and 2), we have to find the one that makes the integral S a minimum (or maximum or at least stationary). To be definite, I shall suppose that we wish to find a minimum. Notice that the function f in (6.4) is a function of three variables $f = f(y, y', x)$, but because the integral follows the path $y = y(x)$ the integrand $f[y(x), y'(x), x]$ is actually a function of just the one variable x.

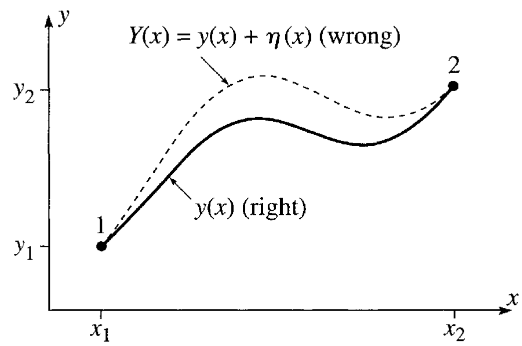

Figure 6.3 The path $y = y(x)$ between points 1 and 2 is the "right" path, the one for which the integral S of (6.4) is a minimum. Any other path $Y(x)$ is "wrong," in that it gives a larger value for S.

Let us denote the correct solution to our problem by $y = y(x)$. Then the integral S in (6.4) evaluated for $y = y(x)$ is less than for any neighboring curve $y = Y(x)$, as sketched in Figure 6.3. It is convenient to write the "wrong" curve $Y(x)$ as

$$Y(x) = y(x) + \eta(x) \tag{6.6}$$

where $\eta(x)$ (Greek "eta") is just the difference between the wrong $Y(x)$ and the right $y(x)$. Since $Y(x)$ must pass through the endpoints 1 and 2, $\eta(x)$ must satisfy

$$\eta(x_1) = \eta(x_2) = 0. \tag{6.7}$$

There are infinitely many choices for the difference $\eta(x)$; for example, we could choose $\eta = (x - x_1)(x_2 - x)$ or $\eta(x) = \sin[\pi(x - x_1)/(x_2 - x_1)]$.

The integral S taken along the wrong curve $Y(x)$ must be larger than that along the right curve $y(x)$, no matter how close the former is to the latter. To express this requirement, I shall introduce a parameter α and redefine $Y(x)$ to be

$$Y(x) = y(x) + \alpha\eta(x). \tag{6.8}$$

The integral S taken along the curve $Y(x)$ now depends on the parameter α, so I shall call it $S(\alpha)$. The right curve $y(x)$ is obtained from (6.8) by setting $\alpha = 0$. Thus the requirement that S is minimum for the right curve $y(x)$ implies that $S(\alpha)$ is a minimum at $\alpha = 0$. With this result, we have converted our problem to the traditional problem from elementary calculus of making sure that an ordinary function [namely $S(\alpha)$] has a minimum at a specified point ($\alpha = 0$). To ensure this, we must just check that the derivative $dS/d\alpha$ is zero when $\alpha = 0$.

If we write out the integral $S(\alpha)$ in detail, it looks like this:

$$\begin{aligned} S(\alpha) &= \int_{x_1}^{x_2} f(Y, Y', x)\, dx \\ &= \int_{x_1}^{x_2} f(y + \alpha\eta, y' + \alpha\eta', x)\, dx. \end{aligned} \tag{6.9}$$

To differentiate (6.9) with respect to α, we note that α appears in the integrand f, so we need to evaluate $\partial f/\partial\alpha$. Since α appears in two of the arguments of f, this gives two terms, namely (using the chain rule)

$$\frac{\partial f(y+\alpha\eta, y'+\alpha\eta', x)}{\partial\alpha} = \eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'},$$

and for $dS/d\alpha$ (which has to be zero)

$$\frac{dS}{d\alpha} = \int_{x_1}^{x_2}\frac{\partial f}{\partial\alpha}\,dx = \int_{x_1}^{x_2}\left(\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'}\right)dx = 0. \tag{6.10}$$

This condition must be true for any $\eta(x)$ satisfying (6.7); that is, for any choice of the "wrong" path $Y(x) = y(x) + \alpha\eta(x)$.

To take advantage of the condition (6.10), we need to rewrite the second term on the right using integration by parts[1] (remember that η' means $d\eta/dx$):

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

Because of the condition (6.7), the first term on the right (the "endpoint term") is zero. Thus[2]

$$\int_{x_1}^{x_2}\eta'(x)\frac{\partial f}{\partial y'}dx = -\int_{x_1}^{x_2}\eta(x)\,\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)dx. \tag{6.11}$$

Substituting this identity into (6.10), we find that

$$\int_{x_1}^{x_2}\eta(x)\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right)dx = 0. \tag{6.12}$$

This condition must be satisfied for any choice of the function $\eta(x)$. Therefore, as I shall argue in a moment, the factor in large parentheses must be zero:

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} = 0 \quad \text{(Euler–Lagrange Equation)} \tag{6.13}$$

for all x (in the relevant interval $x_1 \le x \le x_2$). This is the so-called **Euler–Lagrange equation** (named for the Swiss mathematician Leonhard Euler, 1707–1783, and the Italian-French physicist and mathematician Joseph Lagrange, 1736–1813), which lets

[1] If you are used to thinking of integration by parts in the form $\int v\,du = [uv] - \int u\,dv$, then you will find it helpful to recognize that another way to say the same thing is: $\int u'v\,dx = [uv] - \int uv'\,dx$. In words: In the integral $\int u'v\,dx$, you can move the prime from the u to the v if you change the sign and add the endpoint contribution $[uv]$.

[2] This is the simple form in which integration by parts often appears in physics: Provided the endpoint term $[uv]$ is zero (as often happens), integration by parts lets you move the differentiation from the u to the v as long as you change the sign; that is, $\int u'v\,dx = -\int uv'dx$.

us find the path for which the integral S is stationary. Before I illustrate its use, I need to discuss the step from (6.12) to (6.13), which is by no means obvious.

Equation (6.12) has the form $\int \eta(x)g(x)\,dx = 0$. I would certainly not claim that this condition alone implies that $g(x) = 0$ for all x. However, (6.12) holds for any choice of the function $\eta(x)$, and if $\int \eta(x)g(x)\,dx = 0$ for *any* $\eta(x)$, then we *can* conclude that $g(x) = 0$ for all x. To prove this, we must assume that all functions concerned are continuous, but, as physicists, we would take for granted that this is the case.[3] Now, to prove the assertion, let us assume the contrary, that $g(x)$ is nonzero in some interval between x_1 and x_2. Then choose a function $\eta(x)$ that has the same sign as $g(x)$ (that is, η is positive where g is positive and η is negative where g is negative). Then the integrand is continuous, satisfies $\eta(x)g(x) \geq 0$, and is nonzero at least in some interval. Under these conditions $\int \eta(x)g(x)\,dx$ cannot be zero. This contradiction implies that $g(x)$ *is* zero for all x.

This completes the proof of the Euler–Lagrange equation. The procedure for using it is this: (1) Set up the problem so that the quantity whose stationary path you seek is expressed as an integral in the standard form

$$S = \int_{x_1}^{x_2} f[y(x), y'(x), x]\,dx, \tag{6.14}$$

where $f[y(x), y'(x), x]$ is the function appropriate to your problem. (2) Write down the Euler–Lagrange equation (6.13) in terms of the function $f[y(x), y'(x), x]$. (3) Finally, solve (if possible) the differential equation (6.13) for the function $y(x)$ that defines the required stationary path. I shall illustrate this procedure with a couple of examples in the next section.

6.3 Applications of the Euler–Lagrange Equation

Let us start with the problem that began this chapter, finding the shortest path between two points in a plane.

EXAMPLE 6.1 Shortest Path between Two Points

We saw that the length of a path between points 1 and 2 is given by the integral (6.2) as

$$L = \int_1^2 ds = \int_{x_1}^{x_2} \sqrt{1 + y'^2}\,dx.$$

This has the standard form (6.14), with the function f given by

$$f(y, y', x) = (1 + y'^2)^{1/2}. \tag{6.15}$$

[3] The claimed result is clearly false if discontinuous functions are admitted. For instance, if we made $g(x)$ nonzero at just one point, then $\int \eta(x)g(x)\,dx$ would still be zero.

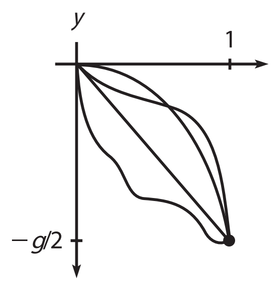

Fig. 6.2

Theorem 6.1 *If the function $x_0(t)$ yields a stationary value (that is, a local minimum, maximum, or saddle point) of S, then*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_0}\right) = \frac{\partial L}{\partial x_0}. \tag{6.15}$$

It is understood that we are considering the class of functions whose endpoints are fixed. That is, $x(t_1) = x_1$ and $x(t_2) = x_2$.

Proof: We will use the fact that if a certain function $x_0(t)$ yields a stationary value of S, then any other function very close to $x_0(t)$ (with the same endpoint values) yields essentially the same S, up to first order in any deviations. This is actually the definition of a stationary value. The analogy with regular functions is that if $f(b)$ is a stationary value of f, then $f(b+\epsilon)$ differs from $f(b)$ only at second order in the small quantity ϵ. This is true because $f'(b) = 0$, so there is no first-order term in the Taylor series expansion around b.

Assume that the function $x_0(t)$ yields a stationary value of S, and consider the function

$$x_a(t) \quad x_0(t) + a\beta(t), \tag{6.16}$$

[5] A saddle point is a point where there are no first-order changes in S, and where some of the second-order changes are positive and some are negative (like the middle of a saddle, of course).

[6] This follows from $y = -gt^2/2$, but pretend that we don't know this formula.

where a is a number, and where $\beta(t)$ satisfies $\beta(t_1) = \beta(t_2) = 0$ (to keep the endpoints of the function fixed), but is otherwise arbitrary. When producing the action $S[x_a(t)]$ in (6.14), the t is integrated out, so S is just a number. It depends on a, in addition to t_1 and t_2. Our requirement is that there be no change in S at first order in a. How does S depend on a? Using the chain rule, we have

$$\begin{aligned} \frac{\partial}{\partial a} S[x_a(t)] &= \frac{\partial}{\partial a} \int_{t_1}^{t_2} L\, dt = \int_{t_1}^{t_2} \frac{\partial L}{\partial a}\, dt \\ &= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x_a} \frac{\partial x_a}{\partial a} + \frac{\partial L}{\partial \dot{x}_a} \frac{\partial \dot{x}_a}{\partial a} \right) dt. \end{aligned} \tag{6.17}$$

In other words, a influences S through its effect on x, and also through its effect on $\dot{x}$. From Eq. (6.16), we have

$$\frac{\partial x_a}{\partial a} = \beta, \quad \text{and} \quad \frac{\partial \dot{x}_a}{\partial a} = \dot{\beta}, \tag{6.18}$$

so Eq. (6.17) becomes[7]

$$\frac{\partial}{\partial a} S[x_a(t)] = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x_a} \beta + \frac{\partial L}{\partial \dot{x}_a} \dot{\beta} \right) dt. \tag{6.19}$$

Now comes the one sneaky part of the proof. We will integrate the second term by parts (you will see this trick many times in your physics career). Using

$$\int \frac{\partial L}{\partial \dot{x}_a} \dot{\beta}\, dt = \frac{\partial L}{\partial \dot{x}_a} \beta - \int \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_a} \right) \beta\, dt, \tag{6.20}$$

Eq. (6.19) becomes

$$\frac{\partial}{\partial a} S[x_a(t)] = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x_a} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_a} \right) \beta\, dt + \left. \frac{\partial L}{\partial \dot{x}_a} \beta \right|_{t_1}^{t_2}. \tag{6.21}$$

But $\beta(t_1) = \beta(t_2) = 0$, so the last term (the "boundary term") vanishes. We now use the fact that $(\partial/\partial a) S[x_a(t)]$ must be zero for *any* function $\beta(t)$, because we are assuming that $x_0(t)$ yields a stationary value. The only way this can be true is if the quantity in parentheses above (evaluated at $a = 0$) is identically equal to zero, that is,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}_0} \right) = \frac{\partial L}{\partial x_0}. \qquad \blacksquare \tag{6.22}$$

[7] Note that nowhere do we assume that x_a and $\dot{x}_a$ are independent variables. The partial derivatives in Eq. (6.18) are very much related, in that one is the derivative of the other. The use of the chain rule in Eq. (6.17) is still perfectly valid.

path $y(x)$. If, for example, $y(x) = 5\,x$ (hence with the boundary conditions $y(0) = 0$ and $y(1) = 5$), we would write

$$I = \int_0^1 (5\,x)^2 dx = \frac{25}{3} x^3 \Big|_0^1 = \frac{25}{3}. \tag{3.14}$$

The argument for I is then a path, an entire function $y(x)$. To make this explicit, we instead write I as

$$I[y(x)] = \int_{x_a}^{x_b} F(x, y(x), y'(x))\, dx \tag{3.15}$$

with square brackets around the argument: I is not a function, but is called a **functional**, traditionally said to be a function of a function.

In general, a functional may take as argument several functions, not just one. But for now let us focus on the case of a functional depending on a single function. The question we want to address is then: how do we make such a functional stationary? This means we are looking for conditions that identify a set of paths $y(x)$ for which the functional $I[y(x)]$ is stationary or "locally horizontal." To do this, we can build upon the simpler example of making stationary a function. For any path $y(x)$, we look at a shifted path

$$y(x) \to y(x) + \delta y(x), \tag{3.16}$$

where $\delta y(x)$ is a function that is small everywhere, but is otherwise arbitrary. However, we require that at the endpoints of the integration in Eq. (3.15), the shifts vanish; *i.e.*, $\delta y(a) = \delta y(b) = 0$. This means that we do not perturb the boundary conditions on trial paths that are fed into $I[y(x)]$, because we only need to find the path that makes stationary the functional amongst the subset of all possible paths that *satisfy* the given fixed boundary conditions at the endpoints. We illustrate this in Figure 3.3. In this restricted set of trial paths, our functional extremization condition now looks very much like (3.11):

$$\delta I[y(x)] \equiv I[y(x) + \delta y(x)] - I[y(x)] = 0, \tag{3.17}$$

to linear order in $\delta y(x)$. We say: "the variation of the functional I is zero." For a function $f(x_1, x_2, \ldots)$, the condition amounted to setting all first derivatives of f to zero. Hence, we need to figure out how to differentiate a functional! Alternatively, we need to expand the functional $I[y(x) + \delta y(x)]$ in $\delta y(x)$ to linear order to identify its "first derivative."

Fortunately, we can deduce all the operations of functional calculus by thinking of a functional in the following way. Imagine that the input to the functional, the path $y(x)$, is evaluated only on a finite discrete set of points:

$$a < x < b \to x = a + n\,\epsilon \le b, \tag{3.18}$$

for n a non-negative integer and ϵ small (see Figure 3.4). In the limit $\epsilon \to 0$ with the upper bound n going to ∞, we recover the original continuum problem.

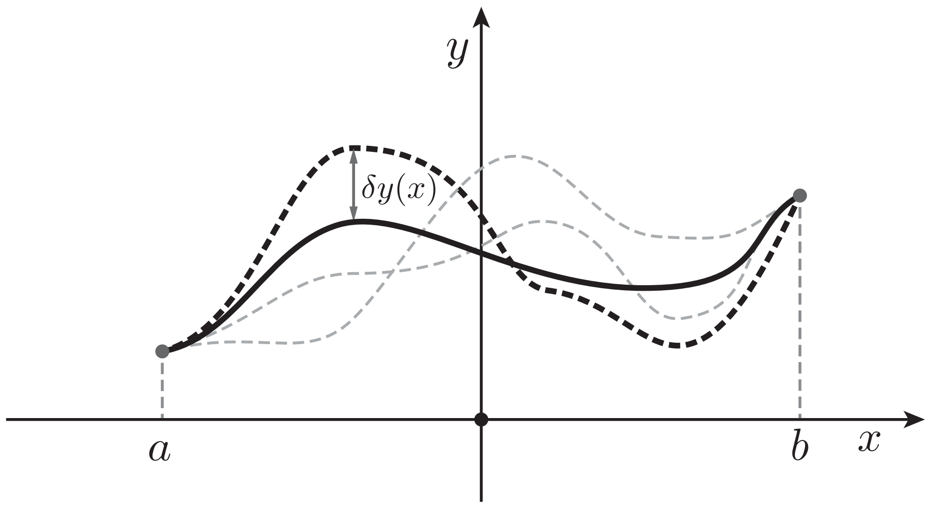

Fig. 3.3 A path $y(x)$ that can be used as input to the functional $I[f(x)]$. We look for that special path from which an arbitrary small displacement $\delta y(x)$ leaves the functional unchanged to linear order in $\delta y(x)$. Note that $\delta y(a) = \delta y(b) = 0$.

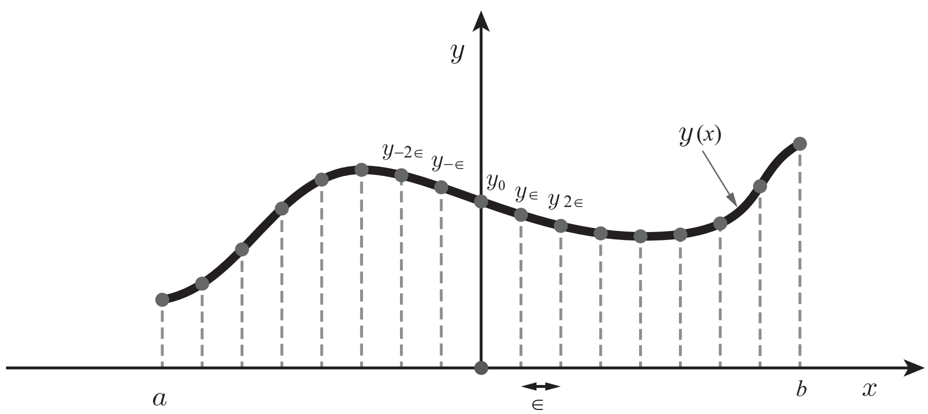

Fig. 3.4 The discretization of a smooth path. In functional calculus, the functional $y(x)$ can be treated as a collection of discrete points.

The functional is simply a function of a finite number of variables $y(a), y(a+\epsilon)$, $y(a+2\,\epsilon), \ldots$. In the limit $\epsilon \to 0$, the set becomes infinitely dense. One can therefore view a functional as a function of an *infinite* number of variables. We can perform all needed operations on I in the discretized regime where I is treated as a function, and then take the $\epsilon \to 0$ limit at the end of the day.

Effectively, we may think of x in $y(x)$ as a discrete index y_x. We then have $I[y(x)] \to I(y_x)$, a function with a large but finite number of variables y_x, with $x \in \{a, \ldots, b\}$ a finite set. A functional then becomes a much more familiar animal: a function. The integral I may also depend upon $y'(x)$, which can be written in discretized form as $y'(x) \to (y_x - y_{x-\epsilon})/\epsilon$ by the definition of the derivative operation. We write it in shorthand as $y'(x) \to y'_x$, and the integration in Eq. (3.15) becomes a sum: $\int dx \to \sum_x \epsilon$. To summarize, we have a discretized form of our original functional:

$$I = \sum_x F(x, y_x, y'_x)\, \epsilon. \tag{3.19}$$

We can now apply the shifts $y_x \to y_x + \delta y_x$, which also implies $y'_x \to y'_x + \delta y'_x$, where $\delta y'_x = (\delta y_x - (\delta y)_{x-\epsilon})/\epsilon = d(\delta y_x)/dx$. We then need the analogue of Eq. (3.11), or

$$\delta f = \frac{\partial f}{\partial x_i}\delta x_i = 0, \tag{3.20}$$

with $f \to I$ and $x_i \to y_x$. Starting from Eq. (3.19), we have

$$\delta I = \sum_x \left(\frac{\partial F}{\partial y_x}\delta y_x + \frac{\partial F}{\partial y'_x}\delta y'_x \right) \epsilon = 0. \tag{3.21}$$

In the $\epsilon \to 0$ limit, we retrieve the integral form

$$\delta I[y(x)] = \int_{x_a}^{x_b} \left(\frac{\partial F}{\partial y(x)}\delta y(x) + \frac{\partial F}{\partial y'(x)}\frac{d}{dx}\left(\delta y(x)\right) \right) dx = 0. \tag{3.22}$$

Integrating the second term by parts, we get

$$\int_{x_a}^{x_b} \frac{\partial F}{\partial y'(x)}\frac{d}{dx}\left(\delta y(x)\right) = \delta y(x)\frac{\partial F}{\partial y'(x)}\bigg|_{x_a}^{x_b} - \int_{x_a}^{x_b} \delta y(x)\frac{d}{dx}\left(\frac{\partial F}{\partial y'(x)}\right) dx, \tag{3.23}$$

where the first term on the right vanishes because we have fixed the endpoints so that $\delta y(a) = \delta y(b) = 0$. Therefore, Eq. (3.22) becomes

$$\delta I[y(x)] = \int_{x_a}^{x_b} \left(\frac{\partial F}{\partial y(x)} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'(x)}\right)\right) \delta y(x)\, dx = 0. \tag{3.24}$$

This integral might be zero because the integrand is zero for all x, or because there are positive and negative portions that cancel one another out. However, since *arbitrary* smooth deviation functions $\delta y(x)$ are permitted, the first alternative has to be the right one. For example, if $a < x_0 < b$ and the integrand happens to be positive from a to x_0 and negative from x_0 to b, so that by cancellation the overall integral is zero, the deviation function $\delta y(x)$ could be changed so that $\delta y(x) = 0$ from x_0 to b, which would force the integral to be positive. Therefore, the requirement that the integral vanishes for *arbitrary* smooth functions $\delta y(x)$ requires that

$$\frac{\partial F}{\partial y(x)} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'(x)}\right) = 0, \tag{3.25}$$

which is known as **Euler's equation**. This equation was worked out by both Euler and Lagrange at around the same time, but we will call it simply "Euler's equation" because we reserve the term "Lagrange equations" for essentially the same equation when used in classical mechanics, as we shall see in Chapter 4.

thereby completing a demonstration of the equivalence between the Newtonian and Lagrangian formulations of mechanics.

- Introduce the Hamiltonian formulation of mechanics and demonstrate its use in several examples.

10.1 Hamilton's Variational Principle: An Example

Hamilton's variational principle states that the integral

$$J = \int_{t_1}^{t_2} L\,dt$$

taken along a path of the possible motion of a physical system is an extremum when evaluated along the path of motion that is the one actually taken. $L = T - V$ is the *Lagrangian* of the system, or the difference between its kinetic and potential energy. In other words, out of the myriad ways in which a system could change its configuration during a time interval $t_2 - t_1$, the actual motion that does occur is the one that either maximizes or minimizes the preceding integral. This statement can be expressed mathematically as

$$\delta J = \delta \int_{t_1}^{t_2} L\,dt = 0 \tag{10.1.1}$$

in which δ is an operation that represents a variation of any particular system parameter by an infinitesimal amount away from that value taken by the parameter when the integral in Equation 10.1.1 is an extremum. For example, the δ that occurs explicitly in Equation 10.1.1 represents a variation in the entire integral about its extremum value. Such a variation is obtained by varying the coordinates and velocities of a dynamical system away from the values actually taken as the system evolves in time from t_1 to t_2, under the constraint that the variation in all parameters is zero at the endpoints of the motion at t_1 and t_2. That is, the variation of the system parameters between t_1 and t_2 is completely arbitrary under the provisos that the motion must be completed during that time interval and that all system parameters must assume their unvaried values at the beginning and end of the motion.

Let us apply Hamilton's variational principle to the case of a particle dropped from rest in a uniform gravitational field. We will see that the integral in Equation 10.1.1 is an extremum when the path taken by the object is the one for which the particle obeys Newton's second law. Let the height of the particle above ground at any time t be denoted by y and its speed by $\dot{y}$. Then δy and $\delta \dot{y}$ represent small, virtual displacements of y and $\dot{y}$ away from the true position and speed of the particle at any time t during its actual motion. (Figure 10.1.1). The potential energy of the particle is mgy, and its kinetic energy is $m\dot{y}^2/2$. The Lagrangian is $L = m\dot{y}^2/2 - mgy$. The variation in the integral of the Lagrangian is given by

$$\delta J = \delta \int_{t_1}^{t_2} L\,dt = \delta \int_{t_1}^{t_2} \left[\frac{m\dot{y}^2}{2} - mgy\right] dt = \int_{t_1}^{t_2} (m\dot{y}\,\delta\dot{y} - mg\,\delta y)\,dt \tag{10.1.2}$$

The variation in the speed can be transformed into a coordinate variation by noting that

$$\delta\dot{y} = \frac{d}{dt}\,\delta y \tag{10.1.3a}$$

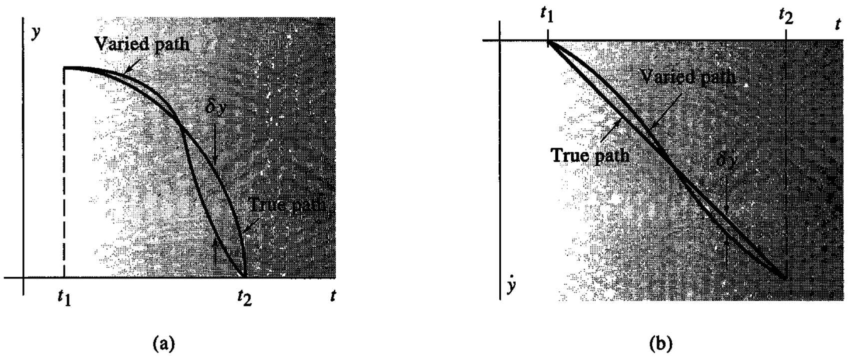

Figure 10.1.1 (a) Variation of the coordinate of a particle from its true path taken in free-fall. (b) Variation in the speed of a particle from the true value taken during free-fall.

Integrating the first term in Equation 10.1.2 by parts gives

$$\int_{t_1}^{t_2} m\dot{y}\,\delta\dot{y}\,dt = \int_{t_1}^{t_2} m\dot{y}\frac{d}{dt}\delta y\,dt = m\dot{y}\,\delta y\Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} m\ddot{y}\,\delta y\,dt \qquad (10.1.3b)$$

The integrated term on the right-hand side is identically zero because the parameters of the admissible paths of motion do not vary at the endpoints of the motion. Hence, we obtain

$$\delta J = \delta\int_{t_1}^{t_2} L\,dt = \int_{t_1}^{t_2}(-m\ddot{y} - mg)\delta y\,dt = 0 \qquad (10.1.4)$$

Because δy represents a completely arbitrary variation of the parameter y away from its true value throughout the motion of the particle (except at the endpoints where the variation is constrained to be zero), the only way in which Equation 10.1.4 can be zero under such conditions is for the term in parentheses to be identically zero at all times. Thus,

$$-mg - m\ddot{y} = 0 \qquad (10.1.5)$$

which, as advertised, is Newton's second law of motion for a particle falling in a uniform gravitational field.

The solution to this equation of motion is $y(t) = -\frac{1}{2}gt^2$ (assuming that the object is dropped from rest at $y_0 = 0$). We now show that any solution that varies from this one yields an integral $J = \int L\,dt$ that is not an extremum. We can represent any possible variation in y away from the true solution by expressing it parametrically as $y(\alpha, t)$ such that

$$y(\alpha,t) = y(0,t) + \alpha\eta(t) \qquad (10.1.6)$$

When the parameter $\alpha = 0$, then $y = y(0, t) = y(t)$, the true solution. $\eta(t)$ is any function of time whose first derivative is continuous on the interval $[t_1, t_2]$ and whose values, $\eta(t_1)$

and $\eta(t_2)$, vanish, thus ensuring that $y(\alpha, t)$ attains its true value at those times regardless of the value of α. Because our choice of $\eta(t)$ is arbitrary, consistent only with the constraints expressed in Equation 10.1.6, the quantity $\alpha\eta(t)$ generates any variation $\delta y(t)$ taken away from the true dynamical path that we wish. An example of a possible variation away from the true dynamical path was depicted in Figure 10.1.1.

The integral J is now a function of the parameter α

$$J(\alpha) = \int_{t_1}^{t_2} L[y(\alpha,t), \dot{y}(\alpha,t); t]\, dt \tag{10.1.7}$$

We now proceed to calculate this integral for the case of the falling body. The expression for $\dot{y}$ in terms of the parameter α is

$$\dot{y}(\alpha, t) = \dot{y}(0, t) + \alpha\dot{\eta}(t) \tag{10.1.8}$$

where $\dot{y}(0,t) = -gt$. The kinetic and potential energies of the falling body are

$$T = \tfrac{1}{2}m\dot{y}^2 = \tfrac{1}{2}m[-gt + \alpha\dot{\eta}(t)]^2 \tag{10.1.9a}$$

$$V = mgy = mg\left[-\tfrac{1}{2}gt^2 + \alpha\eta(t)\right] \tag{10.1.9b}$$

The integral $J(\alpha)$ is, thus,

$$\begin{aligned} J(\alpha) &= \int_{t_1}^{t_2} m\left(\frac{\dot{y}^2}{2} - gy\right) dt \\ &= \int_{t_1}^{t_2} m\left\{g^2t^2 - \alpha g[t\dot{\eta}(t) + \eta(t)] + \tfrac{1}{2}\alpha^2\dot{\eta}^2(t)\right\} dt \end{aligned} \tag{10.1.10}$$

The integral of the term linear in α, in the center square brackets in Equation 10.1.10, is

$$\int_{t_1}^{t_2} [t\dot{\eta}(t) + \eta(t)]\, dt = t\eta(t)\Big|_{t_2}^{t_1} - \int_{t_1}^{t_2} \eta(t)\, dt + \int_{t_1}^{t_2} \eta(t)\, dt \equiv 0 \tag{10.1.11}$$

The first term in Equation 10.1.11 vanishes because $\eta(t_1)$ and $\eta(t_2) = 0$. Thus, the term linear in α completely vanishes, and we obtain

$$J(\alpha) = \tfrac{1}{3}g^2\left(t_2^3 - t_1^3\right) + \tfrac{1}{2}\alpha^2 \int_{t_1}^{t_2} \dot{\eta}^2(t)\, dt \tag{10.1.12}$$

The last term in Equation 10.1.12 is quadratic in the parameter α, and because the integral of $\dot{\eta}^2(t)$ must be positive for any $\dot{\eta}(t)$, $J(\alpha)$ exhibits the behavior pictured in Figure 10.1.2. The value of the integral is a minimum when

$$\left.\frac{\partial J(\alpha)}{\partial \alpha}\right|_{\alpha=0} = 0 \tag{10.1.13}$$

which occurs at $\alpha = 0$.

Even though this result was based on a specific example, it is true for any integral J of a function *of the function* y (and its first derivative) that has the parametric form given by Equation 10.1.6. The resulting $J(\alpha)$ does not depend on α to first order, and its partial derivative vanishes at $\alpha = 0$, making the integral an extremum only when y is equal to the solution obtained from Newton's second law of motion.