

## Chapter 7 The Classical Mechanics of the Special Theory of Relativity

example. A vector  $V$  in two-dimensional space with basis vectors  $e_1$  and  $e_2$  can be written as

$$V = V^1 e_1 + V^2 e_2$$

In general, it is not necessary that any of the basis vectors be normalized ( $e_1 \cdot e_1 \neq 1$ ,  $e_2 \cdot e_2 \neq 1$ ) or that they be orthogonal ( $e_1 \cdot e_2 \neq 0$ ). This means that the magnitude of the scalar product is not conveniently obtained from a simple sum of squares

$$\begin{aligned} V \cdot V &= \sum_{i,j=1}^2 V^i V^j = (V^1)^2 e_1 \cdot e_1 + V^1 V^2 (e_1 \cdot e_2 + e_2 \cdot e_1) + (V^2)^2 e_2 \cdot e_2 \\ &\neq \sum_{i=1}^2 V^i V^i, \end{aligned}$$

and it does not have the value  $\sqrt{(V^1)^2 + (V^2)^2}$ . One way to obtain the magnitude of the vector is to define the **dual space** with basis vectors  $\omega^1$  and  $\omega^2$  (cf. Eq. (7.43)), which have the properties

$$e_1 \cdot \omega^1 = \omega^1 \cdot e_1 = e_2 \cdot \omega^2 = \omega^2 \cdot e_2 = 1$$

and

$$e_1 \cdot \omega^2 = \omega^2 \cdot e_1 = e_2 \cdot \omega^1 = \omega^1 \cdot e_2 = 0.$$

We say that the vector basis,  $e_i$ , is orthonormal to the 1-form basis  $\omega^i$ . The 1-form,  $v$ , corresponding to the vector  $V$  may be written as

$$v = v_1 \omega^1 + v_2 \omega^2.$$

This vector has a (magnitude)<sup>2</sup> of

$$(\text{magnitude})^2 = v \cdot V = V \cdot v = V^1 v_1 + V^2 v_2.$$

When we want to require an object to be expressed in terms of its coordinate basis vectors we will write with a Roman letter (e.g.,  $u$ ) and use Greek letters when it is to be expressed in terms of the basis 1-forms (e.g.,  $\eta$ ). This same approach provides the scalar product of two vectors  $V$  and  $U$  in terms of their associated 1-forms  $v$  and  $u$  as

$$\text{scalar product} = V \cdot u = v \cdot U = u \cdot V = U \cdot v = V^1 u_1 + V^2 u_2 = v_1 U^1 + v_2 U^2.$$

These results are easily generalized to more dimensions, to spaces that have an indefinite metric, and even to more general spaces, such as those discussed in Section 7.11. For example, in a four-dimensional Minkowski space, the 1-form,  $v$ , associated with the vector  $V$ , is  $v_0 = V^0$ ,  $v_1 = -V^1$ ,  $v_2 = -V^2$ ,  $v_3 = -V^3$ , so the squared length of the vector  $V$  is

$$V^0 v_0 + V^1 v_1 + V^2 v_2 + V^3 v_3 = V^0 V^0 - V^1 V^1 - V^2 V^2 - V^3 V^3.$$

The Lorentz transformations can be expressed in terms of the basis vectors. If we let  $x^0, x^1, x^2, x^3$  be the coordinates in a frame  $S$  and  $x^{\alpha'} = x^{\alpha'}(x^0, x^1, x^2, x^3)$  be the transformed coordinates in the frame  $S'$ , then the Lorentz transformation can be written as

$$x^{\alpha'} = L^{\alpha'}_{\beta} x^{\beta} \quad \text{and} \quad x^{\alpha} = L^{\alpha}_{\beta'} x^{\beta'}, \quad (7.50)$$

where  $L^{\alpha}_{\beta'}$  is the inverse transformation of  $L^{\alpha'}_{\beta}$ . The basis vectors transform as

$$e_{\alpha'} = L^{\beta}_{\alpha'} e_{\beta} \quad \text{and} \quad e_{\alpha} = L^{\beta'}_{\alpha} e_{\beta'}. \quad (7.51)$$

Any vector transforms as  $v = v^{\alpha} e_{\alpha} = v^{\beta'} e_{\beta'}$ , so  $\langle \eta, v \rangle = \eta_{\alpha} v^{\alpha} = \eta_{\alpha'} v^{\alpha'}$ . This means that 1-forms transform as  $\eta = \eta_{\alpha} \omega^{\alpha} = \eta_{\alpha'} \omega^{\alpha'}$ , and it follows that

$$\omega^{\alpha'} = L^{\alpha'}_{\beta} \omega^{\beta} \quad \text{and} \quad \omega^{\alpha} = L^{\alpha}_{\beta'} \omega^{\beta'}, \quad (7.52)$$

so

$$v^{\alpha'} = L^{\alpha'}_{\beta} v^{\beta} \quad \text{and} \quad v^{\alpha} = L^{\alpha}_{\beta'} v^{\beta'}, \quad (7.53)$$

and

$$\eta_{\alpha'} = L^{\beta}_{\alpha'} \eta_{\beta} \quad \text{and} \quad \eta_{\alpha} = L^{\beta'}_{\alpha} \eta_{\beta'}. \quad (7.54)$$

To convert vectors, sum on the second (lowered) index of the transformation matrix. To convert 1-forms, sum on the first (raised) index. In tensor notation, vectors are columns, while 1-forms are rows.

Scalars, vectors and 1-forms are simple examples of geometric objects called *tensors*. A tensor is a functional into which we insert  $p$  vectors and  $n$  1-forms to produce a mapping onto a scalar. We describe a tensor by saying that it has a rank given by the numbers  $n$  and  $p$ , where  $n$  is the number of 1-forms insertions possible and  $p$  is the number of possible vector insertions. A tensor,  $Q$ , with  $n$  1-form slots and  $p$  vector slots is written as  $Q$  of rank  $\binom{n}{p}$ . A tensor  $H$  of rank  $\binom{n}{p}$  is a functional into which we can insert  $n$  1-forms  $\sigma, \lambda, \dots, \beta$  and  $p$  vectors  $u, v, \dots, w$  to produce a scalar. For example, the energy momentum vector  $(E/c, \mathbf{p})$  is a tensor of rank  $\binom{1}{0}$ , since contracting it with a 1-form produces a scalar. An example of an ordinary second-rank tensor is the quadrupole tensor of rank  $\binom{2}{0}$ .

Although the components of 1-forms are written with their indices down, the number of 1-form slots is written as the upper of the two numbers used to give the rank of a tensor. This is because in component notation the object generated will have that number of indices to be contracted with 1-forms. For example, if  $S$  is a tensor of rank  $\binom{2}{1}$ ,

$$S(\sigma_{\alpha} \omega^{\alpha}, \lambda_{\beta} \omega^{\beta}, v^{\gamma} e_{\gamma}) = \sigma_{\alpha} \lambda_{\beta} v^{\gamma} S(\omega^{\alpha}, \omega^{\beta}, e_{\gamma}) = S^{\alpha \beta}_{\gamma} \sigma_{\alpha} \lambda_{\beta} v^{\gamma}, \quad (7.55)$$

where the  $S^{\alpha\beta}_{\gamma}$  are called the components of the tensor  $S$  in the chosen coordinate frame. The output of  $S$  is a scalar (see Eq. (7.55)), so if we repeat this calculation in another Lorentz frame, we obtain the transformation law for the tensor components under a coordinate transformation,

$$S^{\alpha'\beta'}_{\gamma'} = S^{\alpha\beta}_{\gamma} L^{\alpha'}_{\alpha} L^{\beta'}_{\beta} L^{\gamma}_{\gamma'}. \quad (7.56)$$

The metric tensor can be used to convert indices from vector to 1-form or 1-form to vector; for example,

$$S^{\alpha}_{\beta\gamma} = g_{\beta\sigma} S^{\alpha\sigma}_{\gamma}. \quad (7.57)$$

Hence, any tensor of rank  $\binom{n}{p}$  can be converted by the metric tensor, without loss of information, to any arrangement of tensor and 1-form indices desired as long as the total number of indices ( $n + p$ ) is conserved. All of these objects are different coordinate forms of the same geometric object (tensor).

Consider our two-dimensional example with a vector,  $u$ , whose components are  $(a, b)$  and a 1-form,  $\sigma$ , with components  $(c, d)$ . If we examine a tensor  $W$  of rank  $\binom{1}{1}$ , then, from Eq. (7.55),

$$W(\sigma, u) = W^{\alpha}_{\beta} \sigma_{\alpha} u^{\beta} = W^0_0 ca + W^0_1 cb + W^1_0 da + W^1_1 db.$$

Physically, by using sets of vectors,  $u$ 's, and 1-forms,  $\sigma$ 's, and measuring the value of the scalar field  $W(\sigma, u)$ , the values of the components of  $W^{\alpha}_{\beta}$  can be determined in one frame. And from Eq. (7.56), specialized to the number and type of components, the values in all inertial frames are known. In a Minkowski space with pseudo-Cartesian coordinates, the components of the tensor  $W$  of rank  $\binom{1}{1}$  can be converted to a corresponding tensor of rank  $\binom{0}{2}$  using the metric tensor in Eq. (7.33)  $\{g_{00} = 1, g_{11} = g_{22} = g_{33} = -1\}$  and the expression in Eq. (7.57) to give the following relations:

$$\begin{aligned} W_{00} &= g_{00} W^0_0 = W^0_0, & W_{01} &= g_{00} W^0_1 = W^0_1, \\ W_{10} &= g_{11} W^1_0 = -W^1_0, & \text{and} & & W_{11} &= g_{11} W^1_1 = -W^1_1. \end{aligned}$$

Given any two vectors, we can construct a second-rank tensor by the operation called *tensor product*,  $T = u \otimes v$ . The tensor product is a machine whose output is a number when the two vectors and the two 1-forms are inserted

$$(u \otimes v)(\sigma, \lambda) = \langle \sigma, u \rangle \langle \lambda, v \rangle. \quad (7.58)$$

The components of the tensor product are

$$T^{\alpha\beta} = u^{\alpha} v^{\beta}. \quad (7.59)$$

In our two-dimensional example of vector  $u$  with components  $(a, b)$  and vector  $v$  with components  $(c, d)$ , Eq. (7.59) becomes written in tensor form

$$(T^{\alpha\beta}) = \begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix}.$$

This process can be continued and could include 1-forms as well as vectors; for example, two vectors ( $\mathbf{u}$ ,  $\mathbf{v}$ ) and a 1-form ( $\sigma$ ) would be written as  $\mathbf{u} \otimes \mathbf{v} \otimes \sigma$ .

Other useful operations include the *gradient*, *contraction*, the *divergence*, and the *wedge product*. First, let us consider the gradient operation. We used  $d$  for the gradient operation on scalars. For a higher-rank tensor, the gradient is often denoted by  $\nabla$ . In three-dimensional Cartesian space,  $\nabla$  is the operator

$$\nabla = i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z},$$

which may also be written as

$$\partial_r = e_1 \frac{\partial}{\partial x^1} + e_2 \frac{\partial}{\partial x^2} + e_3 \frac{\partial}{\partial x^3}.$$

Returning to 4-dimensions, an example of a more general case, let  $\mathbf{S}$  be a  $\binom{0}{3}$  rank tensor, then by definition,  $\nabla \mathbf{S}(\mathbf{u}, \mathbf{v}, \mathbf{w}, \xi) = \partial_\xi \mathbf{S}(\mathbf{u}, \mathbf{v}, \mathbf{w})$  with the vectors  $\mathbf{u}$ ,  $\mathbf{v}$ ,  $\mathbf{w}$  held fixed, and

$$\nabla \mathbf{S}(\mathbf{u}, \mathbf{v}, \mathbf{w}, \xi) = \partial_\xi (S_{\alpha\beta\gamma} u^\alpha v^\beta w^\gamma) = \frac{\partial S_{\alpha\beta\gamma}}{\partial x^\delta} \xi^\delta u^\alpha v^\beta w^\gamma = S_{\alpha\beta\gamma,\delta} \xi^\delta u^\alpha v^\beta w^\gamma. \quad (7.60)$$

That is, the gradient operates only on the coefficients in the definition of the tensor, not on the included vector fields. Since the vectors and 1-forms in Eq. (7.60) are arbitrary and constant, we can rewrite the preceding as

$$\partial_\xi (S_{\alpha\beta\gamma}) = \frac{\partial S_{\alpha\beta\gamma}}{\partial x^\delta} \xi^\delta = S_{\alpha\beta\gamma,\delta} \xi^\delta, \quad (7.60')$$

where the  $\xi^\delta$  define the direction of the gradient, and the last equality shows clearly that the derivative does not operate on the vector given by  $\xi^\delta$ .

In Minkowski spacetime, contracting the energy momentum vector ( $E/c, \mathbf{p}$ ) with the charge-current 1-form ( $\rho c, -\mathbf{J}$ ) produces the scalar ( $E\rho - \mathbf{p} \cdot \mathbf{J}$ ). This idea can be extended to reduce the rank of a tensor by a process called contraction. The contraction operation can be performed on any tensor whose total rank (sum of vector and 1-form indices) is equal to or greater than 2. To do this, enter a basis vector in one slot and the corresponding 1-form basis in another slot and sum over the basis, thereby producing a lower-rank tensor. For example, consider the 4-index tensor whose components are  $R_{\alpha\mu}{}^\beta{}_\nu$ . We can form a two-index tensor by the inserting a basis 1-form into the first slot of the tensor definition, and the related basis vector in the third slot, and summing over the basis set. Formally,

$$R(e_\alpha, u, w^\alpha, v) = M(u, v), \quad (7.61)$$

or in component form

$$M_{\mu\nu} u^\mu v^\nu = R_{\alpha\mu}{}^\alpha{}_\nu u^\mu v^\nu, \quad (7.62)$$

which can be written as

$$M_{\mu\nu} = R_{\alpha\mu}{}^\alpha{}_\nu. \quad (7.62')$$

In three-dimensional Cartesian space, the divergence of a vector  $\mathbf{V}$  is the scalar quantity  $\nabla \cdot \mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}$ , while in 4-dimensional space the 4-divergence is  $\frac{\partial V^\mu}{\partial x^\mu}$ . In Minkowski spacetime the 4-divergence operator is often denoted by the same symbol,  $\nabla$ , in italics, or by  $\square$  whose components are

$$\square_\mu = \nabla_\mu = \omega_\mu^\alpha \frac{1}{\partial x^\alpha},$$

with  $\omega_\mu^\alpha$  the 1-form basic components. For example, the continuity equation in electromagnetic theory is

$$\frac{\partial j^\mu}{\partial x^\mu} = \square \cdot J = \nabla \cdot J = \frac{\partial(\rho c)}{\partial ct} + \nabla \cdot j = \frac{\partial \rho}{\partial t} + \nabla \cdot j = 0.$$

The operator  $\nabla^2$  (sometimes written as  $\square^2$ ) is called the d'Alembertian and is

$$\square^2 = \nabla^2 = \nabla \cdot \nabla = g^{\mu\nu} \frac{\partial}{\partial x^\mu} \frac{\partial}{\partial x^\nu} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right)$$

where the last equality is the expression in Minkowski space with Cartesian coordinates. The 4-divergence operator on tensors reduces the rank of the tensor by 1. For spacetime tensors, the divergence is written as  $\nabla \cdot S$  and, considering as an example a tensor  $S$  with a slot for a 1-form and three vector slots,

$$\square \cdot S(u, v) = \nabla \cdot S(u, v) = \nabla \cdot S(\omega^\alpha, u, v, e_\alpha) = S^\alpha{}_\beta{}_\gamma{}_\alpha u^\beta v^\gamma. \quad (7.63)$$

That is, the gradient of Eq. (7.60) is taken along a basis direction, and then a contraction is formed between this direction and one of the 1-form slots in the tensor. In component form, this reduces to

$$\nabla_\alpha S^\alpha{}_\beta{}_\gamma = S^\alpha{}_\beta{}_\gamma{}_\alpha. \quad (7.63')$$

The final tensor operator we need is the *wedge product*, also called the *bivector* or *biform*, which is

$$u \wedge v = u \otimes v - v \otimes u, \quad (7.64)$$

where the tensor product,  $\otimes$ , was defined in Eq. (7.58). The wedge product is an antisymmetric vector product. In component form, Eq. (7.64) becomes

$$(u \wedge v)^{\alpha\beta} = u^\alpha v^\beta - v^\alpha u^\beta. \quad (7.64')$$

Successive  $\wedge$  operations can be strung together just like the  $\otimes$  operator. The wedge product is useful whenever we deal with antisymmetric expressions. In particular, when we look at the electromagnetic field in the next section, we will discover that the fundamental field tensor, called *Faraday*, can be expressed in terms of the wedge product.

Consider the two-dimensional example used previously, where  $\mathbf{u} = u^1 \mathbf{e}_1 + u^2 \mathbf{e}_2$  and  $\mathbf{v} = v^1 \mathbf{e}_1 + v^2 \mathbf{e}_2$ . The wedge product in Eq. (7.64') has components  $W = u \wedge v$  given by

$$W = \begin{pmatrix} u^1 v^1 - v^1 u^1 & u^1 v^2 - u^2 v^1 \\ u^2 v^1 - v^2 u^1 & u^2 v^2 - v^2 u^2 \end{pmatrix} = \begin{pmatrix} 0 & u^1 v^2 - v^1 u^2 \\ u^2 v^1 - v^2 u^1 & 0 \end{pmatrix}.$$

Although the examples given above assumed a certain combination of 1-form slots and vector slots, we must stress that the metric tensor can be used to produce a tensor with indices in any desired position.

## 7.6 ■ FORCES IN THE SPECIAL THEORY; ELECTROMAGNETISM

The preceding material has been concerned with the kinematics of the special theory. The dynamics of the theory follows from the assumption that Newton's laws are correct for objects at rest in the rest frame of the observer, nearly correct for objects moving slowly relative to the speed of light, and require generalizations to covariant equations. The correct generalization of the three-velocity to the four-velocity was given in Eq. (7.27). So we must generalize the force law,

$$\mathbf{F}' = \frac{d(m\mathbf{v}')}{dt}, \quad (7.65)$$

to a covariant form.

Since Maxwell's equations are assumed to be a correct description, we shall briefly consider a covariant reformulation of electromagnetic theory as a guide for the correct form of the force laws of mechanics. The vector and scalar electromagnetic potentials form a four-vector  $A^\mu = (\phi/c, \mathbf{A})$ . If the potentials satisfy the Lorentz condition (in SI units), which is the vanishing of the four-divergence of the electromagnetic potential 4-vector,

$$\square \cdot A = \nabla \cdot A = \frac{\partial A^\mu}{\partial x^\mu} = \nabla \cdot \mathbf{A} + \mu_0 \epsilon_0 \frac{\partial \phi}{\partial t} = 0, \quad (7.66)$$

they separately satisfy the wave equations of the form (where  $\mu_0 \epsilon_0 = 1/c^2$ )

$$\square^2 \mathbf{A} = \nabla^2 \mathbf{A} = \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} = \mu_0 \mathbf{j} \quad (7.67a)$$

for the space components and for the time component

$$\square^2 \phi = \nabla^2 \phi = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = \frac{\rho}{\epsilon_0}. \quad (7.67b)$$

In terms of  $\phi$  and  $\mathbf{A}$ , the Lorentz force is  $\mathbf{F} = e\{-\nabla\phi + \frac{1}{c^2} \frac{\partial \mathbf{A}}{\partial t} + \frac{1}{c} [\mathbf{v} \times (\nabla \times \mathbf{A})]\}$ . This suggests that we should generalize the Lorentz force law to

$$\frac{dp_\mu}{d\tau} = e \left( \frac{\partial(u^\nu A_\nu)}{\partial x^\mu} - \frac{dA_\mu}{d\tau} \right). \quad (7.68)$$

For the three-momentum,  $\mathbf{p}_3$ , and three-velocity,  $\mathbf{v}$ , Eq. (7.68) becomes

$$\frac{dp_3}{dt} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \quad (7.68')$$

with  $\mathbf{E}$  the electric field,  $\mathbf{B}$  the magnetic field, and  $e$  the electric charge. The geometric approach is to define a tensor  $F$ , named **Faraday**, whose components will be the electromagnetic field tensor and write, with  $u$  the 4-velocity,

$$\frac{dp}{d\tau} = eF(u). \quad (7.69)$$

In component notation, this becomes

$$\frac{dp^\mu}{d\tau} = eF^\mu{}_\beta u^\beta. \quad (7.70)$$

This produces Maxwell's equations, provided (according to Eq. (7.68))  $F^\alpha{}_\beta$  is given by

$$F^\alpha{}_\beta = \begin{pmatrix} 0 & E_x & E_y & E_z \\ E_x & 0 & cB_z & -cB_y \\ E_y & -cB_z & 0 & cB_x \\ E_z & cB_y & -cB_x & 0 \end{pmatrix}. \quad (7.71)$$

In Minkowski space, the indices are raised and lowered by the metric tensor (Eq. (7.33)), so

$$F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -cB_z & cB_y \\ E_y & cB_z & 0 & -cB_x \\ E_z & -cB_y & cB_x & 0 \end{pmatrix}. \quad (7.71')$$

and

$$F_{\alpha\beta} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{pmatrix}. \quad (7.71'')$$

The Faraday tensor can be written in at least two different ways using either the tensor product, Eq. (7.58), or the wedge product, Eq. (7.64), as

$$\mathbf{F} = F_{\alpha\beta} dx^\alpha \otimes dx^\beta = \frac{1}{2} F_{\alpha\beta} dx^\alpha \wedge dx^\beta.$$

The latter expression explicitly shows the antisymmetry.

We can write Maxwell's equation in their normal component form using geometric notation:

$$\nabla F = 0 \quad \text{and} \quad \nabla \cdot F = J, \quad (7.72)$$

where  $J$  is the 4-current density with components  $(\rho c, \mathbf{j})$ , where  $\rho$  is the charge density and  $\mathbf{j}$  is the three-current density. The first of these equations produces (using three-dimensional notation)  $\nabla \cdot \mathbf{B} = 0$  and  $\partial \mathbf{B} / \partial t + \nabla \times \mathbf{E} = 0$ , while the second gives  $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$  and  $(1/c^2) \partial \mathbf{E} / \partial t - \nabla \times \mathbf{B} = -\mu_0 \mathbf{j}$ .

Following the guide provided by the covariant formulation of electromagnetic theory, the proper generalization of Newton's second law, Eq. (7.65), is

$$\frac{dp^\mu}{d\tau} = K^\mu, \quad (7.73)$$

where  $K^\mu$  is a 4-vector force, known as the *Minkowski force*. The spatial components of  $K^\mu$  are not the components of the force in Eq. (7.65), but rather they are quantities that reduce to the  $F^i$  as  $\beta \rightarrow 0$ . The exact form clearly results from the Lorentz transformation properties of the forces present. Some aspects of the 4-force are listed in Table 7.1.

The general question (which cannot be uniquely resolved) is, How do we find the proper relativistic expression for force? Electromagnetism is used to justify the special theory, so we should expect no problem with it. As we saw in the previous paragraphs, this is trivial for electromagnetic forces because the special theory and the Lorentz transformations are constructed to make Maxwell's electromagnetic theory covariant. For example, the electromagnetic force is given by Eq. (7.68) as

$$K_\mu = -q \left( \frac{\partial u_\nu A^\nu}{\partial x^\mu} - \frac{dA_\mu}{d\tau} \right), \quad (7.74)$$

with  $q$  the charge on the particles and  $A_\mu$  the components of the four-potential given by  $(\phi/c, \mathbf{A})$ . Note that  $\phi$  is the scalar potential and  $\mathbf{A}$  is the three-dimensional electromagnetic vector potential. So the ordinary force,  $F_i$ , and the spatial component of the Minkowski electromagnetic force,  $K_i$ , are related by

$$F_i = K_i \sqrt{1 - \beta^2}. \quad (7.75)$$

What about other forces? Two methods are commonly used to deduce acceptable transformation properties of forces and hence the correct relativistic form of the forces.

The first method is to argue that there are only four fundamental forces in nature—gravitational, weak nuclear, electromagnetic, and strong nuclear. A cor-

rect relativistic theory must provide valid expressions for these four forces. These expressions, if stated in covariant form, will automatically provide the transformation properties of the forces. In this approach, since we understand electromagnetic forces, it remains to find expressions for the other three fundamental forces in a covariant form in some frame and assume this is correct in all inertial frames. It is assumed the transformations involve no terms that vanish in the chosen frame; for example, there is no need to arbitrarily add terms proportional to  $(v/c)^3$ . This program has been carried out for two of the remaining three forces (weak nuclear and strong nuclear) and for weak gravitational forces. It fails completely for strong gravitational effects. It is beyond the scope of the present text to probe more deeply into this question.

The second approach of determining the correct relativistic force is to simply define force as being the time rate of change of the momentum. Then we write

$$\frac{dp_i}{dt} = F_i \quad (7.76)$$

where the  $p_i$  in Eq. (7.76) is some relativistic generalization of the Newtonian momentum that reduces to  $mv_i$  in the limit of small  $\beta$ . The simplest generalization is the one given in Eq. (7.36). This second approach has thus far failed to produce any results other than those predicted by the first approach.

## 7.7 ■ RELATIVISTIC KINEMATICS OF COLLISIONS AND MANY-PARTICLE SYSTEMS

The formulations of the previous sections enable us to generalize relativistically the discussion of Section 3.11 on the transformation of collision phenomena between various systems. The subject is of considerable interest in experimental high-energy physics. While the forces between elementary particles are only imperfectly known, and are certainly far from classical, so long as the particles involved in a reaction are outside the region of mutual interaction their mean motion can be described by classical mechanics. Further, the main principle involved in the transformations—conservation of the four-vector of momentum—is valid in both classical and quantum mechanics. The actual collision or reaction is taken as occurring at a point—or inside a very small black box—and we look only at the behavior of the particles before and after.

Because of the importance to high-energy physics, this aspect of relativistic kinematics has become an elaborately developed field. It is impossible to give a comprehensive discussion here. All that we can do is provide some of the important tools, and cite a few simple examples that may illustrate the flavor of the techniques employed. Although many collision experiments involve colliding beams, we shall, for simplicity, confine our attentions to problems where one of the particles is at rest in the laboratory frame. The generalization to both particles moving in the laboratory frame is straightforward.

The notion of a point designated as the center of mass obviously presents difficulties in a Lorentz-invariant theory. But the center-of-mass system can be suitably generalized as the Lorentz frame of reference in which the total spatial linear momentum of all particles is zero. That such a Lorentz frame can always be found follows from the theorem that the total momentum 4-vector is timelike for a system of mass points.

One such frame is the center-of-momentum frame. This is a frame in which the components of the spatial momentum of the initial particles add to zero. Such a frame obviously exists. Let us define  $E$  and  $\mathbf{p}$  in Eq. (7.36) to be

$$E = \sum_{i=1}^n E_i \quad \text{and} \quad \mathbf{p} = \sum_{i=1}^n \mathbf{p}_i \quad (7.77)$$

where the sum is over the particles involved. The left-hand side of Eq. (7.38) becomes

$$\sum_{r,s} m_r m_s c^2 - \sum_{r,s} m_r m_s \gamma_r \gamma_s (\mathbf{v}_r \cdot \mathbf{v}_s). \quad (7.78)$$

This clearly is positive (*hint*: separate the negative terms in which  $r = s$ ), so it is possible to find a frame in which the three-momentum,  $\mathbf{p}$ , equals zero. The Lorentz system, in which the spatial components of the total momentum are zero, is termed the *center-of-momentum system*, or more loosely, and somewhat incorrectly, as the center-of-mass system, and will be designated by the abbreviation "C-O-M system."

As an example, let us consider a particle of mass  $m_1$  and momentum  $p^1$  in the  $x$ -direction, which suffers a head-on collision with a particle of mass  $m_2$  at rest in an experimenter's frame (called the laboratory frame). The initial 4-momentum is

$$p^\mu = ([m_1 \gamma + m_2]c, m_1 \gamma v^1, 0, 0). \quad (7.79)$$

The length squared of momentum has the magnitude

$$p^\mu p_\mu = (m_1^2 + m_2^2 + 2m_1 \gamma m_2)c^2. \quad (7.79')$$

When components are given, we shall follow the practice of denoting the primed frame by primes on the indices. The two particles are denoted by subscripts 1 and 2 respectively.

In the C-O-M system, the total momentum is

$$([m_1 \gamma'_1 + m_2 \gamma'_2]c, 0, 0, 0), \quad (7.80)$$

since by definition the space part of the momentum vanishes,

$$m_1 \gamma'_1 \beta'_1 c + m_2 \gamma'_2 \beta'_2 c = 0, \quad (7.81)$$

where  $\beta'_1$  and  $\beta'_2$  are the velocities of  $m_1$  and  $m_2$ , respectively, in the C-O-M frame.

The boost,  $\beta'$ , needed to go from the laboratory to the C-O-M frame, has the value

$$\beta'_2 = -\beta'. \quad (7.81')$$

Since all velocities are parallel, the velocity addition formula Eq. (7.15) gives the velocity  $\beta'_1$  of mass  $m_1$  in the C-O-M system in terms of  $\beta'$  and its velocity  $\beta = v/c$  in the laboratory frame,

$$\beta'_1 = \frac{\beta - \beta'}{1 - \beta\beta'}. \quad (7.82)$$

The total squared momentum in the C-O-M frame given in Eq. (7.80) can be rewritten using the results of Eqs. (7.81) and (7.82) as

$$p^\mu p_\mu = \frac{m_2^2 \beta^2 (1 - \beta'^2) c^2}{\beta - \beta'}. \quad (7.83)$$

Equating Eqs. (7.79') and (7.83) gives a single equation that can be solved for the boost velocity  $\beta'$ . There are two real roots, one of which corresponds to the physically meaningful case of  $\beta' < 1$ .

Since the spatial momentum in the C-O-M frame is zero, there is clearly more energy,  $p^0$ , in this frame than in the laboratory frame.\* The excess energy in the C-O-M frame,  $\Delta E$ , is obtained by subtracting the time component of Eq. (7.79) from the time component of Eq. (7.80).

The total momentum four vector is conserved, which automatically implies both conservation of spatial linear momentum and conservation of total energy (including rest mass energy). Our major tools for making use of the conservation principle are Lorentz transformations to and from the C-O-M system, and the formation of Lorentz invariants (world scalars) having the same value in all Lorentz frames. Since energy and momentum are combined into one conservation law, the relativistic results are more easily obtained than the nonrelativistic results of previous chapters. The transformations between laboratory system and C-O-M system are merely special cases of the Lorentz transformation.

As an example of the use of Lorentz invariants, let us consider a reaction initiated by two particles that produces another set of particles with masses  $m_r$ ,  $r = 3, 4, 5, \dots$ . In the C-O-M system, the transformed total momentum is

$$P^\mu = (E'/c, 0, 0, 0). \quad (7.84)$$

It is often convenient to look on the C-O-M system as the proper (or rest) system of a composite mass particle of mass  $M = E'/c^2$ .† The square of the magnitude of

\*For a single particle, the energy has a minimum value,  $mc^2$ , in the rest frame. The C-O-M frame is not the rest frame of either particle.

†Although it is customary in high-energy physics to use units in which  $c = 1$ , it seems more helpful in an introductory exposition such as this to retain the powers of  $c$  throughout.

$P$  must be invariant in all Lorentz systems and conserved in the reaction. Hence, we have

$$P_\mu P^\mu = P_{\mu'} P^{\mu'} = \frac{E'^2}{c^2} = M^2 c^2. \quad (7.85)$$

But for the initial particles,  $P_\mu P^\mu$  can be evaluated as

$$P_\mu P^\mu = (m_1^2 + m_2^2)c^2 - 2p_{1\mu} p_2^\mu. \quad (7.86)$$

The energy in the C-O-M system, or equivalent mass  $M$ , is therefore given in terms of the incident particles as

$$E'^2 \equiv M^2 c^4 = (m_1^2 + m_2^2)c^4 + 2(E_1 E_2 - c^2 \mathbf{p}_1 \cdot \mathbf{p}_2). \quad (7.87)$$

Suppose now that, one particle, say 2, was initially stationary in the laboratory system. Since then  $\mathbf{p}_2 = 0$  and  $E_2 = m_2 c^2$ , the C-O-M energy becomes

$$E'^2 \equiv M^2 c^4 = (m_1^2 + m_2^2)c^4 + 2m_2 c^2 E'_1. \quad (7.88)$$

If the excess of  $E_1$  over the rest mass energy be denoted by  $T_1$ , [cf. Eq. (7.39)] that is, the kinetic energy, this can be written

$$E'^2 \equiv M^2 c^4 = (m_1 + m_2)^2 c^4 + 2m_2 c^2 T_1. \quad (7.89)$$

It is clear that the available energy in the C-O-M system increases only slowly with incident kinetic energy. Even in the “ultrarelativistic” region, where the kinetic energy of motion is very large compared to the rest mass energy,  $E'$  increases only as the square root of  $T_1$ .

The effect of the proportionally small amount of incident energy available in the C-O-M system is shown dramatically in terms of the threshold energies. It is obvious that the lowest energy at which a reaction (other than elastic scattering) is possible is when the reaction products are at rest in the C-O-M system. Any finite kinetic energy requires a higher  $E'$  or equivalently higher incident energy. The total four-momentum in the C-O-M system after the reaction, denoted by  $P^{\mu''}$  has the magnitude at threshold given by

$$P_{\mu''} P^{\mu''} = c^2 \left( \sum_r m_r \right)^2, \quad (7.90)$$

which, by conservation of momentum, must be the same as Eq. (7.85). For a stationary target, the incident energy of motion as threshold is then given as a consequence of Eq. (7.89) by

$$\frac{T_1}{m_1 c^2} = \frac{\left( \sum_r m_r \right)^2 - (m_1 + m_2)^2}{2m_1 m_2}.$$

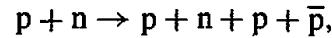
If the  $Q$  value of the reaction is defined as\*

$$Q = \left[ \sum_r m_r - (m_1 + m_2) \right] c^2, \quad (7.91)$$

this threshold energy becomes

$$\frac{T_1}{m_1 c^2} = \frac{Q^2 + 2Q(m_1 + m_2)c^2}{2m_1 m_2 c^4}. \quad (7.92)$$

A common illustration of the application of Eq. (7.92) is the historic production of an antiproton,  $\bar{p}$ , by the reaction, involving a proton  $p$ ,



where  $n$  is a nucleon, either neutron or proton. The masses of all particles involved are nearly equal at 938 MeV equivalent rest mass energy and we select  $Q = 2mc^2$ . Equation (7.92) then says that the incident particle kinetic energy at threshold must be

$$T_1 = 6mc^2 = 5.63 \text{ GeV},$$

which is 3 times the energy represented by  $Q$ ! If, however, the reaction was initiated by two nucleons incident on each other with equal and opposite velocity, then the laboratory system is the same as the C-O-M system. All of the kinetic energy is available in this case to go into production of the proton-antiproton pair, and each of the incident particles at threshold need have a kinetic energy of motion equivalent to only the mass of one proton, 938 MeV. It is no wonder so much effort has been put into constructing colliding beam machines!

Another instructive example of a threshold calculation is photomeson production, say, by the reaction

$$\gamma + p \rightarrow \Sigma^0 + K^+, \quad (7.93)$$

where  $\gamma$  stands for an incoming photon. For the purposes of classical mechanics, a photon is a zero-mass particle with spatial momentum  ${}^0\mathbf{p}$  and energy  ${}^0pc$ .<sup>†</sup> In calculating  $Q$ , the mass  $m_1$  of the photon is zero:

$$Q = (m_{\Sigma^0} + m_{K^+} - m_p)c^2 = 749 \text{ MeV}.$$

\* $Q$  here has the opposite sign to the convention adopted in Eq. (3.112).

<sup>†</sup>The square of the magnitude of the photon momentum four-vector is zero, so the vector can be described as "lightlike." The C-O-M theorem is imperiled only if all of the particles are photons, and even then only if the photons are going in the same direction.

Equation (7.92) is rewritten for a reaction involving an incident photon as

$$T_1 = {}^0 p c = \frac{Q^2 + 2Qm_2c^2}{2m_2c^2}.$$

From the value of  $Q$  and the rest mass energy  $m_2$  of the proton, the threshold energy for the reaction Eq. (7.93) is then

$$T_1 = 1.05 \text{ GeV},$$

which is only slightly higher than  $Q$ .

We can also easily find the energy of the reaction products in the laboratory system at threshold. The C-O-M system is the rest system for the mass  $M$ , with  $P^{0'} = Mc$ . In any other system, the zeroth component of the 4-vector is  $P^0 = Mc\gamma$ . But in the laboratory system

$$P^0 = \frac{1}{c}(E_1 + E_2) = \frac{1}{c}(E_1 + m_2c^2),$$

where the last form holds only for a stationary target particle. Hence, the C-O-M system moves relative to the laboratory system such that

$$\gamma = \frac{E_1 + m_2c^2}{Mc^2}. \quad (7.94)$$

But at threshold all the reaction products are at rest in the C-O-M system so that  $M = \sum_r m_r$ , and therefore

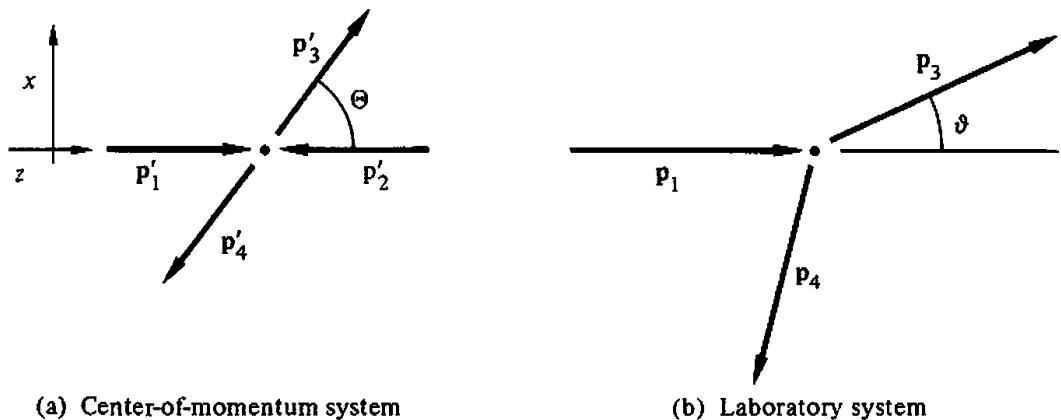
$$\gamma = \frac{T_1 + (m_1 + m_2)c^2}{\sum_r m_r c^2} \quad (\text{threshold}). \quad (7.95)$$

The kinetic energy of the  $s$ th reaction product in the laboratory system is then

$$T_s = m_s c^2(\gamma - 1). \quad (7.96)$$

Thus, the antiproton at threshold has a kinetic energy  $T_{\bar{p}} = mc^2 = 938 \text{ MeV}$ . In contrast, the  $K^+$  meson emerges at threshold with 494 MeV.

In Section 3.11, the kinematic transformations of a two-body nonrelativistic collision were investigated. Eq. (3.117') gives the reduction in energy of an incident particle after elastic scattering from a stationary target, as a function of the scattering angle in the C-O-M system. The derivation of the relativistic analog provides another interesting example of the methods of relativistic kinematics. Use of Lorentz invariants here is not particularly helpful; instead direct Lorentz transformations are made between the laboratory and C-O-M systems. Figure 7.4 illustrates the relations of the incident and scattered *spatial* momentum vectors in both systems. The incident and scattered momentum vectors define a plane, in-



**FIGURE 7.4** Momentum vectors for relativistic elastic scattering in C-O-M and laboratory Lorentz frames.

variant in orientation under Lorentz transformation, here taken to be the  $xz$  plane with the incident direction along the  $z$  axis. Because the collision is elastic, the masses of the incident particle,  $m_1$ , and of the stationary target,  $m_2$ , remain unchanged; that is,  $m_3 = m_1$ ,  $m_4 = m_2$ . Primes on the vectors denote C-O-M values, unprimed vectors are in the laboratory system. To distinguish clearly between before and after the scattering, the indexes 3 and 4 will be retained for the vectors *after* scattering. We have only to remember that 3 denotes the scattered incident particle, and 4 the recoiling target particle. Components of the separate particle 4-vectors will always have two indices: the first for the particle, the second for the component.

The Lorentz transformation from the laboratory to the C-O-M system is defined by the  $\gamma$  of Eq. (7.94) with  $M$  given by Eq. (7.89):

$$\gamma = \frac{E_1 + m_2 c^2}{\sqrt{2m_2 c^2 E_1 + (m_1^2 + m_2^2)c^4}} = \frac{T_1 + (m_1 + m_2)c^2}{\sqrt{2m_2 c^2 T_1 + (m_1 + m_2)^2 c^4}}. \quad (7.97)$$

The quantity  $\beta$  can be found from  $\gamma$ , or more directly by arguments similar to those used to obtain  $\gamma$ . In the C-O-M system, the spatial part of the total momentum four-vector is zero; in any other system, the spatial part is  $Mc\beta\gamma$ . However, in the laboratory system the spatial part is  $\mathbf{p}_1$ . Hence, by Eq. (7.94)  $\beta$  must be given as

$$\beta = \frac{\mathbf{p}_1 c}{E_1 + m_2 c^2} = \frac{\mathbf{p}_1 c}{T_1 + (m_1 + m_2)c^2}. \quad (7.98)$$

Because  $\beta$  is along the  $z$  axis, the Lorentz transformation takes (with  $\beta_x = \beta_y = 0$ ) the form given by Eq. (7.11), and the components of  $p_1^\mu$  in the C-O-M system are given by

$$\begin{aligned} p'_1 &= p_1^{3'} = \gamma \left( p_1 - \frac{\beta E_1}{c} \right) \\ \frac{E'_1}{c} &= p_1^{0'} = \gamma \left( \frac{E_1}{c} - \beta p_1 \right). \end{aligned} \quad (7.99)$$

After the collision,  $\mathbf{p}'_3$  is no longer along the  $z$  axis, but since the collision is elastic, its magnitude is the same as that of  $\mathbf{p}'_1$ . If  $\Theta$  is the angle between  $\mathbf{p}'_3$  and the incident direction, as in Section 3.11, then the components of  $\mathbf{p}'_3$  in the C-O-M system are

$$p_3^{1'} = p'_1 \sin \Theta, \quad p_3^{3'} = p'_1 \cos \Theta, \quad p_3^{0'} = p_1^{0'} = \frac{E'_1}{c}. \quad (7.100)$$

The transformation back to the laboratory system is the same Lorentz transformation but with relative velocity  $-\beta$ . Hence, the components of  $\mathbf{p}_3$  are

$$\begin{aligned} p_3^1 &= p_3^{1'} = p'_1 \sin \Theta \\ p_3^3 &= \gamma(p_3^{3'} - \beta p_3^{0'}) = \gamma \left( p'_1 \cos \Theta + \frac{\beta E'_1}{c} \right) \\ p_3^0 &= \gamma(p_3^{0'} + \beta p_3^{3'}) = \gamma \left( \frac{E'_1}{c} + \beta p'_1 \cos \Theta \right). \end{aligned} \quad (7.101)$$

If  $E'_1$  and  $p'_1$  are substituted in the last of Eqs. (7.101), from Eqs. (7.99) we obtain, after a little simplification, an expression for the energy of the scattered particle in terms of its incident properties:

$$E_3 = E_1 - \gamma^2 \beta (1 - \cos \Theta) (p_1 c - \beta E_1). \quad (7.102)$$

In Eq. (7.102),  $\gamma$  and  $\beta$  must be expressed terms of the incident quantities through Eqs. (7.97) and (7.98), resulting in the relation

$$\gamma^2 \beta (p_1 c - \beta E_1) = \frac{m_2 p_1^2 c^2}{2m_2 E_1 + (m_1^2 + m_2^2)c^2}. \quad (7.103)$$

With the help of the relation between  $p_1$  and  $E_1$ , Eq. (7.38'), this can be written

$$\gamma^2 \beta (p_1 c - \beta E_1) = \frac{m_2 T_1 (T_1 + 2m_1 c^2)}{2m_2 T_1 + (m_1 + m_2)^2 c^2}. \quad (7.104)$$

Some further algebraic manipulation then enables us to rewrite Eq. (7.102) as

$$\frac{T_3}{T_1} = 1 - \frac{2\rho(1 - \mathcal{E}_1/2)}{(1 + \rho)^2 + 2\rho\mathcal{E}_1} (1 - \cos \Theta), \quad (7.105)$$

where  $\rho = m_1/m_2$ , as in Section 3.11 for elastic scattering, and  $\mathcal{E}_1$  is the kinetic energy of the incident particle in units of the rest mass energy,

$$\mathcal{E}_1 = \frac{T_1}{m_1 c^2}. \quad (7.106)$$

Equation (7.105) is the relativistic counterpart of Eq. (3.117'). It is easy to see that Eq. (7.105) reduces to the nonrelativistic case as  $\mathcal{E}_1 \rightarrow 0$ , and that if  $\rho = 1$  (equal masses), the relativistic corrections cancel completely. Equation (7.105) implies that the minimum energy after scattering, in units of  $m_1 c^2$ , is given by

$$(\mathcal{E}_3)_{\min} = \mathcal{E}_1 \frac{(1 - \rho)^2}{(1 + \rho)^2 + 2\rho\mathcal{E}_1}. \quad (7.107)$$

In the nonrelativistic limit, the minimum fractional energy after scattering is

$$\frac{(\mathcal{E}_3)_{\min}}{\mathcal{E}_1} = \left( \frac{1 - \rho}{1 + \rho} \right)^2; \quad \mathcal{E}_1 \ll 1, \quad (7.108)$$

which is a well-known result, easily obtained from Eq. (3.117'). Equation (7.108) says that in the nonrelativistic region a particle of mass  $m_1$  cannot lose much kinetic energy through scattering from a much heavier particle, that is, when  $\rho \ll 1$ , which clearly agrees with common sense. However, in the ultrarelativistic region, when  $\rho\mathcal{E}_1 \gg 1$ , the minimum energy after scattering is independent of  $\mathcal{E}_1$ :

$$(T_3)_{\min} = \frac{(m_2 - m_1)^2 c^2}{2m_2}; \quad \rho\mathcal{E}_1 \gg 1. \quad (7.109)$$

Since the condition on  $\mathcal{E}_1$  is equivalent to requiring  $T_1 \gg m_2 c^2$ , it follows from Eq. (7.109) that such a particle can lose a large fraction of its energy even when scattered by a much heavier particle. This behavior is unexpected, but it should be remembered that for particles at these energies, traveling very close to the speed of light, even a slight change in velocity corresponds to a large change in energy.

Finally, we may easily obtain the relation between the scattering angles in the C-O-M and laboratory system by noting that (first index particle, second component)

$$\tan \vartheta = \frac{p_{31}}{p_{33}} = \frac{\sin \Theta}{\gamma \left( \cos \Theta + \frac{\beta E'_1}{p'_1 c} \right)}. \quad (7.110)$$

By Eq. (7.36),

$$\frac{p'_1 c}{E'_1} = \frac{v'_1}{c} \equiv \beta'_1, \quad (7.111)$$

so that  $\tan \vartheta$  can also be written

$$\tan \vartheta = \frac{\sin \Theta}{\gamma(\cos \Theta + \beta/\beta'_1)}. \quad (7.112)$$

In terms of initial quantities, Eqs. (7.99) show that

$$\frac{\beta E'_1}{p'_1 c} = \frac{\beta \left( \frac{E_1}{c} - \beta p_1 \right)}{p_1 - \frac{\beta E_1}{c}}. \quad (7.113)$$

This can be further reduced by employing the relations (cf. Eq. (7.98))

$$\frac{\beta}{p_1 - \frac{\beta E_1}{c}} = \frac{1}{m_2 c}. \quad (7.114)$$

$$E_1 - \beta p_1 = \frac{m_1(m_1 + m_2)c^4 + m_2c^2T_1}{(m_1 + m_2)c^2 + T_1}. \quad (7.115)$$

The final expression for  $\tan \vartheta$  can then be written as

$$\tan \vartheta = \frac{\sin \Theta}{\gamma [\cos \Theta + \rho g(\rho, \mathcal{E}_1)]}, \quad (7.116)$$

where  $g(\rho, \mathcal{E}_1)$  is the function

$$g(\rho, \mathcal{E}_1) = \frac{1 + \rho(1 + \mathcal{E}_1)}{(1 + \mathcal{E}_1) + \rho}, \quad (7.117)$$

and  $\gamma$ , by Eq. (7.97), takes the form

$$\gamma(\rho, \mathcal{E}_1) = \frac{1 + \mathcal{E}_1 + \rho}{\sqrt{(1 + \rho)^2 + 2\rho\mathcal{E}_1}}. \quad (7.118)$$

Again, in the nonrelativistic region,  $\gamma$  and  $g$  tend to unity, and Eq. (7.116) reduces to Eq. (3.107). The correction function  $g(\rho, \mathcal{E}_1)$  never really amounts to much, approaching the constant limit  $\rho$  as  $\mathcal{E}_1$  becomes very large. The important factor affecting the transformed angle is  $\gamma$ , which of course increases indefinitely as  $\mathcal{E}_1$  increases. It does not affect the bounds of the angular distribution, when  $\Theta = 0$  or  $\pi$ , but its presence means that at other angles  $\vartheta$  is always smaller than it would be nonrelativistically. The Lorentz transformation from C-O-M to the laboratory system, which does not affect the transverse component of the momentum, thus always tends to distort the scattered angular distribution into the forward direction.

## 7.8 ■ RELATIVISTIC ANGULAR MOMENTUM

In Chapter 1, it was proven that the nonrelativistic angular momentum obeys an equation of motion much like that for the linear momentum, but with torques

replacing forces. It was shown that for an isolated system obeying the law of action and reaction the total angular momentum is conserved, and that in the C-O-M system it is independent of the point of reference. All of these statements have their relativistic counterparts, at times involving some additional restrictions.

For a single particle, let us define an antisymmetric tensor of rank  $\binom{2}{0}$  in Minkowski space using the formalism of Eq. (7.64)

$$m = x \wedge p \quad (7.119)$$

whose elements would be

$$m^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu. \quad (7.120)$$

The  $3 \times 3$  subtensor  $m^{IJ}$  clearly corresponds, as was seen in Section 5.1, with the spatial angular momentum of the particle. An equation of motion for  $m^{\mu\nu}$  can be found by taking its derivative with respect to the particle's proper time and making use of Eq. (7.73) giving

$$\frac{dm}{d\tau} = u \wedge p + x \wedge K = x \wedge K, \quad (7.121)$$

where the first term vanishes by the antisymmetry of the wedge product and  $K$  is the Minkowski force. In component notation, Eq. (7.121) becomes

$$\frac{dm^{\mu\nu}}{d\tau} = x^\mu K^\nu - x^\nu K^\mu. \quad (7.122)$$

This suggests we define the relativistic generalization of the torque by

$$N = x \wedge K, \quad (7.123)$$

whose components are

$$N^{\mu\nu} = x^\mu K^\nu - x^\nu K^\mu. \quad (7.124)$$

Thus,  $m$  obeys the equations of motion

$$\frac{dm}{d\tau} = N, \quad (7.125)$$

whose component form is

$$\frac{dm^{\mu\nu}}{d\tau} = N^{\mu\nu}, \quad (7.126)$$

with Eq. (1.11) as the nonrelativistic limiting form.

For a system involving a collection of particles, a total angular momentum 4-tensor can be defined (analogously to the total linear momentum 4-vector) as

$$M = \sum_s m_s \quad (7.127)$$

or in component form

$$M^{\mu\nu} = \sum_s m_s^{\mu\nu}, \quad (7.128)$$

where the index  $s$  denotes the  $s$ th particle. It is more difficult to form an equation of motion for  $M$  because each particle has its own proper time. (For the same reason, we did not attempt it even for  $P$ .) Nevertheless, plausible arguments can be given for the conservation of  $M$  under certain circumstances. If the system is completely isolated and the particles do not interact with each other or with the outside world (including fields), then  $m$  for each particle is conserved by Eq. (7.126), and therefore  $M$  is also conserved. Even if the particles interact, but the interaction takes place only through binary collisions at a point, there still could be conservation as can be seen from the following argument. Instantaneously when the two particles collide they are traveling together and have the same proper time. In other words, their world lines cross and they share the same *event*. One can therefore write an equation of motion of the form of Eq. (7.126) for the sum of their angular momenta. If the impulsive forces of contact are equal and opposite—as we would expect from conservation of linear momentum in the collision—then the sum of the impulsive torques cancel. Hence relativistic angular momentum is also conserved through such collisions. Note that unlike the nonrelativistic case covariance requires that the interactions are assumed to be instantaneous point collisions.

The relativistic angular momentum obeys the same kind of theorem regarding translation of the reference point as does its nonrelativistic counterpart. In the definition, Eq. (7.120) or Eq. (7.128), the reference point (really reference “event”) is the arbitrary origin of the Lorentz system. With respect to some other reference event  $\mathbf{a}_\lambda$ , the total angular momentum is

$$M(\mathbf{a}_\lambda) = \sum_s (\mathbf{x}_s - \mathbf{a}_\lambda) \wedge \mathbf{p}_s, \quad (7.129)$$

$$= M(0) - \mathbf{a}_\lambda \wedge \mathbf{P} \quad (7.130)$$

As in the nonrelativistic case, the change in the angular momentum components is equal to the angular momentum, relative to the origin, that the whole system would have if it were located at  $\mathbf{a}_\lambda$ .

In Chapter 1, one particular reference point played an important role—the center of mass. We can find something similar here, at least in one Lorentz frame, by examining the nature of the mixed time and space components of  $M^{\mu\nu}$ , namely,  $M^{0j} = -M^{j0}$ . By definition, in some particular Lorentz frame, these components are given by

$$M^{0j} = \sum_s (x_s^0 p_s^j - x_s^j p_s^0) \quad (7.131)$$

$$= c \sum_s \left( t p_s^j - \frac{x_s^j E_s}{c^2} \right). \quad (7.132)$$

In the C-O-M frame, the total linear momentum  $p = \sum p_s$  vanishes, and  $M^{0j}$  in this frame has the form

$$M^{0j} = -c \sum_s \frac{x_s^j E_s}{c^2}. \quad (7.133)$$

If the system is such that the total angular momentum is conserved, as described above, then along with other components  $M$  is conserved and hence

$$\sum_s x_s^j E_s = \text{constant}.$$

Conservation of total linear momentum means that  $E = \sum E_s$  is also conserved. It is therefore possible to define a *spatial* point  $R_j$ ,

$$R_j = \frac{\sum_s x_s^j E_s}{\sum_s E_s}, \quad (7.134)$$

associated with the system, which is stationary in the C-O-M coordinate frame. In the nonrelativistic limit, where to first approximation  $E_s = m_s c^2$ , Eq. (7.134) reduces to the usual definition, Eq. (1.21). Thus, a meaningful center of mass (sometimes called *center of energy*) can be defined in special relativity only in terms of the angular-momentum tensor, and only for a particular frame of reference. Finally, it should be noted that by Eq. (7.130) the spatial part of the angular momentum tensor,  $M$ , is independent of reference point in the C-O-M system, exactly as in the nonrelativistic case.

Except for the special case of point collisions, we have so far carefully skirted the problem of finding the motion of a relativistic particle given the Minkowski forces. To this more general problem we address ourselves in the next section, within the nominal framework of the Lagrangian formulation.

## 7.9 ■ THE LAGRANGIAN FORMULATION OF RELATIVISTIC MECHANICS

Having established the appropriate generalization of Newton's equation of motion for special relativity, we can now seek to establish a Lagrangian formulation of the resulting relativistic mechanics. Generally speaking, there are two ways in which this has been attempted. One method makes no pretense at a manifestly covariant formulation and instead concentrates on reproducing, for some particular Lorentz frame, the spatial part of the equation of motion, Eq. (7.76). The forces  $F_i$  may or may not be suitably related to a covariant Minkowski force. The other method sets out to obtain a covariant Hamilton's principle and ensuing Lagrange's equations in which space and time are treated in common fashion as coordinates in a four-dimensional configuration space. The basis for the first method is at times quite shaky, especially when the forces are not relativistically well formulated. Most of

the time, however, the equations of motion so obtained, while not manifestly covariant, are relativistically correct for some particular Lorentz frame. The second method, on the other hand, seems clearly to be the proper approach, but it quickly runs into difficulties that require skillful handling if they are to be solvable, even for a single particle. For a system of more than one particle, it breaks down almost from the start. No satisfactory formulation for an interacting multiparticle system exists in classical relativistic mechanics except for some few special cases.

This section follows the first method, seeking to find a Lagrangian that leads to the relativistic equations of motion in terms of the coordinates of some particular inertial system. Within these limitations there is no great difficulty in constructing a suitable Lagrangian. It is true that the method of Section (1.4), deriving the Lagrangian from D'Alembert's principle, will not work here. While the principle itself remains valid in any given Lorentz frame, the derivation there is based on  $p_i = m_i v_i$ , which is no longer valid relativistically. But we may also approach the Lagrangian formulation from the alternative route of Hamilton's principle (Section 2.1) and attempt simply to find a function  $L$  for which the Euler-Lagrange equations, as obtained from the variational principle

$$\delta I = \delta \int_{t_1}^{t_2} L dt = 0. \quad (7.135)$$

agree with the known relativistic equations of motion, Eq. (7.76).

A suitable relativistic Lagrangian for a single particle acted on by conservative forces independent of velocity would be\*

$$L = -mc^2\sqrt{1 - \beta^2} - V, \quad (7.136)$$

where  $V$  is the potential, depending only upon position, and  $\beta^2 = v^2/c^2$ , with  $v$  the speed of the particle in the Lorentz frame under consideration. That this is the correct Lagrangian can be shown by demonstrating that the resultant Lagrange equations,

$$\frac{d}{dt} \left( \frac{\partial L}{\partial v^i} \right) - \frac{\partial L}{\partial x^i} = 0,$$

agree with Eq. (7.76). Since the potential is velocity independent  $v_i$  occurs only in the first term of (7.136) and therefore

$$\frac{\partial L}{\partial v^i} = \frac{mv^i}{\sqrt{1 - \beta^2}} = P^i. \quad (7.137)$$

The equations of motion derived from the Lagrangian (7.136) are then

\*We do not choose  $L = mc^2\sqrt{1 - \sqrt{1 - \beta^2}} - V$  because we want  $h$  in Eq. (7.139) to be the total energy

$$\frac{d}{dt} \frac{mv^i}{\sqrt{1 - \beta^2}} = -\frac{\partial V}{\partial x^i} = F^i,$$

which agree with (7.76). Note that the Lagrangian is no longer  $L = T - V$  but that the partial derivative of  $L$  with velocity is still the momentum. Indeed, it is this last fact that ensures the correctness of the Lagrange equations, and we could have worked backward from Eq. (7.137) to supply at least the velocity dependence of the Lagrangian.

We can readily extend the Lagrangian (7.136) to systems of many particles and change from Cartesian to any desired set of generalized coordinates  $q$ . The canonical momenta,  $\mathcal{P}$ , will still be defined by

$$\mathcal{P}_i = \frac{\partial L}{\partial \dot{q}^i}, \quad (7.138)$$

so that the connection between cyclic coordinates and conservation of the corresponding momenta remains just as in the nonrelativistic theory. Further, just as in Section (2.7), if  $L$  does not contain the time explicitly, there exists a constant of the motion

$$h = \dot{q}^i \mathcal{P}_i - L. \quad (7.139)$$

However, the identification of  $h$  with the energy for, say, a Lagrangian of the form of Eq. (7.136) cannot proceed along the same route as in Section (2.7). Note that  $L$  in Eq. (7.136) is not at all a homogeneous function of the velocity components. Nonetheless, direct evaluation of Eq. (7.139) from Eq. (7.136) shows that in this case  $h$  is indeed the total energy:

$$h = \frac{mv_i v^i}{\sqrt{1 - \beta^2}} + mc^2 \sqrt{1 - \beta^2} + V,$$

which, on collecting terms, reduces to

$$h = \frac{mc^2}{\sqrt{1 - \beta^2}} + V = T + V + mc^2 = E. \quad (7.140)$$

The quantity  $h$  is thus again seen to be the total energy  $E$ , which is therefore a constant of the motion under these conditions.

The introduction of velocity-dependent potentials produces no particular difficulty here and can be performed in exactly the same manner as in Section 1.5 for nonrelativistic mechanics. Thus, the Lagrangian for a single particle of charge,  $q$ , in an electromagnetic field is

$$L = -mc^2 \sqrt{1 - \beta^2} - q\phi + q\mathbf{A} \cdot \mathbf{v}. \quad (7.141)$$

Note that the *canonical* momentum is no longer  $mu$ ; there are now additional terms arising from the velocity dependent part of the potential:

$$\mathcal{P}^i = mu^i + qA^i. \quad (7.142)$$

This phenomenon is not a relativistic one of course; exactly the same additional term was found in the earlier treatment (cf. Eq. (2.47)). The formulation of Eq. (7.141) is not manifestly covariant. But we can confidently expect that the results will hold in all Lorentz frames as a consequence of the relativistic covariance of the Lorentz force derivable from the velocity dependent potential in Eq. (7.141).

Almost all of the procedures devised previously for the solution of specific mechanical problems thus can be carried over into relativistic mechanics. A few simple examples will be considered here by way of illustration.

1. *Motion under a constant force; hyperbolic motion.* It will be no loss of generality to take the  $x$  axis as the direction of the constant force. The Lagrangian is therefore

$$L = -mc^2\sqrt{1 - \beta^2} - max, \quad (7.143)$$

where  $\beta$  is  $\dot{x}/c$  and  $a$  is the constant magnitude of the force per unit mass. Either from Eq. (7.143) or directly on the basis of Eq. (7.76), the equation of motion is easily found to be

$$\frac{d}{dt}\left(\frac{\beta}{\sqrt{1 - \beta^2}}\right) = \frac{a}{c}.$$

The first integration leads to

$$\frac{\beta}{\sqrt{1 - \beta^2}} = \frac{at + \alpha}{c}$$

or

$$\beta = \frac{at + \alpha}{\sqrt{c^2 + (at + \alpha)^2}},$$

where  $\alpha$  is a constant of integration. A second integration over  $t$  from 0 to  $t$  and  $x$  from  $x_0$  to  $x$ ,

$$x - x_0 = c \int_0^t \frac{(at' + \alpha)dt'}{\sqrt{c^2 + (at' + \alpha)^2}},$$

leads to the complete solution

$$x - x_0 = \frac{c}{a} \left[ \sqrt{c^2 + (at + \alpha)^2} - \sqrt{c^2 + \alpha^2} \right]. \quad (7.144)$$

If the particle starts at rest from the origin so that  $x_0 = 0$  and  $v_0 = 0 = \alpha$ , then Eq. (7.144) can be written as

$$\left( x + \frac{c^2}{a} \right)^2 - c^2 t^2 = \frac{c^4}{a^2},$$

which is the equation of a hyperbola in the  $x, t$  plane. (Under the same conditions the nonrelativistic motion is of course a parabola in the  $x, t$  plane). The nonrelativistic limit is obtained from Eq. (7.144) by considering  $(at + \alpha)$  small compared to  $c$ ; the usual freshman-physics formula for  $x$  as a function of  $t$  is then easily obtained, recognizing that in this limit  $\alpha \rightarrow v_0$ .

The motion described in this example arises in reasonably realistic situations. It corresponds, for example, to the acceleration of electrons to relativistic speeds in the laboratory system by means of a constant and uniform electric field. The illustration considered next is more academic, but is of interest as an example of the techniques employed.

*2. The relativistic one-dimensional harmonic oscillator.* The Lagrangian in this case is of the form of Eq. (7.136) with

$$V(x) = \frac{1}{2} k x^2. \quad (7.145)$$

Since  $L$  is then not explicitly a function of time and  $V$  is not velocity dependent, the total energy  $E$  is constant. Equation (7.140) may now be solved for the velocity  $x$  as

$$\frac{1}{c^2} \left( \frac{dx}{dt} \right)^2 = 1 - \frac{m^2 c^4}{(E - V)^2}. \quad (7.146)$$

For the moment, we shall postpone substituting in the particular form of  $V(x)$  and generalize the problem slightly to include any potential sharing the qualitative characteristics of Eq. (7.145). Thus, let us suppose that  $V(x)$  is any potential function symmetric about the origin and possessing a minimum at that point. Then providing  $E$  lies between  $V(0)$  and the maximum of  $V$ , the motion will be oscillatory between limits  $x = -b$  and  $x = +b$ , determined by

$$V(\pm b) = E.$$

The period of the oscillatory motion is, by Eq. (7.146), to be obtained from

$$\tau = \frac{4}{c} \int_0^b \frac{dx}{\sqrt{1 - \frac{m^2 c^4}{(E - V(x))^2}}}. \quad (7.147)$$

Equation (7.147), when specialized to the particular Hooke's law form (7.145) for  $V(x)$ , can be expressed in terms of elliptic integrals. We shall instead examine the first-order relativistic corrections when the potential energy is always small compared to the rest mass energy  $mc^2$ . A change of notation is helpful. The energy  $E$  can be written as

$$E = mc^2(1 + \mathcal{E})$$

so that here

$$\frac{E - V(x)}{mc^2} = 1 + \mathcal{E} - \kappa x^2 = 1 + \kappa(b^2 - x^2), \quad (7.148)$$

where

$$\kappa = \frac{k}{2mc^2}. \quad (7.149)$$

To the order  $(\kappa b^2)^2$ , the period, Eq. (7.147) then reduces to

$$\tau \approx \frac{4}{c} \int_0^b \frac{dx}{\sqrt{2\kappa(b^2 - x^2)}} \left[ 1 - \frac{3\kappa}{4}(b^2 - x^2) \right]. \quad (7.150)$$

The integral in Eq. (7.150) can be evaluated by elementary means, most simply by changing variable through  $x = b \sin \phi$ ; the final result is

$$\tau \approx \frac{2\pi}{c} \frac{1}{\sqrt{2\kappa}} \left( 1 - \frac{3}{8}\kappa b^2 \right) = 2\pi \sqrt{\frac{m}{k}} \left( 1 - \frac{3}{16} \frac{kb^2}{mc^2} \right).$$

Note that the expression in front of the bracket is  $\tau_0$ , the nonrelativistic period of the harmonic oscillator. In special relativity, the period of the harmonic oscillator is thus not independent of the amplitude; instead, there is an amplitude dependent correction given approximately by

$$\frac{\Delta v}{v_0} = -\frac{\Delta \tau}{\tau_0} \approx \frac{3}{16} \frac{kb^2}{mc^2} = \frac{3}{8}\mathcal{E}. \quad (7.151)$$

*3. Motion of a charged particle in a constant magnetic field.* In principle, we should start from a Lagrangian of the form of Eq. (7.141) with the scalar potential  $\phi = 0$  and  $\mathbf{A}$  appropriate to a constant magnetic field (Eq. 5.106). But we know such a Lagrangian corresponds to the Lorentz force on the charged particle of charge  $q$ , given by

$$\mathbf{F} = q(\mathbf{v} \times \mathbf{B}) \quad (7.152)$$

(cf. Eq. 1.60). Hence, the equation of motion must be

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{v} \times \mathbf{B}) = \frac{q}{m\gamma}(\mathbf{p} \times \mathbf{B}). \quad (7.153)$$

The nature of the force, Eq. (7.152), is clearly such that the magnetic field does no work on the particle:  $\mathbf{F} \cdot \mathbf{v} = 0$ . Hence,  $E$  must be a constant, as also  $p$  and  $\gamma$  by Eq. (7.38'). Further, by Eq. (7.152), there is no component of the force parallel to  $\mathbf{B}$ , and the momentum component along that direction must remain constant. It is therefore no loss of generality to consider the motion only in the plane perpendicular to  $\mathbf{B}$  and to let  $\mathbf{p}$  represent the projection of the total linear momentum on to that plane. Equation (7.153) then says that the vector  $\mathbf{p}$  (whose magnitude is constant) is precessing around the direction of the magnetic field with a frequency

$$\Omega = \frac{qB}{m\gamma} \quad (7.154)$$

referred to as the cyclotron frequency. In the nonrelativistic limit  $\gamma \rightarrow 1$ . This agrees with the cyclotron resonance expression found in solid state physics texts. Because  $\gamma$  is constant, the velocity vector in the plane is also of constant magnitude and rotating with the same frequency. The particle must therefore move uniformly in a circular orbit in the plane with angular speed  $\Omega$ . Since the centrifugal force,  $F$ , equals  $mu^2/r$ , it follows that the magnitude of the linear momentum in the plane must be given by

$$p = m\gamma r\Omega.$$

Combining this expression with Eq. (7.154) leads to the relation between the circle radius and the momentum:

$$r = \frac{p}{qB}. \quad (7.155)$$

The radius of curvature into which the particle motion is bent depends only upon the particle properties through the ratio  $p/q$  ( $= Br$ ), which is sometimes called the *magnetic rigidity* of the particle. Note that while  $\Omega$  (Eq. (7.154)) shows relativistic corrections through the presence of  $\gamma$ , the relation between  $r$  and  $p$  is the same both relativistically and nonrelativistically. Recall that in both Eqs. (7.154) and (7.155)  $p$  is the magnitude of the momentum perpendicular to  $B$ , but in calculating  $\gamma$  we must use both the perpendicular and parallel components to find  $\beta$ .\*

## 7.10 ■ COVARIANT LAGRANGIAN FORMULATIONS

The Lagrangian procedure as given above certainly predicts the correct relativistic equations of motion. Yet it is a relativistic formulation only "in a certain sense."

\*The Larmor precession frequency  $\omega_L$  of Eq. (5.104) has an extra factor of 2, and corresponds to the precession of a magnetic moment in a constant magnetic field. This is a physically different case from that of the cyclotron resonance of a charged particle moving at a constant speed in a magnetic field.

No effort has been made to keep to the ideal of a covariant four-dimensional form for all the laws of mechanics. Thus, the time  $t$  has been treated as a parameter entirely distinct from the spatial coordinates, while a covariant formulation would require that space and time be considered as entirely similar coordinates in world space. Clearly some invariant parameter should be used, instead of  $t$ , to trace the progress of the system point in configuration space. Further, the examples of Lagrangian functions discussed in the previous section do not have any particular Lorentz transformation properties. Hamilton's principle must itself be manifestly covariant, which can only mean in this case that the action integral must be a world scalar. If the parameter of integration is a Lorentz invariant, then the Lagrangian function itself must be a world scalar in any covariant formulation. Finally, instead of being a function of  $x_i$  and  $\dot{x}_i$ , the Lagrangian should be a function of the coordinates in Minkowski space and of their derivatives with respect to the invariant parameter.

We shall consider primarily a system of only one particle. The natural choice of the invariant parameter in such a system would seem to be the particle's proper time  $\tau$ . But the various components of the generalized velocity,  $u^\nu$ , must then obey the relation

$$u \cdot u = u_\nu u^\nu = c^2, \quad (7.35)$$

which shows they are not independent. Therefore, we shall instead assume the choice of some Lorentz-invariant quantity  $\theta$  with no further specification than that it be a monotonic function of the progress of the world point along the particle's world line. For the purpose of this discussion, a superscript prime will be used to denote differentiation with respect to  $\theta$ :

$$x'^\nu \equiv \frac{dx^\nu}{d\theta},$$

while a dot over the letter indicates differentiation with respect to  $t$ . A suitably covariant Hamilton's principle must therefore appear as

$$\delta I = \delta \int_{\theta_1}^{\theta_2} \Lambda(x^\mu, x'^\mu) d\theta, \quad (7.156)$$

where the Lagrangian function  $\Lambda$  must be a world scalar and the  $(x^\mu, x'^\mu)$  means a function of all or any of these. Note that this formulation includes what would have ordinarily been called "time-dependent Lagrangians," because  $\Lambda$  is considered a function of  $x^0$ . The Euler-Lagrange equations corresponding to Eq. (7.156) are

$$\frac{d}{d\theta} \left( \frac{\partial \Lambda}{\partial x'^\mu} \right) - \frac{\partial \Lambda}{\partial x^\mu} = 0. \quad (7.157)$$

The problem is to find the form of  $\Lambda$  such that Eqs. (7.157) are equivalent to the equations of motion, Eq. (7.73).

One way of seeking  $\Lambda$  is to transform the action integral from the usual integral over  $t$  to one over  $\theta$ , and to treat the time  $t$  appearing explicitly in the Lagrangian

not as a parameter but as an additional generalized coordinate. Since  $\theta$  must be a monotonic function of  $t$  as measured in some Lorentz frame, we have

$$\frac{dx^i}{dt} = \frac{dx^i}{d\theta} \frac{d\theta}{dt} = c \frac{x'^i}{x'^0}. \quad (7.158)$$

Hence, the action integral is transformed as

$$I = \int_{t_1}^{t_2} L(x^j, t, \dot{x}^j) dt = \frac{1}{c} \int_{\theta_1}^{\theta_2} L\left(x^j, c \frac{x'^j}{x'^0}\right) x'^0 d\theta.$$

It would seem therefore that a recipe for a suitable  $\Lambda$  is given by the relation

$$\Lambda(x^\mu, x'^\mu) = \frac{x'^\phi}{c} L\left(x^\mu, c \frac{x'^j}{x'^0}\right). \quad (7.159)$$

The Lagrangian obtained this way is however a strange creature, unlike any Lagrangian we have so far met. Note that no matter what the functional form of  $L$ , the new Lagrangian  $\Lambda$  is a homogeneous function of the generalized velocities in the first degree:

$$\Lambda(x^\mu, ax'^\mu) = a\Lambda(x^\mu, x'^\mu). \quad (7.160)$$

This is not a phenomenon of relativistic physics per se; it is a mathematical consequence of enlarging configuration space to include  $t$  as a dynamical variable and using some other parameter to mark the system-point's travel through the space. A Lagrangian obeying Eq. (7.160) is often called (somewhat misleadingly) a homogeneous Lagrangian and the corresponding "homogeneous" problem of the calculus of variations requires special treatment. The most serious of the resulting difficulties will arise in the Hamiltonian formulation, but we can glimpse some of them by noting that in consequence the energy function  $h$ , according to Eq. (2.53), is identically zero. It follows from Euler's theorem on homogeneous functions that if  $\Lambda$  is homogeneous to first degree in  $x'^\mu$ , then

$$\Lambda = x'^\mu \frac{\partial \Lambda}{\partial x'^\mu}.$$

We can then show (cf. Derivation 10 at the end of this chapter) that as a result the function  $\Lambda$  *identically* satisfies the relation

$$\left[ \frac{d}{d\theta} \left( \frac{\partial \Lambda}{\partial x'^\mu} \right) - \frac{\partial \Lambda}{\partial x^\mu} \right] x'^\mu = 0. \quad (7.161)$$

Thus, if any three of the Lagrangian Eqs. (7.157) are satisfied, it will follow, solely as a consequence of the homogeneous property of  $\Lambda$ , that the fourth is satisfied identically.

Being thus forewarned to tread carefully, so to speak, let us carry out this transformation for a free particle. From Eq. (7.136), the "relativistic" but "noncovari-

ant" Lagrangian for the free particle is

$$L = -mc\sqrt{c^2 - \dot{x}^i \dot{x}_i}.$$

By the transformation of Eq. (7.159), a possible covariant Lagrangian is then

$$\Lambda = -mc\sqrt{x'_\mu x'^\mu}. \quad (7.162)$$

With this Lagrangian, the Euler–Lagrange equations are equivalent to

$$\frac{d}{d\theta} \left( \frac{mcx'}{\sqrt{x'^\mu x'_\mu}} \right) = 0.$$

The parameter  $\theta$  must be a monotonic function of the proper time  $\tau$  so that derivatives with respect to  $\theta$  are related to those in terms of  $\tau$  according to

$$x' \equiv \frac{dx}{d\theta} = \frac{d\tau}{d\theta} u.$$

Hence, the Lagrangian equations correspond to

$$\frac{d}{d\tau} \left( \frac{mcu}{\sqrt{u_v u^v}} \right) = \frac{d(mu)}{d\tau} = 0,$$

which are Eqs. (7.73) for a free particle. As we have seen above, the fourth of these equations says that the kinetic energy  $T$  is conserved, which is indeed not new but can be derived from the other three equations.

We have thus been led to a covariant Lagrangian procedure that works, at least for a single free particle, but only in a tortuous fashion. The elaborate superstructure can be greatly simplified however by a few bold pragmatic steps. First of all, we can avoid using  $\theta$  and work in terms of the proper time  $\tau$  directly by a procedure introduced in a slightly different context by Dirac. The constraint on the generalized velocities in terms of  $\tau$ , Eq. (7.35), is not a true dynamical constraint on the motion; rather it is a geometric consequence of the way in which  $\tau$  is defined. Equation (7.35) says in effect that we cannot roam over the full four-dimensional  $u$  space; we are confined to a particular three-dimensional surface in the space. Dirac calls relations such as Eq. (7.35) *weak equations*. We can with impunity treat  $u^\nu$  as unconstrained quantities, and only *after* all differentiation operations have been carried out, need the condition of Eq. (7.35) be imposed. Certainly the procedure would have worked above for the free particle Lagrangian. There would have been no difference if  $\theta$  were set equal to  $\tau$  from the start and Eq. (7.35) applied only in the last step. The covariant Lagrange equations can with this proviso therefore be written directly in terms of  $\tau$ :

$$\frac{d}{d\tau} \left( \frac{\partial \Lambda}{\partial u^\nu} \right) - \frac{\partial \Lambda}{\partial x^\nu} = 0. \quad (7.163)$$

Secondly, it is not a sacrosanct physical law that the action integral in Hamilton's principle must have the same value whether expressed in terms of  $t$  or in terms of  $\theta$  (or  $\tau$ ). It *needn't* be given by the prescription of Eq. (7.159). All that is required is that  $\Lambda$  be a world scalar (or function of a world scalar) that leads to the correct equations of motion. It doesn't *have* to be homogeneous to first degree in the generalized velocities. For example, a suitable  $\Lambda$  for a free particle would clearly be the quadratic expression

$$\Lambda = \frac{1}{2}mu_\nu u^\nu. \quad (7.164)$$

Many other possibilities are available.\* We shall use Eq. (7.162) for the "kinetic energy" part of the Lagrangian in all subsequent discussions; many present and future headaches will thereby be avoided.

If the particle is not free, but is acted on by external forces, then interaction terms have to be added to the Lagrangian of Eq. (7.164) that would lead to the corresponding Minkowski forces. Very little can be said at this time about the additional terms, other than they must be Lorentz-invariant. For example, if  $G^\mu$  were some (external) four-vector, then  $G_\mu x^\mu$  would be suitable interaction term. If in some particular Lorentz frame  $G_1 = ma$  and all other components vanish, then we would have an example of a constant force such as discussed in the previous section. In general, these terms will represent the interaction of the particle with some external field. The specific form will depend upon the covariant formulation of the field theory. We have only one example of a field already expressed in a covariant way—the electromagnetic field—and it is instructive therefore to examine the Lagrangian for a particle in an electromagnetic field.

A suitable Lagrangian can easily be seen to be

$$\Lambda(x^\mu, u^\mu) = \frac{1}{2}mu_\mu u^\mu + qu^\mu A_\mu(x^\lambda). \quad (7.165)$$

The corresponding Lagrange's equations are then

$$\frac{d}{d\tau}(mu^\nu) = -\frac{q}{d\tau}\frac{dA^\nu}{dx^\lambda} + \frac{\partial}{\partial x^\nu}(qu^\mu A_\mu),$$

which are exactly the generalized equations of motion Eq. (7.73), with the Minkowski force  $K_\nu$  on a charged particle, Eq. (7.74). Note that again the "mechanical momentum" four-vector  $p^\mu$  differs from the *canonical* momentum  $P^\mu$ :

\*In general,  $\Lambda$  can have the form  $mf(u_\nu u^\nu)$ , where  $f(y)$  is any function of  $y$  such that

$$\left.\frac{\partial f}{\partial y}\right|_{y=u^2} = \frac{1}{2}.$$

In Eq. (7.164), we have used  $f(u_\nu u^\nu) = \frac{1}{2}u_\nu u^\nu$ . The choice

$$f(u_\nu u^\nu) = -c\sqrt{u_\nu u^\nu}$$

corresponds to Eq. (7.162)

$$\mathcal{P}^\mu = \frac{\partial \Lambda}{\partial u^\mu} = mu^\mu + qA^\mu = p^\mu + qA^\mu \quad (7.166)$$

by a term linear in the electromagnetic potential. The canonical momentum,  $\mathcal{P}$ , conjugate to  $x^0$  is now

$$\mathcal{P}^0 = \frac{E}{c} + q\frac{\phi}{c} = \frac{1}{c}\bar{E},$$

where  $E$  is the mechanical energy and  $\bar{E}$  is the total energy of the particle,  $E + q\phi$ . Thus, the momentum conjugate to the *time* coordinate is proportional to the total *energy*. A similar conjugate connection between these two quantities will recur later in nonrelativistic theory. The connection between the magnitude of the spatial “mechanical” momentum and the energy  $E$  is still given by Eq. (7.38'). From Eq. (7.166), it is seen that the canonical momenta conjugate to  $x$  form the components of a spatial Cartesian vector  $\mathcal{P}$  related to  $\mathbf{p}$  by

$$\mathcal{P} = \mathbf{p} + q\mathbf{A}. \quad (7.167)$$

In terms of  $\mathcal{P}$ , Eq. (7.100) can be rewritten as

$$E^2 = (\mathcal{P} - q\mathbf{A})^2 + m^2c^4, \quad (7.168)$$

which is a useful relation between the energy  $E$  and the canonical momentum vector  $\mathcal{P}$ .

The interaction term in the Lagrangian of Eq. (7.165) is an example of a vector field interaction (as is also a term of the form  $G_\mu x^\mu$ ). We could also have a simple scalar field interaction where the term added to the Lagrangian would be some world scalar  $\psi(x^\mu)$ . Or more complicated invariant interaction terms can be created involving an external tensor field. The nature of such Lagrangians properly stems from the physical field theory involved and cannot concern us further here.

So far we have spoken only of systems comprising a single mass particle. Multiparticle systems introduce new complications. One obvious problem is finding an invariant parameter to describe the evolution of the system—each particle in the system has its own proper time. With a little thought, however, we could imagine ways of solving this difficulty. For example, the proper time associated with the C-O-M system involves a symmetric treatment of all the particles and might prove suitable. We could also include in the picture interactions of the particles with external fields very much as was done for a single particle. The great stumbling block however is the treatment of the type of interaction that is so natural and common in nonrelativistic mechanics—direct interaction between particles.

At first sight, it would seem indeed that such interactions are impossible in relativistic mechanics. To say that the force on a particle depends upon the positions or velocities of other particles at the same time implies propagation of effects with infinite velocity from one particle to another—“action at a distance.” In special relativity, where signals cannot travel faster than the speed of light,

action-at-a-distance seems outlawed. And in a certain sense this seems to be the correct picture. It has been proven that if we require certain properties of the system to behave in the normal way (such as conservation of total linear momentum), then there can be no covariant direct interaction between particles except through contact forces.

There have been many attempts in recent years to get around this “no-interaction” theorem. After all, we have seen that electromagnetic forces can be expressed covariantly, and a static electric field gives rise to the Coulomb law of attraction, which has the same form as the supposedly banned Newtonian gravitational attraction. Some of these attempts have led to approximately covariant Lagrangians, correct through orders of  $v^2/c^2$ . Others involve formulations of mechanics at variance with our normal structures; most for example cannot be stated in terms of a simple Hamilton’s principle.

### 7.11 ■ INTRODUCTION TO THE GENERAL THEORY OF RELATIVITY

Thus far we have been careful to use the term “special theory of relativity” and not to introduce the term “special relativity,” by which we endeavored to make clear that it is the theory that is special, not the relativity. The special theory uses ideal inertial frames that are assumed to exist over all of spacetime. The general theory not only removes that requirement, but also has a spacetime whose nature is part of the solution to the question of motion. To paraphrase John A. Wheeler: “*Matter tells space how to bend, and space returns the compliment by telling matter how to move.*” The general theory is often interpreted in terms of non-Euclidean geometry, so terms like geodesic (paths of shortest distance) and curvature of spacetime are often used. In this brief section we can only outline the formalism of the general theory to show how the full tensor notation is used.

Five principles guided Einstein in the development of the general theory:

1. *Mach’s principle*—the special theory used inertial frames. E. Mach observed that Newtonian inertial frames were not rotating with respect to the fixed stars. This suggests Mach’s principle, whereby inertial properties are determined by the presence of other bodies in the universe.
2. *Principle of equivalence*—whereby the gravitational mass for each body in the universe can be consistently and universally chosen to equal its inertial mass. To the best accuracy of all experiments performed to date, the ratio of the gravitational mass (the mass that appears in Newton’s force law for gravity) to the inertial mass (the mass that appears in the second law) of any object is independent of both the total mass and of the composition of the object. This means that no local experiments can distinguish nonrotating free fall in a gravitational field from uniform motion in the absence of any gravitational fields. Likewise, local experiments cannot distinguish between being at rest in a uniform gravitational field and undergoing uniform acceleration in the absence of any gravitational field (that is, in a rocket).

3. *Principle of covariance*—in the special theory, all inertial observers are equivalent. The general theory extends this idea by postulating the principle of covariance. This principle is that all observers, inertial or not, observe the same laws of physics. That means the laws of physics can be expressed in terms of tensors, since tensors are geometric objects defined independent of any coordinate system.
4. *Correspondence principle*—in weak gravitational fields with velocities small compared to light, the general theory should make predictions that approximate the predictions of gravitational behavior in Newtonian mechanics. As gravitational fields go to zero, the correspondence principle states the predictions of the general theory should approach those of the special theory.
5. *Principle of minimal gravitational coupling*—this principle postulates that no terms explicitly containing the curvature should be added in making the transition from the special theory to the general theory.

Newton's first law tells us that in the absence of external force bodies move along straight lines without acceleration. The preceding guiding principles suggest that in the general theory, objects will move along the geodesics of spacetime. For example, let us consider a family of geodesics that start out parallel. If gravitational effects in the region under consideration are uniform, the geodesics will remain parallel. If there is a nonuniform gravitational field, the geodesics should start to approach or recede. The change in separation, or *geodesic deviation*, is the proper measure of the gravitational field. Near Earth's surface, we often assume the gravitational field is uniform over small regions. Thus, we assume two falling bodies released side by side fall parallel. An experiment for larger separations or longer fall times measures the nonuniformity of Earth's gravitational field.

To illustrate this, let us consider an example of two balls separated horizontally by a distance,  $d$ , which are dropped at the same time from the same height high above Earth. Very close to either ball, and neglecting the gravitational mass of the balls, local experiments will give results that allow us to treat the local region as an inertial frame. Locally, gravity can be made to vanish by a choice of coordinate frame. Let us choose this local free-fall frame for our observations. Locally this satisfies the conditions for an inertial frame. However, as the balls fall toward Earth, their separation,  $d$ , decreases. This change in separation, rather than the fall toward Earth, is the local measure of the gravitational effect of Earth since it can not be eliminated by a choice of frame. This is reflected by the general theory statement that only the tides (differential effects) are real gravitational effects. Any other gravitational effects can be locally eliminated by freely falling.

Now consider two geodesics as shown in Figure 7.5. We can define two vector fields at any point. One field, denoted by  $u$ , gives the 4-velocity of motion along the geodesic, while the other field, denoted by  $\xi$ , gives the separation to the next geodesic. We assume at some time,  $\tau$ , there were test particles at the head and tail of the  $\xi$  vector.

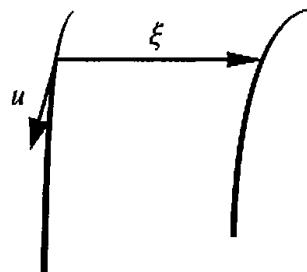


FIGURE 7.5 Tangent vector,  $u$ , and deviation vector,  $\xi$ .

We shall use the proper time at the tail of the deviation vector and have the head point to where the other test particle is at that time. In general, as the motion progresses, the proper time of the first test particle will not be the same proper time for the other test particle. A straightforward calculation, in the Newtonian limit, for the example of two falling balls, gives for the space components of  $\xi$  perpendicular to the direction toward Earth's center,

$$\frac{d^2 \xi^i}{dt^2} = R \xi^i, \quad (7.169)$$

where  $R$  depends upon the distance to Earth's center and other physical constants. Equation (7.169) says the acceleration in the separation of two geodesics is proportional to their separation. A two-dimensional example is the geodesics on the surface of a sphere. Consider two initially parallel geodesics on a sphere. These geodesics will meet after they have traveled one-quarter of the circumference of the sphere. For this case, Eq. (7.169) has  $R = 1/a^2$ , where  $a$  is the radius of the sphere.

If we analyze this problem in three or more dimensions, the relative acceleration is written as  $D^2 \xi^i / ds^2$  where  $ds$  is the length of the travel along the geodesic and we use a  $D$  for the derivative since our coordinate system is completely arbitrary. The twists and turns in the coordinate system can cause changes in the components of  $\xi$  even if its magnitude is not changing. As he developed more of the theory, Einstein discovered that the mathematicians—in particular, Riemann—had already developed the mathematical tools needed. The metric serves the role of potentials and derivatives of the metric give the geometric forces. Since the derivatives of the metric are not tensors, a combination of the derivatives and the metric must be used. There are also problems introduced by the freedom of using any coordinate system. Some of the changes are due to physical forces and others are due to the choice of the coordinate system in analogy to the Coriolis effect in a rotating coordinate system. The correct expression for the deviation of geodesic motion is provided by a tensor named *Riemann*. It is constructed of linear combinations of second derivatives of the metric contracted with the metric. *Riemann* has slots for three vectors and one slot for a single one-form. If we put the tangent vector into the second and fourth slots and the deviation vector into the third slot,

*Riemann* produces

$$\nabla_u \nabla_u \xi + Riemann(\dots, u, \xi, u) = 0, \quad (7.170)$$

where  $\nabla_u \nabla_u = \frac{D^2}{ds^2}$ . In component notation, Eq. (7.170) is

$$\frac{d^2 \xi^\alpha}{ds^2} + R^\alpha_{\beta\gamma\delta} \frac{dx^\beta}{d\tau} \xi^\gamma \frac{dx^\delta}{d\tau} = 0. \quad (7.171)$$

If we contract *Riemann* on slots 1 and 3, we produce a tensor called *Ricci*, defined as

$$Ricci(u, v) = Riemann(w^\alpha, u, e_\alpha, v), \quad (7.172)$$

whose components are

$$R_{\mu\nu} = R^\alpha_{\mu\alpha\nu}. \quad (7.173)$$

Another critical contraction produces the *curvature scalar*, called *R*

$$R = Ricci(w^\alpha, e_\alpha) = R^\alpha_{\alpha}. \quad (7.174)$$

Of all these possible contractions of *Riemann*, only one tensor of rank  $\binom{2}{0}$  retains all the differential symmetries of *Riemann*. That tensor is called *Einstein* (denoted by *G*) and is defined as

$$G = Ricci - \frac{1}{2}gR, \quad (7.175)$$

with components

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R. \quad (7.176)$$

Using *T* to denote the stress-energy tensor, Einstein's field equations make *Einstein* proportional to *T*.

$$G = kT. \quad (7.177)$$

These equations for weak gravitational fields and for speeds much less than light approach Newtonian gravitational theory, and for no gravitational fields produce the results of the special theory. They also correctly predict all the measured first- and second-order corrections to the special theory of relativity in experiments thus far performed. In addition, the theory predicts the existence of gravitational waves from moving masses. Although these waves have not, at this writing, been directly observed, measured changes in the periods of several binary star systems are consistent with the existence of such radiation existing.

Soon after Einstein proposed Eqs. (7.177), astronomers pointed out that the solutions of these equations were not consistent with their observation of a static universe that was neither expanding nor contracting. Einstein modified the equations by adding a term that was proportional to the metric tensor. The constant of

proportionality, called the *cosmological constant*, was denoted by  $\Lambda$  giving

$$G + \Lambda g = kT. \quad (7.178)$$

Soon after that, astronomers decided that the observational data showed that the universe was expanding and the cosmological constant was not needed, and most physicists dropped the term. Einstein said that the cosmological constant was his greatest mistake. However, the early 21st century observational data on distant galaxies suggests that the universe is accelerating as it expands. This would reintroduce the cosmological constant into the field equations. The current terminology, since this would be a  $\Lambda < 0$ , is to refer to the cosmological constant as “dark energy,” since it is a positive contribution to the right-hand side of Eq. (7.178).

## DERIVATIONS

1. Consider a mechanical system of  $n$  particles, with a conservative potential consisting of terms dependent only upon the scalar distance between pairs of particles. Show explicitly that the Lagrangian for the system when expressed in coordinates derived by a Galilean transformation differs in form from the original Lagrangian only by a term that is a total time derivative of a function of the position vectors. This is a special case of invariance under a point transformation (cf. Derivation 10, Chapter 1).
2. Obtain the Lorentz transformation in which the velocity is at an infinitesimal angle  $d\theta$  counterclockwise from the  $x$  axis, by means of a similarity transformation applied to Eq. (7.16). Show directly that the resulting matrix is orthogonal and that the inverse matrix is obtained by substituting  $-v$  for  $v$ .
3. The Einstein addition law can also be obtained by remembering that the second velocity is related directly to the space components of a four-velocity, which may then be transformed back to the initial system by a Lorentz transformation. If the second system is moving with a speed  $v'$  relative to the first in the direction of their  $z$  axes, while a third system is moving relative to the second with an arbitrarily oriented velocity  $v''$ , show by this procedure that the magnitude of the velocity  $v$  between the first and third system is given by

$$\sqrt{1 - \beta^2} = \frac{\sqrt{1 - \beta'^2} \sqrt{1 - \beta''^2}}{1 + \beta' \beta''_z},$$

and that the components of  $v$  are

$$\beta_x = \frac{\beta''_x \sqrt{1 - \beta'^2}}{1 + \beta' \beta''_z}, \quad \beta_y = \frac{\beta''_y \sqrt{1 - \beta'^2}}{1 + \beta' \beta''_z}, \quad \beta_z = \frac{\beta' + \beta''_z}{1 + \beta' \beta''_z}$$

Here  $\beta''_x = v''_x/c$ , and so forth.

4. Show that the magnitude of the velocity of the preceding exercise between the first and the third systems can be given in general by

$$\beta^2 = \frac{(\boldsymbol{\beta}' + \boldsymbol{\beta}'')^2 - (\boldsymbol{\beta}' \times \boldsymbol{\beta}'')^2}{(1 + \boldsymbol{\beta}' \cdot \boldsymbol{\beta}'')^2}.$$

5. Show that the matrix  $\mathbf{R}$  defined by Eq. (7.21) has the form of a spatial rotation by doing the matrix multiplication, and by examining the properties of the  $3 \times 3$  submatrix with elements  $R_{ij}$ . Prove that there cannot be two rotation matrices such that Eq. (7.21) is satisfied; that is,  $\mathbf{R}$  is unique. Finally, show that  $\mathbf{L}$  can similarly be uniquely factored into a rotation and a pure Lorentz transformation in the form

$$\mathbf{L} = \mathbf{P}'\mathbf{R}'.$$

6. Show that to each plane wave there is associated a covariant four-vector involving the frequency and the wave number. From the consequent transformation equations of the components of the four-vector, derive the Doppler-effect equations.
7. From the transformation properties of the world acceleration, show that the components of the acceleration  $\mathbf{a}$  are given in terms of the transformed acceleration  $\mathbf{a}'$  in a system momentarily at rest with respect to the particle by the formulas

$$a'_x = \frac{a_x}{(1 - \beta^2)^{3/2}}, \quad a'_y = \frac{a_y}{1 - \beta^2}, \quad a'_z = \frac{a_z}{1 - \beta^2},$$

the  $x$  axis being chosen in the direction of the relative velocity.

8. By expanding the equation of motion, Eq. (7.73), with Eq. (7.36) for the momentum show that the force is parallel to the acceleration only when the velocity is either parallel or perpendicular to the acceleration. Obtain expressions for the coefficients of the acceleration in these two cases. In the older literature, these coefficients were known as the longitudinal and transverse masses, respectively.
9. A generalized potential suitable for use in a covariant Lagrangian for a single particle

$$\mathcal{U} = -A_{\lambda\nu}(x^\mu)u^\lambda u^\nu$$

where  $A_{\lambda\nu}$  stands for a symmetric world tensor of the second rank and  $u^\nu$  are the components of the world velocity. If the Lagrangian is made up of Eq. (7.164) minus  $\mathcal{U}$ , obtain the Lagrange equations of motion. What is the Minkowski force? Give the components of the force as observed in some Lorentz frame.

10. Show that if  $\Lambda$  satisfies the Lagrange equations, it identically satisfies Eq. (7.161) on the basis of the homogeneity of  $\Lambda$ , by explicitly forming the total derivative with respect to  $\theta$  that occurs in the equation.
11. In special relativity, it is not necessarily obvious that the velocity of system B as observed in system A is the negative of the velocity vector of system A observed in system B. From the orthogonality properties of  $\mathbf{L}$ , prove that the two vectors have the same magnitude and are in fact the negative of each other. For simplicity, a pure Lorentz transformation may be assumed, although this condition is not necessary for the proof.
12. A set of transformations are said to have the group property if they possess the following four characteristics:

- The transformation equivalent to two successive transformations (“product” of transformations) is a member of the set.
- The product operation obeys the associative law.
- The identity transformation is a member of the set.
- The inverse of each transformation in the set is also a member of the set.

Prove that the sets of full Lorentz transformations and of restricted Lorentz transformation have (separately) the group property.

## EXERCISES

13. Show by direct multiplication of the vector form of the Lorentz transformation, Eqs. (7.9), that

$$r'^2 - c^2 t'^2 = r^2 - c^2 t^2.$$

14. A rocket of length  $l_0$  in its rest system is moving with constant speed along the  $z$  axis of an inertial system. An observer at the origin of this system observes the apparent length of the rocket at any time by noting the  $z$  coordinates that can be seen for the head and tail of the rocket. How does this apparent length vary as the rocket moves from the extreme left of the observer to the extreme right? How do these results compare with measurements in the rest frame of the observer? (Note: observe, not measure).
15. A beam of particles moving with uniform velocity collides with a collection of target particles that are at rest in a particular system. Let  $\sigma_0$  be the collision cross section observed in this system. In another system, the incident particles have a normalized velocity  $\beta_1$  and the target particles a normalized velocity  $\beta_2$ . If  $\sigma$  is the observed cross section in this system, show that

$$\sigma = \sigma_0 \sqrt{1 - \frac{(\beta_1 \times \beta_2)^2}{(\beta_1 - \beta_2)^2}}.$$

Remember that collision rate must be invariant under a Lorentz transformation.

16. For a “close” satellite of Earth (semimajor axis approximately the radius of Earth) calculate numerically the value of the Thomas precession rate. Compare the result with the precession rate induced in the orbit because of the oblate figure of Earth. Assume the satellite orbital plane is inclined at  $30^\circ$  to the equator.
17. Two particles with rest masses  $m_1$  and  $m_2$  are observed to move along the observer’s  $z$  axis toward each other with speeds  $v_1$  and  $v_2$ , respectively. Upon collision, they are observed to coalesce into one particle of rest mass  $m_3$  moving with speed  $v_3$  relative to the observer. Find  $m_3$  and  $v_3$  in terms of  $m_1$ ,  $m_2$ ,  $v_1$ , and  $v_2$ . Would it be possible for the resultant particle to be a photon, that is,  $m_3 = 0$ , if neither  $m_1$  nor  $m_2$  are zero?
18. In the  $\beta$  disintegration considered in Exercise 17, Chapter 1, the electron has a mass equivalent to a rest energy of 0.511 MeV, while the neutrino has essentially no mass

What are the total energies carried away by the electron and neutrino? What fraction of the nuclear mass is converted into kinetic energy (including the electron rest energy)?

19. A meson of mass  $m_\pi$  at rest disintegrates into a meson of mass  $m_\mu$  and a neutrino of effectively zero mass. Show that the kinetic energy of motion of the  $\mu$  meson is

$$T = \frac{(m_\pi - m_\mu)^2}{2m_\pi} c^2.$$

20. A  $\pi^+$  meson of rest mass 139.6 MeV collides with a neutron (rest mass 939.6 MeV) stationary in the laboratory system to produce a  $K^+$  meson (rest mass 494 MeV) and a  $\Lambda$  hyperon (rest mass 1116 MeV). What is the threshold energy for this reaction in the laboratory system?

21. A photon may be described classically as a particle of zero mass possessing nevertheless a momentum  $h/\lambda = h\nu/c$ , and therefore a kinetic energy  $h\nu$ . If the photon collides with an electron of mass  $m$  at rest, it will be scattered at some angle  $\theta$  with a new energy  $h\nu'$ . Show that the change in energy is related to the scattering angle by the formula

$$\lambda' - \lambda = 2\lambda_e \sin^2 \frac{\theta}{2},$$

where  $\lambda_e = h/mc$ , is known as the Compton wavelength. Show also that the kinetic energy of the recoil motion of the electron is

$$T = h\nu \frac{2 \left( \frac{\lambda_e}{\lambda} \right) \sin^2 \frac{\theta}{2}}{1 + 2 \left( \frac{\lambda_e}{\lambda} \right) \sin^2 \frac{\theta}{2}}.$$

22. A photon of energy  $\mathcal{E}$  collides at angle  $\theta$  with another photon of energy  $E$ . Prove that the minimum value of  $\mathcal{E}$  permitting formation of a pair of particles of mass  $m$  is

$$\mathcal{E}_{th} = \frac{2m^2c^4}{E(1 - \cos \theta)}.$$

23. The theory of rocket motion developed in Exercise 13, Chapter 1, no longer applies in the relativistic region, in part because there is no longer conservation of mass. Instead, all the conservation laws are combined into the conservation of the world momentum; the change in each component of the rocket's world momentum in an infinitesimal time  $dt$  must be matched by the value of the same component of  $p_v$  for the gases ejected by the rocket in that time interval. Show that if there are no external forces acting on the rocket, the differential equation for its velocity as a function of the mass is

$$m \frac{dv}{dm} + a \left( 1 - \frac{v^2}{c^2} \right) = 0,$$

where  $a$  is the constant velocity of the exhaust gases *relative to the rocket*. Verify that the solution can be put in the form

$$\beta = \frac{1 - \left(\frac{m}{m_0}\right)^{2a/c}}{1 + \left(\frac{m}{m_0}\right)^{2a/c}},$$

$m_0$  being the initial mass of the rocket. Since mass is not conserved, what happens to the mass that is lost?

24. A particle in hyperbolic motion starts from the origin at  $t = 0$ . Find the time  $t_0$  such that if a photon is emitted from the origin after  $t_0$ , it will never catch up with the particle.
25. A particle of rest mass  $m$ , charge  $q$ , and initial velocity  $v_0$  enters a uniform electric field  $\mathbf{E}$  perpendicular to  $v_0$ . Find the subsequent trajectory of the particle and show that it reduces to a parabola as the limit  $c$  becomes infinite.
26. Show that the relativistic motion of a particle in an attractive inverse-square law of force is a precessing ellipse. Compute the precession of the perihelion of Mercury resulting from this effect. (The answer, about  $7''$  per century, is much smaller than the actual precession of  $43''$  per century that can be accounted for correctly only by general relativity. The other planets produce a precession greater than  $5.000''$  per century.)
27. Starting from the equation of motion (7.73), derive the relativistic analog of the virial theorem, which states that for motions bounded in space and such that the velocities involved do not approach indefinitely close to  $c$ , then

$$\overline{L_0} + \overline{T} = -\overline{\mathbf{F} \cdot \mathbf{r}},$$

where  $L_0$  is the form the Lagrangian takes in the absence of external forces. Note that although neither  $L_0$  nor  $T$  corresponds exactly to the kinetic energy in nonrelativistic mechanics, their sum,  $L + T$ , plays the same role as twice the kinetic energy in the nonrelativistic virial theorem, Eq. (3.26).

28. Let  $\mathbf{e}_1$  and  $\mathbf{e}_2$  be the basis vectors for a Cartesian coordinate system in a two-dimensional Euclidean space that contains a crystal whose lattice vectors are  $\mathbf{a} = \mathbf{e}_1$  and  $\mathbf{b} = \mathbf{e}_1 + \mathbf{e}_2$ . Use the underlying Euclidean geometry to determine that the reciprocal lattice vectors are  $\mathbf{A} = \mathbf{e}_1 - \mathbf{e}_2$  and  $\mathbf{B} = \mathbf{e}_2$ . Using the  $\mathbf{a}, \mathbf{b}$  pair as basis vectors, determine the metric tensor  $g$  necessary for  $\mathbf{A}$  and  $\mathbf{B}$  to be the 1-forms as defined by Eqs. (7.34') and (7.49).
29. Using *Maple* or *Mathematica* calculate the Lorentz transformation matrix in Eq. (7.17), then without assuming that the velocities in the frame  $S'$  are small, find the exact Lorentz boost from  $S$  to  $S''$ , (generalization of Eq. (7.20)) and the rotation (generalization of Eq. (7.21)). Show that your results reduce to Eqs. (7.20) and (7.21).
30. Using *Maple* or *Mathematica* or a similar program calculate the Einstein field equations for spherical coordinates assuming  $T_{\mu\nu} = 0$  everywhere except possibly for  $r = 0$ , where the coordinate system is undefined. The most general spherical static metric corresponds to an interval given by

$$ds^2 = e^{\nu(r)} c^2 dt^2 - e^{\lambda(r)} dr^2 - r^2(d\theta^2 + \sin^2 \theta d\phi^2),$$

where  $r$ ,  $\theta$ , and  $\phi$  correspond to the usual three-dimensional spherical coordinates. Solve these equations using an integration constant  $m$  to obtain the Schwarzschild so-

lution for a point source of mass  $m$ . As you will discover, these coordinates have a singularity at  $r = 2m$ . Show that this is a coordinate singularity (a singularity determined by the choice of coordinates) rather than a physical singularity by examining the components of *Riemann* as  $r$  crosses  $2m$ .

31. To show that the word “relativity” in the special theory of relativity does not have its ordinary meaning, consider a disk rotating in an inertial frame about an axis fixed at its center and perpendicular to the disk. Mounted on the edge of the disk are mirrors arranged so that light emitted tangentially from a point on the disk is reflected tangentially around the disk back to the starting location. Compare the behavior of light emitted in the direction of rotation (assumed clockwise) to the behavior of light emitted in the opposite direction. Now consider a pulse of light emitted by a source on the axis and used to synchronize the clocks on the perimeter. Since clocks are commonly synchronized by light and distance in the special theory (elapsed time = distance/ $c$ ), what does this say about the absolute sense of rotation in the special theory?
32. Show that the space components of Eq. (7.68) are identical to the components in the equation on the preceding line.

# CHAPTER

# 8

# The Hamilton Equations of Motion

The Lagrangian formulation of mechanics was developed largely in the first two chapters, and most of the subsequent discussion has been in the nature of application, but still within the framework of the Lagrangian procedure. In this chapter we resume the formal development of mechanics, turning our attention to an alternative statement of the structure of the theory known as the Hamiltonian formulation. Nothing new is added to the physics involved; we simply gain another (and more powerful) method of working with the physical principles already established. The Hamiltonian methods are not particularly superior to Lagrangian techniques for the direct solution of mechanical problems. Rather, the usefulness of the Hamiltonian viewpoint lies in providing a framework for theoretical extensions in many areas of physics. Within classical mechanics it forms the basis for further developments, such as Hamilton–Jacobi theory, perturbation approaches and chaos. Outside classical mechanics, the Hamiltonian formulation provides much of the language with which present-day statistical mechanics and quantum mechanics is constructed. We shall assume in the following chapters that the mechanical systems are holonomic and that the forces are monogenic, that is, derived either from a potential dependent upon position only, or from velocity-dependent generalized potentials of the type discussed in Section 1.5.

## 8.1 ■ LEGENDRE TRANSFORMATIONS AND THE HAMILTON EQUATIONS OF MOTION

In the Lagrangian formulation (nonrelativistic), a system with  $n$  degrees of freedom possesses  $n$  equations of motion of the form

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0. \quad (8.1)$$

As the equations are of second order, the motion of the system is determined for all time only when  $2n$  initial values are specified, for example, the  $n$   $q_i$ 's and  $n$   $\dot{q}_i$ 's at a particular time  $t_1$ , or then  $n$   $q_i$ 's at two times,  $t_1$  and  $t_2$ . We represent the state of the system by a point in an  $n$ -dimensional *configuration space* whose coordinates are the  $n$  generalized coordinates  $q_i$  and follow the motion of the system point in time as it traverses its path in configuration space. Physically, in the Lagrangian viewpoint a system with  $n$  independent degrees of freedom is a

problem in  $n$  independent variables  $q_i(t)$ , and  $\dot{q}_i$  appears only as a shorthand for the time derivative of  $q_i$ . All  $n$  coordinates must be independent. In the Hamiltonian formulation there can be no constraint equations among the coordinates. If the  $n$  coordinates are not independent, a reduced set of  $m$  coordinates, with  $m < n$ , must be used for the formulation of the problem before proceeding with the following steps.

The Hamiltonian formulation is based on a fundamentally different picture. We seek to describe the motion in terms of *first-order* equations of motion. Since the number of initial conditions determining the motion must of course still be  $2n$ , there must be  $2n$  independent first-order equations expressed in terms of  $2n$  *independent variables*. Hence, the  $2n$  equations of the motion describe the behavior of the system point in a *phase space* whose coordinates are the  $2n$  independent variables. In thus doubling our set of independent quantities, it is natural (though not inevitable) to choose half of them to be the  $n$  generalized coordinates  $q_i$ . As we shall see, the formulation is nearly symmetric if we choose the other half of the set to be the generalized or *conjugate momenta*  $p_i$  already introduced by the definition (cf. Eq. (2.44)):

$$p_i = \frac{\partial L(q_j, \dot{q}_j, t)}{\partial \dot{q}_i} \quad (\text{no sum on } j) \quad (8.2)$$

where the  $j$  index shows the set of  $q$ 's and  $\dot{q}$ 's. The quantities  $(q, p)$  are known as the *canonical variables*.\*

From the mathematical viewpoint, it can however be claimed that the  $q$ 's and  $\dot{q}$ 's have been treated as distinct variables. In Lagrange's equations, Eq. (8.1), the partial derivative of  $L$  with respect to  $q_i$  means a derivative taken with all other  $q$ 's and all  $\dot{q}$ 's constant. Similarly, in the partial derivatives with respect to  $\dot{q}$ , the  $q$ 's are kept constant. Treated strictly as a mathematical problem, the transition from Lagrangian to Hamiltonian formulation corresponds to changing the variables in our mechanical functions from  $(q, \dot{q}, t)$  to  $(q, p, t)$ , where  $p$  is related to  $q$  and  $\dot{q}$  by Eqs. (8.2). The procedure for switching variables in this manner is provided by the *Legendre transformation*, which is tailored for just this type of change of variable.

Consider a function of only two variables  $f(x, y)$ , so that a differential of  $f$  has the form

$$df = u dx + v dy, \quad (8.3)$$

where

$$u = \frac{\partial f}{\partial x}, \quad v = \frac{\partial f}{\partial y}. \quad (8.4)$$

\*Unless otherwise specified, in this and subsequent chapters the symbol  $p$  will be used only for the conjugate or canonical momentum. When the forces are velocity dependent, the canonical momentum will differ from the corresponding mechanical momentum (cf. Eq. (2.47)).

## Chapter 8 The Hamilton Equations of Motion

We wish now to change the basis of description from  $x, y$  to a new distinct set of variables  $u, y$ , so that differential quantities are expressed in terms of the differentials  $du$  and  $dy$ . Let  $g$  be a function of  $u$  and  $y$  defined by the equation

$$g = f - ux. \quad (8.5)$$

A differential of  $g$  is then given as

$$dg = df - u dx - x du,$$

or, by (8.3), as

$$dg = v dy - x du,$$

which is exactly in the form desired. The quantities  $x$  and  $v$  are now functions of the variables  $u$  and  $y$  given by the relations

$$x = -\frac{\partial g}{\partial u}, \quad v = \frac{\partial g}{\partial y}, \quad (8.6)$$

which are the analogues of Eqs. (8.4).

The Legendre transformation so defined is used frequently in thermodynamics. The first law of thermodynamics relates the differential change in energy,  $dU$ , to the corresponding change in heat content,  $dQ$ , and the work done,  $dW$ :

$$dU = dQ - dW. \quad (8.7)$$

For a gas undergoing a reversible process, Eq. (8.7) can be written as

$$dU = T dS - P dV, \quad (8.8)$$

where  $U(S, V)$  is written as a function of the entropy,  $S$ , and the volume,  $V$ , where the temperature,  $T$ , and the gas pressure,  $P$ , are given by

$$T = \frac{\partial U}{\partial S} \quad \text{and} \quad P = -\frac{\partial U}{\partial V}. \quad (8.9)$$

The enthalpy,  $H(S, P)$  is generated by the Legendre transformation

$$H = U + PV, \quad (8.10)$$

which gives

$$dH = T dS + V dP. \quad (8.11)$$

where

$$T = \frac{\partial H}{\partial S} \quad \text{and} \quad V = \frac{\partial H}{\partial P}.$$

Additional Legendre transformations,

$$\begin{aligned} F &= U - TS \\ G &= H - TS, \end{aligned} \tag{8.12}$$

generate the Helmholtz free energy,  $F(T, V)$ , and the Gibbs free energy,  $G(T, P)$ .

The transformation from  $(q, \dot{q}, t)$  to  $(q, p, t)$  differs from the type considered in Eqs. (8.3) to (8.12) only in that more than one variable is to be transformed. We begin by writing the differential of the Lagrangian,  $L(q, \dot{q}, t)$ , as

$$dL = \frac{\partial L}{\partial q_i} dq_i + \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial L}{\partial t} dt. \tag{8.13}$$

The canonical momentum was defined in Eq. (2.44) as  $p_i = \partial L / \partial \dot{q}_i$ ; substituting this into the Lagrange equation (8.1), we obtain

$$\dot{p}_i = \frac{\partial L}{\partial \dot{q}_i}, \tag{8.14}$$

so Eq. (8.13) can be written as

$$dL = \dot{p}_i dq_i + p_i d\dot{q}_i + \frac{\partial L}{\partial t} dt. \tag{8.13'}$$

The Hamiltonian  $H(q, p, t)$  is generated by the Legendre transformation

$$H(q, p, t) = \dot{q}_i p_i - L(q, \dot{q}, t), \tag{8.15}$$

which has the differential

$$dH = \dot{q}_i dp_i - \dot{p}_i dq_i - \frac{\partial L}{\partial t}, \tag{8.16}$$

where the term  $p_i d\dot{q}_i$  is removed by the Legendre transformation. Since  $dH$  can also be written as

$$dH = \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial t} dt, \tag{8.17}$$

we obtain the  $2n + 1$  relations

$$\left. \begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ -\dot{p}_i &= \frac{\partial H}{\partial q_i} \end{aligned} \right\} \tag{8.18}$$

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}. \tag{8.19}$$

Equations (8.18) are known as the *canonical equations of Hamilton*; they constitute the desired set of  $2n$  first-order equations of motion replacing the  $n$  second-order Lagrange equations.\*

The first half of Hamilton's equations give the  $\dot{q}_i$ 's as functions of  $(q, p, t)$ . They form therefore the inverse of the constitutive equations (8.2), which define the momenta  $p_i$  as functions of  $(q, \dot{q}, t)$ . It may therefore be said that they provide no new information. In terms of solving mechanical problems by means of the canonical equations, the statement is correct. But within the framework of the Hamiltonian picture, where  $H(q, p, t)$  is some given function obtained no matter how, the two halves of the set of Hamiltonian equations are equally independent and meaningful. The first half says how  $\dot{q}$  depends on  $q$ ,  $p$ , and  $t$ ; the second says the same thing for  $\dot{p}$ .

Of course, the Hamiltonian  $H$  is constructed in the same manner, and has identically the same value, as  $h$ , the energy function defined in Eq. (2.53). But they are functions of different variables: Like the Lagrangian,  $h$  is a function of  $q, \dot{q}$  (and possibly  $t$ ), while  $H$  must always be expressed as a function of  $q, p$  (and possibly  $t$ ). It is to emphasize this difference in functional behavior that different symbols have been given to the quantities even though they have the same numerical values.

Nominally, the Hamiltonian for each problem must be constructed via the Lagrangian formulation. The formal procedure calls for a lengthy sequence of steps:

1. With a chosen set of generalized coordinates,  $q_i$ , the Lagrangian  $L(q_i, \dot{q}_i, t) = T - V$  is constructed.
2. The conjugate momenta are defined as functions of  $q_i, \dot{q}_i$ , and  $t$  by Eqs. (8.2).
3. Equation (8.15) is used to form the Hamiltonian. At this stage we have some mixed function of  $q_i, \dot{q}_i, p_i$ , and  $t$ .
4. Equations (8.2) are then inverted to obtain  $\dot{q}_i$  as functions of  $(q, p, t)$ . Possible difficulties in the inversion will be discussed below.
5. The results of the previous step are then applied to eliminate  $\dot{q}$  from  $H$  so as to express it solely as a function of  $(q, p, t)$ .

Now we are ready to use the Hamiltonian in the canonical equations of motion.

For many physical systems it is possible to shorten this drawn-out sequence quite appreciably. As has been described in Section 2.7, in many problems the Lagrangian is the sum of functions each homogeneous in the generalized veloc-

\*Canonical is used here presumably in the sense of designating a simple, general set of standard equations. It appears that the term was first introduced by C. G. J. Jacobi in 1837 (*Comptes rendus de l'Académie des Sciences de Paris*, 5, p. 61) but in a slightly different context referring to an application of Hamilton's equations of motion to perturbation theory. Although the term rapidly gained common usage, the reason for its introduction apparently remained obscure even to contemporaries. By 1879, only 45 years after Hamilton explicitly introduced his equations, Thomson (Lord Kelvin) and Tait were moved by the adjective "canonical" to exclaim "Why it has been so called would be hard to say."

ties of degree 0, 1, and 2, respectively. In that case,  $H$  by the prescription of Eq. (8.15) is given by (cf. Eqs. (2.53) and (2.55))

$$H = \dot{q}_i p_i - L = \dot{q}_i p_i - [L_0(q_i, t) + L_1(q_i, t)\dot{q}_k + L_2(q_i, t)\dot{q}_k\dot{q}_m] \quad (8.20)$$

(no sum on  $i$  in the square brackets) where  $L_0$  is the part of the Lagrangian that is independent of the generalized velocities,  $L_1$  represents the coefficients of the part of the Lagrangian that is homogeneous in  $\dot{q}_i$  in the first degree, and  $L_2$  is the part that is homogeneous in  $\dot{q}_i$  in the second degree. Further, if the equations defining the generalized coordinates don't depend on time explicitly, then  $L_2\dot{q}_k\dot{q}_m = T$  (the kinetic energy), and if the forces are derivable from a conservative potential  $V$  (that is, work is independent of the path), then  $L_0 = -V$ . When both these conditions are satisfied, the Hamiltonian is *automatically* the total energy:

$$H = T + V = E. \quad (8.21)$$

If either Eq. (8.20) or (8.21) holds, then much of the algebra in steps 3 and 4 above is eliminated.

We can at times go further. In large classes of problems, it happens that  $L_2$  is a quadratic function of the generalized velocities and  $L_1$  is a linear function of the same variables with the following specific functional dependencies:

$$L(q_i, \dot{q}_i, t) = L_0(q, t) + \dot{q}_i a_i(q, t) + \dot{q}_i^2 T_i(q, t), \quad (8.22)$$

where the  $a_i$ 's and the  $T_i$ 's are functions of the  $q$ 's and  $t$ .

The algebraic manipulations required in steps 2–5 can then be carried out, at least formally, once and for all. To show this, let us form the  $\dot{q}_i$ 's into a single column matrix  $\dot{\mathbf{q}}$ . Under the given assumptions the Lagrangian can be written as

$$L(q, \dot{q}, t) = L_0(q, t) + \tilde{\mathbf{q}}\mathbf{a} + \frac{1}{2}\tilde{\mathbf{q}}\mathbf{T}\dot{\mathbf{q}}, \quad (8.23)$$

where the single row matrix  $\tilde{\mathbf{q}}$  has been written explicitly as the transpose of a single column matrix,  $\dot{\mathbf{q}}$ . Here  $\mathbf{a}$  is a column matrix, and  $\mathbf{T}$  is a square  $n \times n$  matrix (much like the corresponding matrix introduced in Section 6.2). The elements of both are in general functions of  $q$  and  $t$ . To illustrate this formalism, let us consider the special case where  $q_i = \{x, y, z\}$  and  $\mathbf{T}$  is diagonal. We would then write

$$\frac{1}{2}\tilde{\mathbf{q}}\mathbf{T}\dot{\mathbf{q}} = \frac{1}{2}(\dot{x}\dot{y}\dot{z}) \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \quad (8.24a)$$

and

$$\tilde{\mathbf{q}}\mathbf{a} = (\dot{x}\dot{y}\dot{z}) \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = a_x \dot{x} + a_y \dot{y} + a_z \dot{z} = \mathbf{a} \cdot \dot{\mathbf{r}}. \quad (8.24b)$$

In this notation the Hamiltonian,  $H = \tilde{\mathbf{q}}\mathbf{p} - L$ , becomes

$$H = \tilde{\mathbf{q}}(\mathbf{p} - \mathbf{a}) - \frac{1}{2}\tilde{\mathbf{q}}\mathbf{T}\tilde{\mathbf{q}} - L_0. \quad (8.24c)$$

The conjugate momenta, considered as a column matrix  $\mathbf{p}$ , is then, by Eq. (8.2), given as

$$\mathbf{p} = \mathbf{T}\dot{\mathbf{q}} + \mathbf{a}, \quad (8.25)$$

which can be inverted (step 4) to the column vector  $\dot{\mathbf{q}}$

$$\dot{\mathbf{q}} = \mathbf{T}^{-1}(\mathbf{p} - \mathbf{a}). \quad (8.26a)$$

This step presupposes that  $\mathbf{T}^{-1}$  exists, which it normally does by virtue of the positive definite property of kinetic energy.

The corresponding equation for  $\tilde{\mathbf{q}}$  is

$$\tilde{\mathbf{q}} = (\tilde{\mathbf{p}} - \tilde{\mathbf{a}})\mathbf{T}^{-1}. \quad (8.26b)$$

To obtain the correct functional form for the Hamiltonian, Eqs. (8.26) must be used to replace  $\dot{\mathbf{q}}$  and  $\tilde{\mathbf{q}}$ , yielding the final form for the Hamiltonian:

$$H(q, p, t) = \frac{1}{2}(\tilde{\mathbf{p}} - \tilde{\mathbf{a}})\mathbf{T}^{-1}(\mathbf{p} - \mathbf{a}) - L_0(q, t). \quad (8.27)$$

If the Lagrangian can be written in the form of Eq. (8.23), then we can immediately skip the intervening steps and write the Hamiltonian as Eq. (8.27). The inverse matrix  $\mathbf{T}^{-1}$  can usually most easily be obtained straightforwardly as

$$\mathbf{T}^{-1} = \frac{\tilde{\mathbf{T}}_c}{|\mathbf{T}|}, \quad (8.28)$$

where  $\mathbf{T}_c$  is the cofactor matrix whose elements  $(\mathbf{T}_c)_{jk}$  are  $(-1)^{j+k}$  times the determinant of the matrix obtained by striking out the  $j$ th row and the  $k$ th column of  $\mathbf{T}$ .

In the example Eq. (8.24a), these three matrices are given explicitly by

$$\mathbf{T} = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix}, \quad \mathbf{T}^{-1} = \begin{bmatrix} \frac{1}{m} & 0 & 0 \\ 0 & \frac{1}{m} & 0 \\ 0 & 0 & \frac{1}{m} \end{bmatrix}, \quad \text{and}$$

$$\tilde{\mathbf{T}}_c = \begin{bmatrix} m^2 & 0 & 0 \\ 0 & m^2 & 0 \\ 0 & 0 & m^2 \end{bmatrix},$$

and the determinant  $|\mathbf{T}| = m^3$ . It is easy to see that for the usual case when  $\mathbf{T}$  is diagonal, then  $\mathbf{T}^{-1}$  is also diagonal with elements that are just the reciprocals of the corresponding elements of  $\mathbf{T}$ .

A number of exercises in applying this formalism to various mechanical systems will be found in the problems at the end of the chapter. Two very simple examples are considered here because they illustrate some important aspects of the technique. First consider the spatial motion of a particle in a central force field, using spherical polar coordinates  $(r, \theta, \phi)$  for the generalized coordinates. The potential energy is some function  $V(r)$  and the kinetic energy is

$$T = \frac{mv^2}{2} = \frac{m}{2}(\dot{r}^2 + r^2 \sin^2 \theta \dot{\phi}^2 + r^2 \dot{\theta}^2). \quad (8.28')$$

Clearly the Hamiltonian has the form of Eq. (8.21) and corresponds to the total energy  $T + V$ . Since  $\mathbf{T}$  is diagonal the form of  $H$  is, by inspection,

$$H(r, \theta, p_r, p_\theta, p_\phi) = \frac{1}{2m} \left( p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2 \theta} \right) + V(r). \quad (8.29)$$

Note that the Hamiltonian would have a different functional form if the generalized coordinates were chosen to be the Cartesian coordinates  $x_i$  of the particle. If we make that choice, then the kinetic energy has the form

$$T = \frac{mv^2}{2} = \frac{m\dot{x}_i \dot{x}_i}{2},$$

so that the Hamiltonian is now

$$H(x_i, p_i) = \frac{p_i p_i}{2m} + V(r). \quad (8.30)$$

It is sometimes convenient to form the canonical momenta  $p_i$  conjugate to  $x_i$  into a vector  $\mathbf{p}$  such that the Hamiltonian can be written as

$$H(x_i, p_i) = \frac{\mathbf{p} \cdot \mathbf{p}}{2m} + V(\sqrt{x_i x^i}). \quad (8.31)$$

We can of course take the components of  $\mathbf{p}$  relative to any coordinate system we desire, curvilinear spherical coordinates, for example. But it is important not to confuse, say,  $p_\theta$  with the  $\theta$  component of  $\mathbf{p}$ , designated as  $(\mathbf{p})_\theta$ . The former is the canonical momentum conjugate to the coordinate  $\theta$ ; the latter is the  $\theta$  component of the momentum vector conjugate to the Cartesian coordinates. Dimensionally, it is clear they are quite separate quantities;  $p_\theta$  is an angular momentum,  $(\mathbf{p})_\theta$  is a linear momentum. *Whenever a vector is used from here on to represent canonical momenta it will refer to the momenta conjugate to Cartesian position coordinates.*

For a second example, let us consider a single (nonrelativistic) particle of mass  $m$  and charge  $q$  moving in an electromagnetic field. By Eq. (1.63), the Lagrangian for this system is

$$L = T - V = \frac{1}{2}mv^2 - q\phi + q\mathbf{A} \cdot \mathbf{v},$$

where the scalar potential term,  $-q\phi$ , is the  $L_0$  term of the Lagrangian as expressed in Eq. (8.22) and the vector potential term,  $q\mathbf{A} \cdot \mathbf{v}$ , is the  $L_1$  term.

## Chapter 8 The Hamilton Equations of Motion

Using Cartesian position coordinates as generalized coordinates, the Lagrangian can also be written as

$$L = \frac{m\dot{x}_i \dot{x}_i}{2} + qA_i \dot{x}_i - q\phi, \quad (8.32)$$

where the potentials  $\phi$  and  $A$  are in general functions of  $x_i$  and the time.

There is now a linear term in the generalized velocities such that the matrix  $a$  has the elements  $qA_{ij}$ . Because of this linear term in  $V$ , the Hamiltonian is *not*  $T + V$ . However, it is still in this case the total energy since the “potential” energy in an electromagnetic field is determined by  $\phi$  alone. The canonical momenta, either by Eq. (8.2) or Eq. (8.25), are

$$p_i = m\dot{x}_i + qA_i, \quad (8.33)$$

and the Hamiltonian (cf. Eq. (8.27)) is

$$H = \frac{(p_i - qA_i)(p_i - qA_i)}{2m} + q\phi, \quad (8.34)$$

which is the total energy of the particle. Again, the momenta  $p_i$  can be formed into a vector  $\mathbf{p}$  and  $H$  written as

$$H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 + q\phi, \quad (8.35)$$

and remembering that  $\mathbf{p}$  refers only to momenta conjugate to  $x_i$ .

It is clear that Hamilton's equations of motion do not treat the coordinates and momenta in a completely symmetric fashion. The equation for  $\dot{p}$  has a minus sign that is absent in the equation for  $\dot{q}$ . Considerable ingenuity has been exercised in devising nomenclature schemes that result in entirely symmetric equations, or combine the two sets into one. Most of these schemes have only curiosity value, but one has proved to be an elegant and powerful tool for manipulating the canonical equations and allied expressions.

For a system of  $n$  degrees of freedom, we construct a column matrix  $\boldsymbol{\eta}$  with  $2n$  elements such that

$$\eta_i = q_i, \quad \eta_{i+n} = p_i; \quad i \leq n. \quad (8.36)$$

Similarly, the column matrix  $\partial H / \partial \boldsymbol{\eta}$  has the elements

$$\left( \frac{\partial H}{\partial \boldsymbol{\eta}} \right)_i = \frac{\partial H}{\partial q_i}, \quad \left( \frac{\partial H}{\partial \boldsymbol{\eta}} \right)_{i+n} = \frac{\partial H}{\partial p_i}; \quad i \leq n. \quad (8.37)$$

Finally, let  $\mathbf{J}$  be the  $2n \times 2n$  square matrix composed of four  $n \times n$  zero and unit matrices according to the scheme

$$\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix} \quad (8.38a)$$

with the following transpose matrix, which is its inverse

$$\tilde{\mathbf{J}} = \begin{bmatrix} \mathbf{0} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}, \quad (8.38b)$$

which means

$$\tilde{\mathbf{J}}\mathbf{J} = \mathbf{J}\tilde{\mathbf{J}} = \mathbf{1} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}, \quad (8.38c)$$

so

$$\tilde{\mathbf{J}} = -\mathbf{J} = \mathbf{J}^{-1} \quad (8.38d)$$

and

$$\mathbf{J}^2 = -\mathbf{1}, \quad (8.38e)$$

and the determinant is

$$|\mathbf{J}| = +1. \quad (8.38f)$$

Here  $\mathbf{0}$  is the  $n \times n$  matrix all of whose elements is zero, and  $\mathbf{1}$  is the standard  $n \times n$  unit matrix. Hamilton's equations of motion can then be written in compact form as

$$\dot{\boldsymbol{\eta}} = \mathbf{J} \frac{\partial H}{\partial \boldsymbol{\eta}}. \quad (8.39)$$

For two coordinate variables, this has the expanded form

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\dot{p}_1 \\ -\dot{p}_2 \\ \dot{q}_1 \\ \dot{q}_2 \end{bmatrix}, \quad (8.40)$$

where use was made of Eqs. (8.37) and (8.18). This method of displaying the canonical equations of motion will be referred to as Hamilton's equations in matrix or *symplectic*\* notation. In subsequent chapters we shall frequently employ this matrix form of the equations.

## 8.2 ■ CYCLIC COORDINATES AND CONSERVATION THEOREMS

According to the definition given in Section 2.6, a cyclic coordinate  $q_j$  is one that does not appear explicitly in the Lagrangian; by virtue of Lagrange's equations

\*The term *symplectic* comes from the Greek for "intertwined," particularly appropriate for Hamilton's equations where  $\dot{q}$  is matched with a derivative with respect to  $p$  and  $\dot{p}$  similarly with the negative of a  $q$  derivative. H. Weyl first introduced the term in 1939 in his book *The Classical Groups*.

its conjugate momentum  $p_j$  is then a constant. But comparison of Eq. (8.14) with Eq. (8.16) has already told us that

$$\dot{p}_j = \frac{\partial L}{\partial q_j} = -\frac{\partial H}{\partial q_j}.$$

A coordinate that is cyclic will thus also be absent from the Hamiltonian.\* Conversely if a generalized coordinate does not occur in  $H$ , the conjugate momentum is conserved. The momentum conservation theorems of Section 2.6 can thus be transferred to the Hamiltonian formulation with no more than a substitution of  $H$  for  $L$ . In particular, the connection between the invariance or symmetry properties of the physical system and the constants of the motion can also be derived in terms of the Hamiltonian. For example, if a system is completely self-contained, with only internal forces between the particles, then the system can be moved as a rigid ensemble without affecting the forces or subsequent motion. The system is said to be invariant under a rigid displacement. Hence, a generalized coordinate describing such a rigid motion will not appear explicitly in the Hamiltonian, and the corresponding conjugate momentum will be conserved. If the rigid motion is a translation along some particular direction, then the conserved momentum is the corresponding Cartesian component of the total linear (canonical) momentum of the system. Since the direction is arbitrary, the total vector linear momentum is conserved. The rigid displacement may be a rotation, from whence it follows that the total angular momentum vector is conserved. Even if the system interacts with external forces, there may be a symmetry in the situation that leads to a conserved canonical momentum. Suppose the system is symmetrical about a given axis so that  $H$  is invariant under rotation about that axis. Then  $H$  obviously cannot involve the rotation angle about the axis and the particular angle variable must be a cyclic coordinate. It follows, as in Section 2.6, that the component of the angular momentum about that axis is conserved.<sup>†</sup>

The considerations concerning  $h$  in Section 2.7 have already shown that if  $L$  (and in consequence of Eq. (8.15), also  $H$ ) is not an explicit function of  $t$ , then  $H$  is a constant of motion. This can also be seen directly from the equations of motion (8.18) by writing the total time derivative of the Hamiltonian as

$$\frac{dH}{dt} = \frac{\partial H}{\partial q_i} \dot{q}_i + \frac{\partial H}{\partial p_i} \dot{p}_i + \frac{\partial H}{\partial t}.$$

In consequence of the equations of motion (8.18), the first two sums on the right cancel each other, and it therefore follows that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \quad (8.41)$$

\*This conclusion also follows from the definition of Eq. (8.15), for  $H$  differs from  $-L$  only by  $p_i \dot{q}_i$ , which does not involve  $q_i$  explicitly.

<sup>†</sup>The relation between conservation laws, symmetry of the Lagrangian, (and the Hamiltonian) of the system is called Noether's theorem. The formal proof is given in Section 13.7.

Thus if  $t$  doesn't appear explicitly in  $L$ , it will also not be present in  $H$ , and  $H$  will be constant in time.

Further, it was proved in Section 2.7 that if the equations of transformation that define the generalized coordinates (1.38),

$$\mathbf{r}_m = \mathbf{r}_m(q_1, \dots, q_n; t),$$

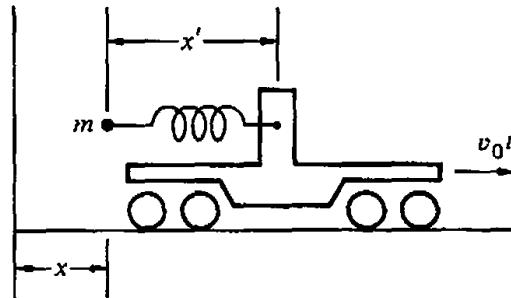
do not depend explicitly upon the time, and if the potential is velocity independent, then  $H$  is the total energy,  $T + V$ . The identification of  $H$  as a constant of the motion and as the total energy are two separate matters, and the conditions sufficient for the one are not enough for the other. It can happen that the Eqs. (1.38) do involve time explicitly but that  $H$  does not. In this case,  $H$  is a constant of the motion but it is *not* the total energy. As was also emphasized in Section (2.6), the Hamiltonian is dependent both in magnitude and in functional form upon the initial choice of generalized coordinates. For the Lagrangian, we have a specific prescription,  $L = T - V$ , and a change of generalized coordinates within that prescription may change the functional appearance of  $L$  but cannot alter its magnitude. On the other hand, use of a different set of generalized coordinates in the definition for the Hamiltonian, Eq. (8.15), may lead to an entirely different quantity for the Hamiltonian. It may be that for one set of generalized coordinates  $H$  is conserved, but that for another it varies in time.

To illustrate some of these points in a simple example, we may consider a somewhat artificial one-dimensional system. Suppose a point mass  $m$  is attached to a spring, of force constant  $k$ , the other end of which is fixed on a massless cart that is being moved uniformly by an external device with speed  $v_0$  (cf. Fig. 8.1). If we take as generalized coordinate the position  $x$  of the mass particle in the stationary system, then the Lagrangian of the system is obviously

$$L(x, \dot{x}, t) = T - V = \frac{m\dot{x}^2}{2} - \frac{k}{2}(x - v_0 t)^2. \quad (8.42)$$

(For simplicity, the origin has been chosen so that the cart passes through it at  $t = 0$ .) The corresponding equation of motion is clearly

$$m\ddot{x} = -k(x - v_0 t).$$



**FIGURE 8.1** A harmonic oscillator fixed to a uniformly moving cart.

An obvious way of solving this equation is to change the unknown to  $x'(t)$  defined as

$$x' = x - v_0 t, \quad (8.43)$$

and noting that  $\ddot{x}' = \ddot{x}$ , the equation of motion becomes

$$m\ddot{x}' = -kx'. \quad (8.44)$$

From Eq. (8.43),  $x'$  is the displacement of the particle relative to the cart; Eq. (8.44) says that to an observer on the cart the particle exhibits simple harmonic motion, as would be expected on the principle of equivalence in Galilean relativity.

Having looked at the nature of the motion, let us consider the Hamiltonian formulation. Since  $x$  is the Cartesian coordinate of the particle, and the potential does not involve generalized velocities, the Hamiltonian relative to  $x$  is the sum of the kinetic and potential energies, that is, the total energy. In functional form the Hamiltonian is given by

$$H(x, p, t) = T + V = \frac{p^2}{2m} + \frac{k}{2}(x - v_0 t)^2. \quad (8.45)$$

The Hamiltonian is the total energy of the system, but since it is explicitly a function of  $t$ , it is *not* conserved. Physically this is understandable; energy must flow into and out of the “external physical device” to keep the cart moving uniformly against the reaction of the oscillating particle.\*

Suppose now we formulated the Lagrangian from the start in terms of the relative coordinate  $x'$ . The same prescription gives the Lagrangian as

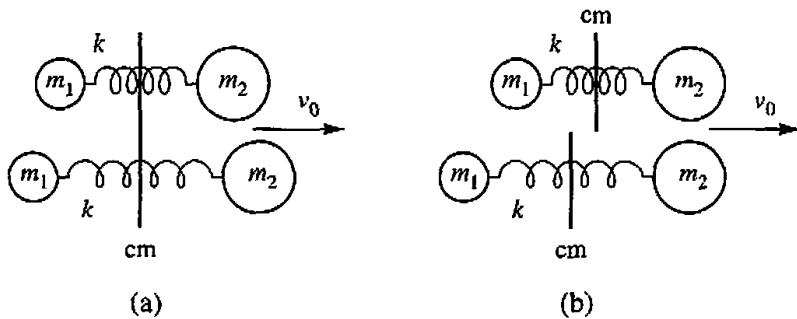
$$L(x', \dot{x}') = \frac{m\dot{x}'^2}{2} + m\dot{x}'v_0 + \frac{mv_0^2}{2} - \frac{kx'^2}{2}. \quad (8.46)$$

In setting up the corresponding Hamiltonian, we note there is now a term linear in  $\dot{x}'$ , with the single component of  $a$  being  $mv_0$ . The new Hamiltonian is now

$$H'(x', p') = \frac{(p' - mv_0)^2}{2m} + \frac{kx'^2}{2} - \frac{mv_0^2}{2}. \quad (8.47)$$

Note that the last term is a constant involving neither  $x'$  nor  $p'$ ; it could, if we wished, be dropped from  $H'$  without affecting the resultant equations of motion. Now  $H'$  is *not* the total energy of the system, but it *is* conserved. Except for the last term, it can be easily identified as the total energy of motion of the particle relative to the moving cart. The two Hamiltonian's are different in magnitude.

\*Put another way, the moving cart constitutes a time-dependent constraint on the particle, and the force of the constraint does do work in actual (*not* virtual) displacement of the system.



**FIGURE 8.2** Vibrating dumbbell under two conditions: (a) freely oscillating, and (b) oscillating with mass  $m_2$  kept at a constant velocity

time dependence, and functional behavior. But the reader can easily verify that both lead to the same motion for the particle.

Additional insight into the problem of the mass cart previously discussed can be gained by considering a dumbbell of two masses connected by a spring of constant  $k$ . We shall consider the case where the center of mass of the dumbbell is in constant motion at a speed  $v_0$  along the direction determined by the spring and allow oscillations of the masses only along this direction. This is shown in Fig. 8.2, where C-O-M denotes the center of mass.

The dumbbell is made to vibrate while its center of mass has an initial velocity  $v_0$ . It will continue with this velocity with uniform translational motion. This translational motion will have no effect on the oscillations. The motion of the center of mass and the motion relative to the center of mass separate as they do in the Kepler problem. Once the motion is started, energy is conserved and the Hamiltonian is the total conserved energy. The situation is different if the mass  $m_2$  moves at the constant speed  $v_0$  since a periodic force is applied. The center of mass and the mass  $m_1$  then oscillate relative to  $m_2$ . Since a changing external force must be applied to the system to keep  $m_2$  at the constant velocity  $v_0$ , the Hamiltonian is no longer conserved, nor is the Hamiltonian the total energy.

### 8.3 ■ ROUTH'S PROCEDURE

It has been remarked that the Hamiltonian formulation is not particularly helpful in the direct solution of mechanical problems. Often we can solve the  $2n$  first-order equations only by eliminating some of the variables, for example, the  $p$  variables, which speedily leads back to the second-order Lagrangian equations of motion. But an important exception should be noted. The Hamiltonian procedure is especially adapted to the treatment of problems involving cyclic coordinates.

Let us consider the situation in Lagrangian formulation when some coordinate, say  $q_n$ , is cyclic. The Lagrangian as a function of  $q$  and  $\dot{q}$  can then be written

$$L = L(q_1, \dots, q_{n-1}; \dot{q}_1, \dots, \dot{q}_n; t).$$

All the generalized velocities still occur in the Lagrangian and in general will be functions of the time. We still have to solve a problem of  $n$  degrees of freedom, even though one degree of freedom corresponds to a cyclic coordinate. A cyclic coordinate in the Hamiltonian formulation, on the other hand, truly deserves its alternative description as "ignorable," for in the same situation  $p_n$  is some constant  $\alpha$ , and  $H$  has the form

$$H = H(q_1, \dots, q_{n-1}; p_1, \dots, p_{n-1}; \alpha; t).$$

In effect, the Hamiltonian now describes a problem involving only  $n - 1$  coordinates, which may be solved completely ignoring the cyclic coordinate except as it is manifested in the constant of integration  $\alpha$ , to be determined from the initial conditions. The behavior of the cyclic coordinate itself with time is then found by integrating the equation of motion

$$\dot{q}_n = \frac{\partial H}{\partial \alpha}.$$

The advantages of the Hamiltonian formulation in handling cyclic coordinates may be combined with the Lagrangian conveniences for noncyclic coordinates by a method devised by Routh. Essentially, we carry out a mathematical transformation from the  $q, \dot{q}$  basis to the  $q, p$  basis only for those coordinates that are cyclic, obtaining their equations of motion in the Hamiltonian form, while the remaining coordinates are governed by Lagrange equations. If the cyclic coordinates are labeled  $q_{s+1}, \dots, q_n$ , then a new function  $R$  (known as the Routhian) may be introduced, defined as

$$R(q_1, \dots, q_n; \dot{q}_1, \dots, \dot{q}_s; p_{s+1}, \dots, p_n; t) = \sum_{i=s+1}^n p_i \dot{q}_i - L, \quad (8.48)$$

which is equivalent to writing

$$\begin{aligned} R(q_1, \dots, q_n; \dot{q}_1, \dots, \dot{q}_s; p_{s+1}, \dots, p_n; t) &= \\ H_{cycl}(p_{s+1}, \dots, p_n) - L_{noncycl}(q_1, \dots, q_s; \dot{q}_1, \dots, \dot{q}_s). \end{aligned} \quad (8.49)$$

It is easy to show for the  $s$  nonignorable coordinates, the Lagrange equations

$$\frac{d}{dt} \left( \frac{\partial R}{\partial \dot{q}_i} \right) - \frac{\partial R}{\partial q_i} = 0, \quad i = 1, \dots, s, \quad (8.50)$$

are satisfied, while for the  $n - s$  ignorable coordinates, Hamilton's equations apply as

$$\frac{\partial R}{\partial q_i} = -\dot{p}_i = 0, \quad \text{and} \quad \frac{\partial R}{\partial p_i} = \dot{q}_i, \quad i = s + 1, \dots, n. \quad (8.51)$$

A simple, almost trivial, example may clarify Routh's procedure and the physical significance of the quantities involved. Consider the Kepler problem investi-

gated in Section 3.7, that of a single particle moving in a plane under the influence of the inverse-square central force  $f(r)$  derived from the potential  $V(r) = -k/r^n$ . The Lagrangian is then

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{k}{r^n}.$$

As noted before, the ignorable coordinate is  $\theta$ , and if the constant conjugate momentum is denoted by  $p_\theta$ , the corresponding Routhian (8.49) is

$$R(r, \dot{r}, p_\theta) = \frac{p_\theta^2}{2mr^2} - \frac{1}{2}m\dot{r}^2 - \frac{k}{r^n}.$$

Physically we see that the Routhian is the equivalent one-dimensional potential  $V'(r)$  minus the kinetic energy of radial motion.

Applying the Lagrange equation (8.50) to the noncyclic radial coordinate  $r$ , we obtain the equation of motion (3.11)

$$\ddot{r} - \frac{p_\theta^2}{mr^3} + \frac{nk}{r^{n+1}} = 0. \quad (8.52)$$

Applying Hamilton's equation (8.51) to the cyclic variable  $\theta$ , we obtain the pair of equations

$$\dot{p}_\theta = 0 \quad \text{and} \quad \frac{p_\theta}{mr^2} = \dot{\theta}. \quad (8.53)$$

whose solution is the same as Eq. (3.8),

$$p_\theta = mr^2\dot{\theta} = l = \text{constant}.$$

Typically, Routh's procedure does not add to the physics of the analysis presented earlier in Chapter 3, but it makes the analysis more automatic. In complicated problems with many degrees of freedom, this feature can be a considerable advantage. It is not surprising therefore that Routh's procedure finds its greatest usefulness in the direct solution of problems relating to engineering applications. But as a fundamental entity, the Routhian is a sterile hybrid, combining some of the features of both the Lagrangian and the Hamiltonian pictures. For the development of various formalisms of classical mechanics, the complete Hamiltonian formulation is more fruitful.

## 8.4 ■ THE HAMILTONIAN FORMULATION OF RELATIVISTIC MECHANICS

As with the Lagrangian picture in special relativity, two attitudes can be taken to the Hamiltonian formulation of relativistic mechanics. The first makes no pretense at a covariant description but instead works in some specific Lorentz or inertial frame. Time as measured in the particular Lorentz frame is then not treated on a

## Chapter 8 The Hamilton Equations of Motion

common basis with other coordinates but serves, as in nonrelativistic mechanics, as a parameter describing the evolution of the system. Nonetheless, if the Lagrangian that leads to the Hamiltonian is itself based on a relativistically invariant physical theory (for example, Maxwell's equations and the Lorentz force), then the resultant Hamiltonian picture will be relativistically correct. The second approach of course attempts a fully covariant description of the Hamiltonian picture, but the difficulties that plagued the corresponding Lagrangian approach (cf. Section 7.9) are even fiercer here. We shall consider the noncovariant method first.

For a single-particle Lagrangian of the form of Eq. (7.136),

$$L = -mc^2\sqrt{1 - \beta^2} - V,$$

we have already shown that the Hamiltonian (in the guise of the energy function  $h$ ) is the total energy of the system:

$$H = T + V.$$

The energy  $T$  can be expressed in terms of the canonical momenta  $p_i$  (Eq. 7.139) through Eq. (7.38):\*

$$T^2 = p^2c^2 + m^2c^4,$$

so that a suitable form for the Hamiltonian is

$$H = \sqrt{p^2c^2 + m^2c^4} + V. \quad (8.54)$$

When the system consists of a single particle moving in an electromagnetic field, the Lagrangian has been given as (cf. Eq. (7.141))

$$L = -mc^2\sqrt{1 - \beta^2} + q\mathbf{A} \cdot \mathbf{v} - q\phi.$$

The term in  $L$  linear in the velocities does not appear explicitly in the Hamiltonian (cf. Eq. (8.54)), as we have seen, whereas the first term leads to the appearance of  $T$  in the Hamiltonian. Thus, the Hamiltonian is again the total particle energy:

$$H = T + q\phi. \quad (8.55)$$

For this system, the canonical momenta conjugate to the Cartesian coordinates of the particle are defined by (cf. Eq. (7.142))

$$p^i = mu^i + qA^i,$$

so that the relation between  $T$  and  $p^i$  is given by Eq. (7.168), and the Hamiltonian has the final form

\*In this section we use  $T$  for the motion energy ( $pc$ ) plus the rest energy ( $mc^2$ ) to avoid confusing it with the total energy  $T + V$

$$H = \sqrt{(\mathbf{p} - q\mathbf{A})^2 c^2 + m^2 c^4} + q\phi. \quad (8.56)$$

It should be emphasized again that  $\mathbf{p}$  here is the vector of the canonical momenta conjugate to the *Cartesian* position coordinates of the particle. We may also note that  $(H - q\phi)/c$  is the zeroth component of the 4-vector

$$mu^v + qA^v$$

(cf. Eqs. (7.27), (7.38'), and (7.166)). While the Hamiltonian (8.56) is not expressed in covariant fashion, it does have a definite transformation behavior under a Lorentz transformation as being, in some Lorentz form, the zeroth component of a 4-vector.

In a covariant approach to the Hamiltonian formulation, time must be treated in the same fashion as the space coordinates; that is, time must be taken as one of the canonical coordinates having an associated conjugate momentum. The foundations of such an extension of the dimensionality of phase space can in fact be constructed even in nonrelativistic mechanics. Following the pattern of Section 7.10, the progress of the system point along its trajectory in phase space can be marked by some parameter  $\theta$ , and  $t$  “released,” so to speak, to serve as an additional coordinate. If derivatives with respect to  $\theta$  are denoted by a superscript prime, the Lagrangian in the  $(q_1, \dots, q_n; t)$  configuration space is (cf. Eq. (7.159))

$$\Lambda(q, q', t, t') = t'L\left(q, \frac{q'}{t'}, t\right). \quad (8.57)$$

The momentum conjugate to  $t$  is then

$$p_t = \frac{\partial \Lambda}{\partial t'} = L + t' \frac{\partial L}{\partial t'}.$$

If we make explicit use of the connection  $\dot{q} = q'/t'$ , this relation becomes

$$p_t = L - \frac{q'_i}{t'} \frac{\partial L}{\partial \dot{q}_i} = L - \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} = -H. \quad (8.58)$$

The momentum conjugate to the time “coordinate” is therefore the negative of the ordinary Hamiltonian.\* While the framework of this derivation is completely non-relativistic, the result is consistent with the identification of the time component of the 4-vector momentum with  $E/c$ . As can be seen from the definition, Eq. (8.2), if  $q$  is multiplied by a constant  $\alpha$ , then the conjugate momentum is divided by  $\alpha$ . Hence, the canonical momentum conjugate to  $ct$  is  $H/c$ .

\*The remaining momenta are unchanged by the shift from  $t$  to  $\theta$ , as can be seen by evaluating the corresponding derivative

$$\frac{\partial \Lambda}{\partial q'_i} = t' \frac{\partial L}{\partial q'_i} = t' \left( \frac{\partial L}{\partial \dot{q}} \frac{1}{t'} \right) = p_i.$$

Thus, there seems to be a natural route available for constructing a relativistically covariant Hamiltonian. But the route turns out to be mined with booby traps. It will be recalled that the covariant Lagrangian used to start the process, Eq. (7.159) or Eq. (8.57), is homogeneous in first degree in the generalized velocities  $q'$ , and for such a Lagrangian the recipe described above for constructing the Hamiltonian formulation breaks down irreparably. If  $L$  is of type  $L_1$ , the corresponding Hamiltonian, call it  $H_c(q, t, p, p_t)$ , is identically zero!

Fortunately, there does not seem to be any compelling reason why the covariant Lagrangian has to be homogeneous in the first degree, at least for classical relativistic mechanics. It has already been seen that for a single free particle the covariant Lagrangian

$$\Lambda(x^\mu, u^\mu) = \frac{1}{2}mu_\mu u^\mu$$

leads to the correct equations of motion. Of course the four-velocity components,  $u^\mu$ , are still not all independent, but the constraint can be treated as a “weak condition” to be imposed only *after* all the differentiations have been carried through. There is now no difficulty in obtaining a Hamiltonian from this Lagrangian, by the same route as in nonrelativistic mechanics; the result is clearly

$$H_c = \frac{p_\mu p^\mu}{2m}. \quad (8.59)$$

For a single particle in an electromagnetic field, a covariant Lagrangian has been found previously: (cf. Eq. (7.165))\*

$$\Lambda(x^\mu, u^\mu) = \frac{1}{2}mu_\mu u^\mu + qu^\mu A_\mu(x_\lambda), \quad (7.147)$$

with the canonical momenta (cf. Eq. (7.167)),

$$p_\mu = mu_\mu + qA_\mu. \quad (7.149)$$

In the corresponding Hamiltonian, the term linear in  $u_\mu$  does not appear explicitly in the Hamiltonian, and the remaining  $L_2$  part in terms of the canonical momenta is

$$H'_c = \frac{(p_\mu - qA_\mu)(p^\mu - qA^\mu)}{2m}. \quad (8.60)$$

Both Hamiltonians, Eqs. (8.59) and (8.60), are constant, with the same value,  $-mc^2/2$ , but to obtain the equations of motion it is the *functional* dependence on the 4-vectors of position and momenta that is important. With a system of one particle, the covariant Hamiltonian leads to eight first-order equations of motion

\*The Legendre transformation process is reversible: Given a Hamiltonian we can obtain the corresponding Lagrangian (cf. Derivation 1). But the difficulties also arise in either direction. If a given Hamiltonian is postulated to be homogeneous in first degree in the momenta, then it is not possible to find an equivalent Lagrangian.

$$\frac{dx^\nu}{d\tau} = \frac{\partial H'_c}{\partial p^\nu}, \quad \frac{dp^\nu}{d\tau} = -\frac{\partial H_c}{\partial x^\nu}. \quad (8.61)$$

We know that these equations cannot be all independent. The space parts of Eqs. (8.61) obviously lead to the spatial equations of motion. We should expect therefore that the remaining two equations tell us nothing new, exactly as in the Lagrangian case. This can be verified by examining the  $\nu = 0$  equations in some particular Lorentz frame. One of them is the constitutive equation for  $p^0$ :

$$u^0 = \frac{\partial H'_c}{\partial p^0} = \frac{1}{m} (p^0 - qA^0)$$

or

$$p^0 = \frac{1}{c} (T + q\phi) = \frac{H'_c}{c}, \quad (8.62)$$

a general conclusion that has been noted before. The other can be written as

$$\frac{1}{\sqrt{1 - \beta^2}} \frac{dp^0}{dt} = -\frac{1}{c} \frac{\partial H_c}{\partial t}$$

or

$$\frac{dH}{dt} = \sqrt{1 - \beta^2} \frac{\partial H_c}{\partial t}. \quad (8.63)$$

As with the covariant Lagrangian formulation, we have the problem of finding suitable covariant potential terms in the Lagrangian to describe the forces other than electromagnetic. In multiparticle systems we are confronted in full measure with the critical difficulties of including interactions other than with fields. In Hamiltonian language, the "no-interaction" theorem already referred to in Section 7.10 says that only in the absence of direct particle interactions can Lorentz invariant systems be described in terms of the usual position coordinates and corresponding canonical momenta. The scope of the relativistic Hamiltonian framework is therefore quite limited and so for the most part we shall confine ourselves to nonrelativistic mechanics.

## 8.5 ■ DERIVATION OF HAMILTON'S EQUATIONS FROM A VARIATIONAL PRINCIPLE

Lagrange's equations have been shown to be the consequence of a variational principle, namely, the Hamilton's principle of Section 2.1. Indeed, the variational method is often the preferable one for deriving Lagrange's equations, for it is applicable to types of systems not usually included within the scope of mechanics. It would be similarly advantageous if a variational principle could be found that

leads directly to the Hamilton's equations of motion. Hamilton's principle,

$$\delta I \equiv \delta \int_{t_1}^{t_2} L dt = 0, \quad (8.64)$$

lends itself to this purpose, but as formulated originally it refers to paths in configuration space. The first modification therefore is that the integral must be evaluated over the trajectory of the system point in phase space, and the varied paths must be in the neighborhood of this phase space trajectory. In the spirit of the Hamiltonian formulation, both  $q$  and  $p$  must be treated as independent coordinates of phase space, to be varied independently. To this end the integrand in the action integral, Eq. (8.64), must be expressed as a function of both  $q$  and  $p$ , and their time derivatives, through Eq. (8.15). Equation (8.64) then appears as

$$\delta I = \delta \int_{t_1}^{t_2} (p_i \dot{q}_i - H(q, p, t)) dt = 0. \quad (8.65)$$

As a variational principle in phase space, Eq. (8.65) is sometimes referred to as the *modified Hamilton's principle*. Although it will be used most frequently in connection with transformation theory (see Chapter 9), the main interest in it here is to show that the principle leads to Hamilton's canonical equations of motion.

The modified Hamilton's principle is exactly of the form of the variational problem in a space of  $2n$  dimensions considered in Section 2.3 (cf. Eq. (2.14)):

$$\delta I = \delta \int_{t_1}^{t_2} f(q, \dot{q}, p, \dot{p}, t) dt = 0, \quad (8.66)$$

for which the  $2n$  Euler–Lagrange equations are

$$\frac{d}{dt} \left( \frac{\partial f}{\partial \dot{q}_j} \right) - \frac{\partial f}{\partial q_j} = 0 \quad j = 1, \dots, n \quad (8.67)$$

$$\frac{d}{dt} \left( \frac{\partial f}{\partial \dot{p}_j} \right) - \frac{\partial f}{\partial p_j} = 0 \quad j = 1, \dots, n. \quad (8.68)$$

The integrand  $f$  as given in Eq. (8.65) contains  $\dot{q}_j$  only through the  $p_i \dot{q}_i$  term, and  $q_j$  only in  $H$ . Hence, Eqs. (8.67) lead to

$$\dot{p}_j + \frac{\partial H}{\partial q_j} = 0. \quad (8.69)$$

On the other hand, there is no explicit dependence of the integrand in Eq. (8.65) on  $\dot{p}_j$ . Equations (8.68) therefore reduce simply to

$$\dot{q}_j - \frac{\partial H}{\partial p_j} = 0. \quad (8.70)$$

Equations (8.69) and (8.70) are exactly Hamilton's equations of motion. Eqs. (8.18). The Euler–Lagrange equations of the modified Hamilton's principle are thus the desired canonical equations of motion.

This derivation of Hamilton's equations from the variational principle is so brief as to give the appearance of a sleight-of-hand trick. One wonders whether something extra has been sneaked in while we were being misdirected by the magician's patter. Is the modified Hamilton's principle equivalent to Hamilton's principle, or does it contain some additional physics? The question is largely irrelevant; the primary justification for the modified Hamilton's principle is that it leads to the canonical equations of motion in phase space. After all, no further argument was given for the validity of Hamilton's principle than that it corresponded to the Lagrangian equations of motion. So long as Hamiltonian can be constructed, the Legendre transformation procedure shows that the Lagrangian and Hamiltonian formulations, and therefore their respective variational principles, have the same physical content.

One question that can be raised however is whether the derivation puts limitations on the variation of the trajectory that are not present in Hamilton's principle. The variational principle leading to the Euler–Lagrange equations is formulated, as in Section 2.2, such that the variations of the independent variables vanish at the end points. In phase space, that would require  $\delta q_i = 0$  and  $\delta p_i = 0$  at the end points, whereas Hamilton's principle requires only the vanishing of  $\delta q_i$  under the same circumstances. A look at the derivation as spelled out in Section 2.2 will show however that the variation is required to be zero at the end points only in order to get rid of the integrated terms arising from the variations in the time derivatives of the independent variables. While the  $f$  function in Eq. (8.66) that corresponds to the modified Hamilton's principle, Eq. (8.65), is indeed a function of  $\dot{q}_j$ , there is no explicit appearance of  $\dot{p}_j$ . Equations (8.68) and therefore (8.70) follow from Eq. (8.65) without stipulating the variations of  $p_j$  at the end points. The modified Hamilton's principle, with the integrand  $L$  defined in terms of the Hamiltonian by Eq. (8.19), leads to Hamilton's equations under the same variation conditions as those in Hamilton's principle.\*

Nonetheless, there are advantages to requiring that the varied paths in the modified Hamilton's principle return to the same end points in both  $q$  and  $p$ , for we then have a more generalized condition for Hamilton's equations of motion. As with Hamilton's principle, if there is no variation at the end points we can add a total time derivative of any arbitrary (twice-differentiable) function  $F(q, p, t)$  to the integrand without affecting the validity of the variational principle. Suppose, for example, we subtract from the integrand of Eq. (8.65) the quantity

\*It may be objected that  $q$  and  $p$  cannot be varied independently, because the defining Eqs. (8.2) link  $p$  with  $q$  and  $\dot{q}$ . We could not then have a variation of  $q$  (and  $\dot{q}$ ) without a corresponding variation of  $p$ . But this entire objection is completely at variance with the intent and the spirit of the Hamiltonian picture. Once the Hamiltonian formulation has been set up, Eqs. (8.2) form no part of it. The momenta have been elevated to the status of independent variables, on an equal basis with the coordinates and connected with them and the time only through the medium of the equations of motion themselves and not by any a priori defining relationship.

$$\frac{d}{dt}(q_i p_i).$$

The modified Hamilton's principle would then read

$$\delta \int_{t_1}^{t_2} (-\dot{p}_i q_i - H(q, p, t)) dt = 0. \quad (8.71)$$

Here the  $f$  integrand of Eq. (8.66) is a function of  $\dot{p}$ , and it is easily verified that the Euler–Lagrange equations (8.67) and (8.68) with this  $f$  again correspond to Hamilton's equations of motion, Eqs. (8.18). Yet the integrand in Eq. (8.71) is not the Lagrangian nor can it in general be simply related to the Lagrangian by a point transformation in configuration space. By restricting the variation of both  $q$  and  $p$  to be zero at the end points, the modified Hamilton's principle provides an independent and general way of setting up Hamilton's equations of motion without a prior Lagrangian formulation. If you will, it does away with the necessity of a linkage between the Hamiltonian canonical variables and a corresponding Lagrangian set of generalized coordinates and velocities. This will be very important to us in the next chapter where we examine transformations of phase space variables that preserve the Hamiltonian form of the equations of motion.

The requirement of independent variation of  $q$  and  $p$ , so essential for the above derivation, highlights the fundamental difference between the Lagrangian and Hamiltonian formulations. Neither the coordinates  $q_i$ , nor the momenta  $p_i$ , are to be considered there as the more fundamental set of variables; both are equally independent. Only by broadening the field of independent variables from  $n$  to  $2n$  quantities are we enabled to obtain equations of motion that are of first order. In a sense, the names “coordinates” and “momenta” are unfortunate, for they bring to mind pictures of spatial coordinates and linear, or at most, angular momenta. A wider meaning must now be given to the terms. The division into coordinates and momenta corresponds to no more than a separation of the independent variables describing the motion into two groups having an almost symmetrical relationship to each other through Hamilton's equations.

## 8.6 ■ THE PRINCIPLE OF LEAST ACTION

Another variational principle associated with the Hamiltonian formulation is known as *the principle of least action*. It involves a new type of variation, which we shall call the  $\Delta$ -variation, requiring detailed explanation. In the  $\delta$ -variation process used in the discussion of Hamilton's principle in Chapter 2, the varied path in configuration space always terminated at end points representing the system configuration at the same time  $t_1$  and  $t_2$  as the correct path. To obtain Lagrange's equations of motion, we also required that the varied path return to the same end points in configuration space, that is,  $\delta q_i(t_1) = \delta q_i(t_2) = 0$ . The  $\Delta$ -variation is less constrained; in general, the varied path over which an integral is evaluated may end at different times than the correct path, and there

may be a variation in the coordinates at the end points. We can however use the same parameterization of the varied path as in the  $\delta$ -variation. In the notation of Section 2.3, a family of possible varied paths is defined by functions (cf. Eq. (2.15))

$$q_i(t, \alpha) = q_i(t, 0) + \alpha \eta_i(t), \quad (8.72)$$

where  $\alpha$  is an infinitesimal parameter that goes to zero for the correct path. Here the functions  $\eta_i$  do not necessarily have to vanish at the end points, either the original or the varied. All that is required is that they be continuous and differentiable. Figure 8.3 illustrates the correct and varied path for a  $\Delta$ -variation in configuration space.

Let us evaluate the  $\Delta$ -variation of the action integral:

$$\Delta \int_{t_1}^{t_2} L dt \equiv \int_{t_1+\Delta t_1}^{t_2+\Delta t_2} L(\alpha) dt - \int_{t_1}^{t_2} L(0) dt, \quad (8.73)$$

where  $L(\alpha)$  means the integral is evaluated along the varied path and  $L(0)$  correspondingly refers to the actual path of motion. The variation is clearly composed of two parts. One arises from the change in the limits of the integral; to first-order infinitesimals, this part is simply the integrand on the actual path times the difference in the limits in time. The second part is caused by the change in the integrand on the varied path, but now between the same time limits as the original integral. We may therefore write the  $\Delta$ -variation of the action integral as

$$\Delta \int_{t_1}^{t_2} L dt = L(t_2)\Delta t_2 - L(t_1)\Delta t_1 + \int_{t_1}^{t_2} \delta L dt. \quad (8.74)$$

Here the variation in the second integral can be carried out through a parameterization of the varied path, exactly as for Hamilton's principle except that the

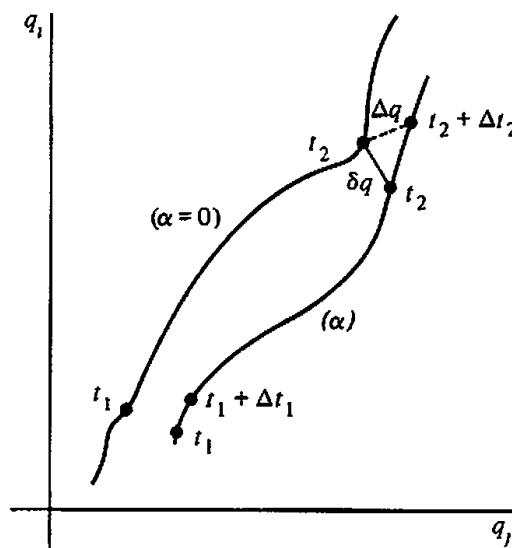


FIGURE 8.3 The  $\Delta$ -variation in configuration space.

variation in  $q_i$  does not vanish at the end points. The end point terms arising in the integration by parts must be retained, and the integral term on the right appears as

$$\int_{t_1}^{t_2} \delta L dt = \int_{t_1}^{t_2} \left[ \frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta q_i dt + \frac{\partial L}{\partial \dot{q}_i} \delta q_i \Big|_1^2.$$

By Lagrange's equations the quantities in the square brackets vanish, and the  $\Delta$ -variation therefore takes the form

$$\Delta \int_{t_1}^{t_2} L dt = (L \Delta t + p_i \delta q_i) \Big|_1^2. \quad (8.75)$$

In Eq. (8.75),  $\delta q_i$  refers to the variation in  $q_i$  at the original end point times  $t_1$  and  $t_2$ . We would like to express the  $\Delta$ -variation in terms of the change  $\Delta q_i$  between  $q_i$  at the end points of the actual path and  $q_i$  at the end points of the varied path, including the change in end point times. It is clear from Fig. 8.3 that these two variations are connected by the relation\*

$$\Delta q_i = \delta q_i + \dot{q}_i \Delta t. \quad (8.76)$$

Hence, Eq. (8.75) can be rewritten as

$$\Delta \int_{t_1}^{t_2} L dt = (L \Delta t - p_i \dot{q}_i \Delta t + p_i \Delta q_i) \Big|_1^2$$

or

$$\Delta \int_{t_1}^{t_2} L dt = (p_i \Delta q_i - H \Delta t) \Big|_1^2. \quad (8.77)$$

To obtain the principle of least action, we restrict our further considerations by three important qualifications:

1. Only systems are considered for which  $L$ , and therefore  $H$ , are not explicit functions of time, and in consequence  $H$  is conserved.
2. The variation is such that  $H$  is conserved on the varied path as well as on the actual path.
3. The varied paths are further limited by requiring that  $\Delta q_i$  vanish at the end points (but not  $\Delta t$ ).

\*Equation (8.76) may be derived formally from the parameter form, Eq. (8.72), of the varied path. Thus, at the upper end point we have

$$\Delta q_i(2) = q_i(t_2 + \Delta t_2, \alpha) - q_i(t_2, 0) = q_i(t_2 + \Delta t_2, 0) - q_i(t_2, 0) + \alpha \eta_i(t + \Delta t_2),$$

which to first order in small quantities  $\alpha$  and  $\Delta t_2$  is

$$\Delta q_i(2) = \dot{q}_i(2) \Delta t_2 + \delta q_i(2),$$

which is what Eq. (8.76) predicts

The nature of the resultant variation may be illustrated by noting that the varied path satisfying these conditions might very well describe the same curve in configuration space as the actual path. The difference will be the speed with which the system point traverses this curve; that is, the functions  $q_i(t)$  will be altered in the varied path. In order then to preserve the same value of the Hamiltonian at all points on the varied path, the times of the end points must be changed. With these three qualifications satisfied, the  $\Delta$ -variation of the action integral, Eq. (8.77), reduces to

$$\Delta \int_{t_1}^{t_2} L dt = -H(\Delta t_2 - \Delta t_1). \quad (8.78)$$

But under the same conditions, the action integral itself becomes

$$\int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} p_i \dot{q}_i dt - H(t_2 - t_1),$$

the  $\Delta$ -variation of which is

$$\Delta \int_{t_1}^{t_2} L dt = \Delta \int_{t_1}^{t_2} p_i \dot{q}_i dt - H(\Delta t_2 - \Delta t_1). \quad (8.79)$$

Comparison of Eqs. (8.78) and (8.79) finally gives the *principle of least action*:\*

$$\Delta \int_{t_1}^{t_2} p_i \dot{q}_i dt = 0. \quad (8.80)$$

By way of caution, note that the modified Hamilton's principle can be written in a form with a superficial resemblance to Eq. (8.80). If the trajectory of the system point is described by a parameter  $\theta$ , as in Sections 7.10 and 8.4, the modified Hamilton's principle appears as

$$\delta \int_{\theta_1}^{\theta_2} (p_i \dot{q}_i - H)t' d\theta = 0. \quad (8.81)$$

It will be recalled (cf. footnote on p. 351) that the momenta  $p_i$  do not change under the shift from  $t$  to  $\theta$ , and that  $\dot{q}_i t' = q'_i$ . Further, the momentum conjugate to  $t$  is  $-H$ . Hence, Eq. (8.81) can be rewritten as

$$\delta \int_{\theta_1}^{\theta_2} \sum_{i=1}^{n+1} p_i q'_i d\theta = 0, \quad (8.82)$$

where  $t$  has been denoted by  $q_{n+1}$ . There should however be no confusion between Eq. (8.82) and the principle of least action. Equations (8.82) involve phase

\*The integral in Eq. (8.80) is usually referred to in the older literature as the action, or action integral, and the first edition of this book followed the same practice. It is now customary to refer to the integral in Hamilton's principle as the action, and we have accepted this usage here. Sometimes the integral in Eq. (8.80) is designated as the *abbreviated action*.

space of  $(2n + 2)$  dimensions, as is indicated by the explicit summation to  $i = n + 1$ , whereas Eq. (8.80) is in the usual configuration space. But most important, the principle of least action is in terms of a  $\Delta$ -variation for constant  $H$ , while Eq. (8.82) employs the  $\delta$ -variation, and  $H$  in principle could be a function of time. Equation (8.82) is nothing more than the modified Hamilton's principle, and the absence of a Hamiltonian merely reflects the phenomenon that the Hamiltonian vanishes identically for the "homogeneous problem."

The least action principle itself can be exhibited in a variety of forms. In non-relativistic mechanics, if the defining equations for the generalized coordinates do not involve the time explicitly, then the kinetic energy is a quadratic function of the  $\dot{q}_i$ 's (cf. Eq. (1.71)):

$$T = \frac{1}{2} M_{jk}(q) \dot{q}_j \dot{q}_k. \quad (8.83)$$

When in addition the potential is not velocity dependent, the canonical momenta are derived from  $T$  only, and in consequence

$$p_i \dot{q}_i = 2T.$$

The principle of least action for such systems can therefore be written as

$$\Delta \int_{t_1}^{t_2} T dt = 0. \quad (8.84)$$

If, further, there are no external forces on the system, as, for example, a rigid body with no net applied forces, then  $T$  is conserved along with the total energy  $H$ . The least action principle then takes the special form

$$\Delta(t_2 - t_1) = 0. \quad (8.85)$$

Equation (8.85) states that of all paths possible between two points, consistent with conservation of energy, the system moves along that particular path for which the time of transit is the least (more strictly, an extremum). In this form the principle of least action recalls Fermat's principle in geometrical optics that a light ray travels between two points along such a path that the time taken is the least. We discussed these considerations in Section 10-8 of the Second Edition when we considered the connection between the Hamiltonian formulation and geometrical optics.

In Section 7.4 we discussed the infinitesimal interval in a metric space giving the interval as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (7.32')$$

where  $g_{\mu\nu}$  was the metric of a possibly curvilinear space and  $ds^2$  was the interval traversed for displacements given by  $dx^\mu$ . We can do something entirely similar here whenever  $T$  is of the form of Eq. (8.83). A configuration space is therefore constructed for which the  $M_{jk}$  coefficients form the metric tensor. In general, the

space will be curvilinear and nonorthogonal. The element of path length in the space is then defined by (cf. Eq. (7.33'))

$$(d\rho)^2 = M_{jk} dq_j dq_k \quad (8.86)$$

so that the kinetic energy has the form

$$T = \frac{1}{2} \left( \frac{d\rho}{dt} \right)^2, \quad (8.87)$$

or equivalently

$$dt = \frac{d\rho}{\sqrt{2T}}. \quad (8.88)$$

Equation (8.88) enables us to change the variable in the abbreviated action integral from  $t$  to  $\rho$ , and the principle of least action becomes

$$\Delta \int_{t_1}^{t_2} T dt = 0 = \Delta \int_{\rho_1}^{\rho_2} \sqrt{T/2} d\rho,$$

or, finally

$$\Delta \int_{\rho_1}^{\rho_2} \sqrt{H - V(q)} d\rho = 0. \quad (8.89)$$

Equation (8.89) is often called *Jacobi's form of the least action principle*. It now refers to the path of the system point in a special curvilinear configuration space characterized by a metric tensor with elements  $M_{jk}$ . The system point traverses the path in this configuration space with a speed given by  $\sqrt{2T}$ . If there are no forces acting on the body,  $T$  is constant, and Jacobi's principle says the system point travels along the shortest path length in the configuration space. Equivalently stated, the motion of the system is then such that the system point travels along the geodesics of the configuration space.

Note that the Jacobi form of the principle of least action is concerned with the *path* of the system point rather than with its motion in *time*. Equation (8.89) is a statement about the element of path length  $d\rho$ ; the time nowhere appears, since  $H$  is a constant and  $V$  depends upon  $q_i$  only. Indeed, it is possible to use the Jacobi form of the principle to furnish the differential equations for the path, by a procedure somewhat akin to that leading to Lagrange's equations. In the form of Fermat's principle, the Jacobi version of the principle of least action finds many fruitful applications in geometrical optics and in electron optics. To go into any detail here would lead us too far afield.

A host of other similar, variational principles for classical mechanics can be derived in bewildering variety. To give one example out of many, the principle of least action leads immediately to *Hertz's principle of least curvature*, which states that a particle not under the influence of external forces travels along the

path of least curvature. By Jacobi's principle such a path must be a geodesic, and the geometrical property of minimum curvature is one of the well-known characteristics of a geodesic. It has been pointed out that variational principles in themselves contain no new physical content, and they rarely simplify the practical solution of a given mechanical problem. Their value lies chiefly as starting points for new formulations of the theoretical structure of classical mechanics. For this purpose, Hamilton's principle is especially fruitful, and to a lesser extent, so also is the principle of least action.

## DERIVATIONS

1. (a) Reverse the Legendre transformation to derive the properties of  $L(q_i, \dot{q}_i, t)$  from  $H(q_i, p_i, t)$ , treating the  $\dot{q}_i$  as independent quantities, and show that it leads to the Lagrangian equations of motion.
- (b) By the same procedure find the equations of motion in terms of the function

$$L'(p, \dot{p}, t) = -\dot{p}_i q_i - H(q, p, t).$$

2. It has been previously noted that the total time derivative of a function of  $q_i$  and  $t$  can be added to the Lagrangian without changing the equations of motion. What does such an addition do to the canonical momenta and the Hamiltonian? Show that the equations of motion in terms of the new Hamiltonian reduce to the original Hamilton's equations of motion.
3. A Hamiltonian-like formulation can be set up in which  $\dot{q}_i$  and  $\dot{p}_i$  are the independent variables with a "Hamiltonian"  $G(\dot{q}_i, \dot{p}_i, t)$ . [Here  $p_i$  is defined in terms of  $q_i, \dot{q}_i$  in the usual manner] Starting from the Lagrangian formulation, show in detail how to construct  $G(\dot{p}_i, \dot{p}_i, t)$ , and derive the corresponding "Hamilton's equation of motion."
4. Show that if  $\lambda_i$  are the eigenvalues of a square matrix, then if the reciprocal matrix exists it has the eigenvalues  $\lambda_i^{-1}$ .
5. Verify that the matrix  $J$  has the properties given in Eqs. (8.38c) and (8.38e) and that its determinant has the value +1.
6. Show that Hamilton's principle can be written as

$$\delta \int_1^2 [2H(\eta, t) + \eta J \dot{\eta}] dt = 0.$$

7. Verify that both Hamiltonians, Eq. (8.45) and Eq. (8.47), lead to the same motion as described by Eq. (8.44).
8. Show that the modified Hamilton's principle, in the form of Eq. (8.71), leads to Hamilton's equations of motion.
9. If the canonical variables are not all independent, but are connected by auxiliary conditions of the form

$$\psi_k(q_i, p_i, t) = 0,$$

show that the canonical equations of motion can be written

$$\frac{\partial H}{\partial p_i} + \sum_k \lambda_k \frac{\partial \psi_k}{\partial p_i} = \dot{q}_i, \quad \frac{\partial H}{\partial q_i} + \sum_k \lambda_k \frac{\partial \psi_k}{\partial q_i} = -\dot{p}_i,$$

where the  $\lambda_k$  are the undetermined Lagrange multipliers. The formulation of the Hamiltonian equations in which  $t$  is a canonical variable is a case in point, since a relation exists between  $p_{n+1}$  and the other canonical variables:

$$H(q_1, \dots, q_{n+1}; p_1, \dots, p_n) + p_{n+1} = 0.$$

Show that as a result of these circumstances the  $2n + 2$  Hamilton's equations of this formulation can be reduced to the  $2n$  ordinary Hamilton's equations plus Eq. (8.41) and the relation

$$\lambda = \frac{dt}{d\theta}.$$

Note that while these results are reminiscent of the relativistic covariant Hamiltonian formulation, they have been arrived at entirely within the framework of nonrelativistic mechanics.

- 10.** Assume that the Lagrangian is a polynomial in  $\dot{q}$  of no higher order than quadratic. Convert the  $2n$  equations (8.2) and (8.14)

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \quad \dot{p}_i = \frac{\partial L}{\partial q_i},$$

into  $2n$  equations for  $\dot{q}_i$  and  $\dot{p}_i$  in terms of  $q$  and  $p$ , using the matrix form of the Lagrangian. Show that these are the same equations as would be obtained from Hamilton's equations of motion.

## EXERCISES

- 11.** A particle is confined to a one-dimensional box. The ends of the box move slowly towards the middle. By slowly we mean the speed of the ends is small when compared to the speed of the particle. Solve the following using Lagrangian formulation and then using the Hamiltonian.
- (a) if the momentum of the particle is  $p_0$  when the walls are a distance  $x_0$  apart, find the momentum of the particle at any later time assuming the collisions with the wall are perfectly elastic. Also assume the motion is nonrelativistic at all times.
  - (b) When the walls are a distance  $x$  apart, what average external force must be applied to each wall in order to move it at a constant speed?
- 12.** Write the problem of central force motion of two mass points in Hamiltonian formulation, eliminating the cyclic variables, and reducing the problem to quadratures.
- 13.** Formulate the double-pendulum problem illustrated by Fig. 1.4, in terms of the Hamiltonian and Hamilton's equations of motion. It is suggested that you find the Hamiltonian both directly from  $L$  and by Eq. (8.27).

14. The Lagrangian for a system can be written as

$$L = a\dot{x}^2 + b\frac{\dot{y}}{x} + c\dot{x}\dot{y} + f\dot{y}^2\dot{x}\dot{z} + g\dot{y} - k\sqrt{x^2 + y^2},$$

where  $a, b, c, f, g$ , and  $k$  are constants. What is the Hamiltonian? What quantities are conserved?

15. A dynamical system has the Lagrangian

$$L = \dot{q}_1^2 + \frac{\dot{q}_2^2}{a + bq_1^2} + k_1 q_1^2 + k_2 \dot{q}_1 \dot{q}_2,$$

where  $a, b, k_1$ , and  $k_2$  are constants. Find the equations of motion in the Hamiltonian formulation.

16. A Hamiltonian of one degree of freedom has the form

$$H = \frac{p^2}{2\alpha} - bqpe^{-\alpha t} + \frac{ba}{2}q^2e^{-\alpha t}(\alpha + be^{-\alpha t}) + \frac{kq^2}{2},$$

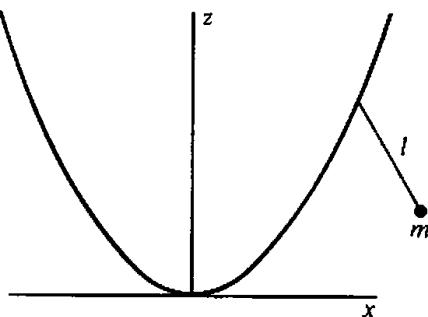
where  $a, b, \alpha$ , and  $k$  are constants.

- (a) Find a Lagrangian corresponding to this Hamiltonian.
- (b) Find an equivalent Lagrangian that is not explicitly dependent on time.
- (c) What is the Hamiltonian corresponding to this second Lagrangian, and what is the relationship between the two Hamiltonians?

17. Find the Hamiltonian for the system described in Exercise 19 of Chapter 5 and obtain Hamilton's equations of motion for the system. Use both the direct and the matrix approach in finding the Hamiltonian.

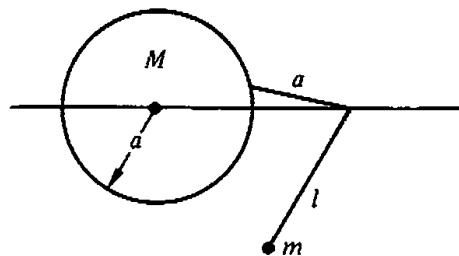
18. Repeat the preceding exercise except this time allow the *pendulum* to move in three dimensions, that is, a spring-loaded spherical pendulum. Either the direct or the matrix approach may be used.

19. The point of suspension of a simple pendulum of length  $l$  and mass  $m$  is constrained to move on a parabola  $z = ax^2$  in the vertical plane. Derive a Hamiltonian governing the motion of the pendulum and its point of suspension. Obtain the Hamilton's equations of motion.



20. Obtain Hamilton's equations of motion for a plane pendulum of length  $l$  with mass point  $m$  whose radius of suspension rotates uniformly on the circumference of a vertical circle of radius  $a$ . Describe physically the nature of the canonical momentum and the Hamiltonian.

21. (a) The point of suspension of a plane simple pendulum of mass  $m$  and length  $l$  is constrained to move along a horizontal track and is connected to a point on the circumference of a uniform flywheel of mass  $M$  and radius  $a$  through a massless connecting rod also of length  $a$ , as shown in the figure. The flywheel rotates about a center fixed on the track. Find a Hamiltonian for the combined system and determine Hamilton's equations of motion.



- (b) Suppose the point of suspension were moved along the track according to some function of time  $x = f(t)$ , where  $x$  reverses at  $x = \pm 2a$  (relative to the center of the fly wheel). Again, find a Hamiltonian and Hamilton's equations of motion.
22. For the arrangement described in Exercise 21 of Chapter 2, find the Hamiltonian of the system, first in terms of coordinates in the laboratory system and then in terms of coordinates in the rotating systems. What are the conservation properties of the Hamiltonians, and how are they related to the energy of the system?
23. (a) A particle of mass  $m$  and electric charge  $e$  moves in a plane under the influence of a central force potential  $V(r)$  and a constant uniform magnetic field  $\mathbf{B}$ , perpendicular to the plane, generated by a static vector potential

$$\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}.$$

Find the Hamiltonian using coordinates in the observer's inertial system.

- (b) Repeat part (a) using coordinates rotating relative to the previous coordinate system about an axis perpendicular to the plane with an angular rate of rotation:

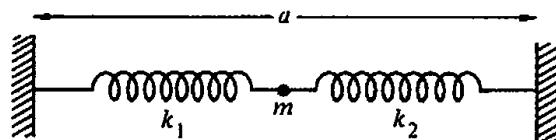
$$\omega = -\frac{eB}{m}$$

24. A uniform cylinder of radius  $a$  and density  $\rho$  is mounted so as to rotate freely around a vertical axis. On the outside of the cylinder is a rigidly fixed uniform spiral or helical track along which a mass point  $m$  can slide without friction. Suppose a particle starts



at rest at the top of the cylinder and slides down under the influence of gravity. Using any set of coordinates, arrive at a Hamiltonian for the combined system of particle and cylinder, and solve for the motion of the system.

25. Suppose that in the previous exercise the cylinder is constrained to rotate uniformly with angular frequency  $\omega$ . Set up the Hamiltonian for the particle in an inertial system of coordinates and also in a system fixed in the rotating cylinder. Identify the physical nature of the Hamiltonian in each case and indicate whether or not the Hamiltonians are conserved.
26. A particle of mass  $m$  can move in one dimension under the influence of two springs connected to fixed points a distance  $a$  apart (see figure). The springs obey Hooke's law and have zero unstretched lengths and force constants  $k_1$  and  $k_2$ , respectively.



- (a) Using the position of the particle from one fixed point as the generalized coordinate, find the Lagrangian and the corresponding Hamiltonian. Is the energy conserved? Is the Hamiltonian conserved?
- (b) Introduce a new coordinate  $Q$  defined by

$$Q = q - b \sin \omega t, \quad b = \frac{k_2 a}{k_1 + k_2}.$$

What is the Lagrangian in terms of  $Q$ ? What is the corresponding Hamiltonian? Is the energy conserved? Is the Hamiltonian conserved?

27. (a) The Lagrangian for a system of one degree of freedom can be written as

$$L = \frac{m}{2} (\dot{q}^2 \sin^2 \omega t + \dot{q} q \omega \sin 2\omega t + q^2 \omega^2).$$

What is the corresponding Hamiltonian? Is it conserved?

- (b) Introduce a new coordinate defined by

$$Q = q \sin \omega t.$$

Find the Lagrangian in terms of the new coordinate and the corresponding Hamiltonian. Is  $H$  conserved?

28. Consider a system of particles interacting with each other through potentials depending only on the scalar distances between them and acted upon by conservative central forces from a fixed point. Obtain the Hamiltonian of the particle with respect to a set of axes, with origin at the center of force, which is rotating around some axis in an inertial system with angular velocity  $\omega$ . What is the physical significance of the Hamiltonian in this case? Is it a constant of the motion?

29. Obtain the Hamiltonian of a heavy symmetrical top with one point fixed, and from it the Hamilton's equations of motion. Relate these to the equations of motion discussed in Section 5.7 and, in particular, show how the solution may be reduced to quadratures. Also use the Routhian procedure to eliminate the cyclic coordinates.
30. In Exercise 16 of Chapter 1, there is given the velocity-dependent potential assumed in Weber's electrodynamics. What is the Hamiltonian for a single particle moving under the influence of such a potential?
31. Treat the nutation of a "fast" top as an example of small oscillations about steady motion, here precession at constant  $\theta$ . Find the frequency of nutation.
32. A symmetrical top is mounted so that it pivots about its center of mass. The pivot in turn is fixed a distance  $r$  from the center of a horizontal disk free to rotate about a vertical axis. The top is started with an initial rotation about its figure axis, which is initially at an angle  $\theta_0$  to the vertical. Analyze the possible nutation of the top as a case of small oscillations about steady motion.
33. Two mass points,  $m_1$  and  $m_2$ , are connected by a string that acts as a Hooke's-law spring with force constant  $k$ . One particle is free to move without friction on a smooth horizontal plane surface, the other hangs vertically down from the string through a hole in the surface. Find the condition for steady motion in which the mass point on the plane rotates uniformly at constant distance from the hole. Investigate the small oscillations in the radial distance from the hole, and in the vertical height of the second particle.
34. A possible covariant Lagrangian for a system of one particle interacting with a field is

$$\Lambda = \frac{1}{2}mu_\lambda u_\lambda + D_{\lambda\nu}(x_\mu)m_{\lambda\nu},$$

where  $D_{\lambda\nu}(x_\mu)$  is an antisymmetric field tensor and  $m_{\lambda\nu}$  is the antisymmetric angular momentum tensor,

$$m_{\lambda\nu} = m(x_\lambda u_\nu - x_\nu u_\lambda).$$

What are the canonical momenta? What is the corresponding covariant Hamiltonian?

35. Consider a Lagrangian of the form

$$L = \frac{1}{2}m(\dot{x}^2 - \omega^2 x^2)e^{\gamma t},$$

where the particle of mass  $m$  moves in one direction. Assume all constants are positive.

- (a) Find the equations of motion.
- (b) Interpret the equations by giving a physical interpretation of the forces acting on the particle.
- (c) Find the canonical momentum and construct the Hamiltonian. Is this Hamiltonian a constant of the motion?
- (d) If initially  $x(0) = 0$  and  $dx/dt = 0$ , what is  $x(t)$  as  $t$  approaches large values?

# CHAPTER

# 9

# Canonical Transformations

When applied in a straightforward manner, the Hamiltonian formulation usually does not materially decrease the difficulty of solving any given problem in mechanics. We wind up with practically the same differential equations to be solved as are provided by the Lagrangian procedure. The advantages of the Hamiltonian formulation lie not in its use as a calculational tool, but rather in the deeper insight it affords into the formal structure of mechanics. The equal status accorded to coordinates and momenta as independent variables encourages a greater freedom in selecting the physical quantities to be designated as "coordinates" and "momenta." As a result we are led to newer, more abstract ways of presenting the physical content of mechanics. While often of considerable help in practical applications to mechanical problems, these more abstract formulations are primarily of interest to us today because of their essential role in constructing the more modern theories of matter. Thus, one or another of these formulations of classical mechanics serves as a point of departure for both statistical mechanics and quantum theory. It is to such formulations, arising as outgrowths of the Hamiltonian procedure, that this and the next chapter are devoted.

## 9.1 ■ THE EQUATIONS OF CANONICAL TRANSFORMATION

There is one type of problem for which the solution of the Hamilton's equations is trivial. Consider a situation in which the Hamiltonian is a constant of the motion, and where *all* coordinates  $q_i$  are cyclic. Under these conditions, the conjugate momenta  $p_i$  are all constant:

$$p_i = \alpha_i,$$

and since the Hamiltonian cannot be an explicit function of either the time or the cyclic coordinates, it may be written as

$$H = H(\alpha_1, \dots, \alpha_n).$$

Consequently, the Hamilton's equations for  $\dot{q}_i$  are simply

$$\dot{q}_i = \frac{\partial H}{\partial \alpha_i} = \omega_i, \quad (9.1)$$

where the  $\omega_i$ 's are functions of the  $\alpha_i$ 's only and therefore are also constant in time. Equations (9.1) have the immediate solutions

$$q_i = \omega_i t + \beta_i, \quad (9.2)$$

where the  $\beta_i$ 's are constants of integration, determined by the initial conditions.

It would seem that the solution to this type of problem, easy as it is, can only be of academic interest, for it rarely happens that all the generalized coordinates are cyclic. But a given system can be described by more than one set of generalized coordinates. Thus, to discuss motion of a particle in a plane, we may use as generalized coordinates either the Cartesian coordinates

$$q_1 = x, \quad q_2 = y,$$

or the plane polar coordinates

$$q_1 = r, \quad q_2 = \theta.$$

Both choices are equally valid, but one of the other set may be more convenient for the problem under consideration. Note that for central forces neither  $x$  nor  $y$  is cyclic, while the second set does contain a cyclic coordinate in the angle  $\theta$ . The number of cyclic coordinates can thus depend upon the choice of generalized coordinates, and for each problem there may be one particular choice for which all coordinates are cyclic. If we can find this set, the remainder of the job is trivial. Since the obvious generalized coordinates suggested by the problem will not normally be cyclic, we must first derive a specific procedure for *transforming* from one set of variables to some other set that may be more suitable.

The transformations considered in the previous chapters have involved going from one set of coordinates  $q_i$  to a new set  $Q_i$ , by transformation equations of the form

$$Q_i = Q_i(q, t). \quad (9.3)$$

For example, the equations of an orthogonal transformation, or of the change from Cartesian to plane polar coordinates, have the general form of Eqs. (9.3). As has been previously noted in Derivation 10 of Chapter 1, such transformations are known as *point transformations*. But in the Hamiltonian formulation the momenta are also independent variables on the same level as the generalized coordinates. The concept of transformation of coordinates must therefore be widened to include the simultaneous transformation of the independent *coordinates* and *momenta*,  $q_i, p_i$ , to a new set  $Q_i, P_i$ , with (invertible) equations of transformation:

$$\begin{aligned} Q_i &= Q_i(q, p, t), \\ P_i &= P_i(q, p, t). \end{aligned} \quad (9.4)$$

Thus, the new coordinates will be defined not only in terms of the old coordinates but also in terms of the old momenta. Equations (9.3) may be said to define

a *point transformation of configuration space*; correspondingly Eqs. (9.4) define a *point transformation of phase space*.

In developing Hamiltonian mechanics, only those transformations can be of interest for which the new  $Q_i, P_i$  are canonical coordinates. This requirement will be satisfied provided there exists some function  $K(Q_i, P_i, t)$  such that the equations of motion in the new set are in the Hamiltonian form

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i}. \quad (9.5)$$

The function  $K$  plays the role of the Hamiltonian in the new coordinate set.\* It is important for future considerations that the transformations considered be problem-independent. That is to say,  $(Q_i, P_i)$  must be canonical coordinates not only for some specific mechanical systems, but for all systems of the same number of degrees of freedom. Equations (9.5) must be the form of the equations of motion in the new coordinates and momenta no matter what the particular initial form of  $H$ . We may indeed be incited to develop a particular transformation from  $(q, p)$  to  $(Q, P)$  to handle, say, a plane harmonic oscillator. But the same transformation must then also lead to Hamilton's equations of motion when applied, for example, to the two-dimensional Kepler problem.

As was seen in Section 8.5, if  $Q_i$  and  $P_i$  are to be canonical coordinates, they must satisfy a modified Hamilton's principle that can be put in the form

$$\delta \int_{t_1}^{t_2} (P_i \dot{Q}_i - K(Q_i, P_i, t)) dt = 0, \quad (9.6)$$

(where summation over the repeated index  $i$  is implied). At the same time the old canonical coordinates of course satisfy a similar principle:

$$\delta \int_{t_1}^{t_2} (p_i \dot{q}_i - H(q, p, t)) dt = 0. \quad (9.7)$$

The simultaneous validity of Eqs. (9.6) and (9.7) does not mean of course that the integrands in both expressions are equal. Since the general form of the modified Hamilton's principle has zero variation at the end points, both statements will be satisfied if the integrands are connected by a relation of the form

$$\lambda(p_i \dot{q}_i - H) = P_i \dot{Q}_i - K + \frac{dF}{dt}. \quad (9.8)$$

Here  $F$  is any function of the phase space coordinates with continuous second derivatives, and  $\lambda$  is a constant independent of the canonical coordinates and the time. The multiplicative constant  $\lambda$  is related to a particularly simple type of transformation of canonical coordinates known as a *scale transformation*.

\*It has been remarked in a jocular vein that if  $H$  stands for the Hamiltonian,  $K$  must stand for the Kamiltonian! Of course,  $K$  is every bit as much a Hamiltonian as  $H$ , but the designation is occasionally a convenient substitute for the longer term "transformed Hamiltonian".

Suppose we change the size of the units used to measure the coordinates and momenta so that in effect we transform them to a set  $(Q', P')$  defined by

$$Q'_i = \mu q_i, \quad P'_i = \nu p_i. \quad (9.9)$$

Then it is clear Hamilton's equations in the form of Eqs. (9.5) will be satisfied for a transformed Hamiltonian  $K'(Q', P') = \mu\nu H(q, p)$ . The integrands of the corresponding modified Hamilton's principles are, also obviously, related as

$$\mu\nu(p_i \dot{q}_i - H) = P'_i \dot{Q}'_i - K', \quad (9.10)$$

which is of the form of Eq. (9.8) with  $\lambda = \mu\nu$ . With the aid of suitable scale transformation, it will always be possible to confine our attention to transformations of canonical coordinates for which  $\lambda = 1$ . Thus, if we have a transformation of canonical coordinates  $(q, p) \rightarrow (Q', P')$  for some  $\lambda \neq 1$ , then we can always find an intermediate set of canonical coordinates  $(Q, P)$  related to  $(Q', P')$  by a simple scale transformation of the form (9.9) such that  $\mu\nu$  also has the same value  $\lambda$ . The transformation between the two sets of canonical coordinates  $(q, p)$  and  $(Q, P)$  will satisfy Eq. (9.8), but now with  $\lambda = 1$ :

$$p_i \dot{q}_i - H = P_i \dot{Q}_i - K + \frac{dF}{dt}. \quad (9.11)$$

Since the scale transformation is basically trivial, the significant transformations to be examined are those for which Eq. (9.11) holds.

A transformation of canonical coordinates for which  $\lambda \neq 1$  will be called an *extended canonical transformation*. Where  $\lambda = 1$ , and Eq. (9.11) holds, we will speak simply of a *canonical transformation*. The conclusion of the previous paragraph may then be stated as saying that any extended canonical transformation can be made up of a canonical transformation followed by a scale transformation. Except where otherwise stated, all future considerations of transformations between canonical coordinates will involve only canonical transformations. It is also convenient to give a specific name to canonical transformations for which the equations of transformation Eqs. (9.4) do not contain the time explicitly; they will be called *restricted canonical transformations*.

The last term on the right in Eq. (9.11) contributes to the variation of the action integral only at the end points and will therefore vanish if  $F$  is a function of  $(q, p, t)$  or  $(Q, P, t)$  or any mixture of the phase space coordinates since these have zero variation at the end points. Further, through the equations of transformation, Eqs. (9.4) and their inverses  $F$  can be expressed partly in terms of the old set of variables and partly of the new. Indeed,  $F$  is useful for specifying the exact form of the canonical transformation only when half of the variables (beside the time) are from the old set and half are from the new. It then acts, as it were, as a bridge between the two sets of canonical variables and is called the *generating function* of the transformation.

To show how the generating function specifies the equations of transformation, suppose  $F$  were given as a function of the old and new generalized space

coordinates:

$$F = F_1(q, Q, t). \quad (9.12)$$

Equation (9.11) then takes the form

$$\begin{aligned} p_i \dot{q}_i - H &= P_i \dot{Q}_i - K + \frac{dF_1}{dt} \\ &= P_i \dot{Q}_i - K + \frac{\partial F_1}{\partial t} + \frac{\partial F_1}{\partial q_i} \dot{q}_i + \frac{\partial F_1}{\partial Q_i} \dot{Q}_i. \end{aligned} \quad (9.13)$$

Since the old and the new coordinates,  $q_i$  and  $Q_i$ , are separately independent, Eq. (9.13) can hold identically only if the coefficients of  $\dot{q}_i$  and  $\dot{Q}_i$  each vanish:

$$p_i = \frac{\partial F_1}{\partial q_i}, \quad (9.14a)$$

$$P_i = -\frac{\partial F_1}{\partial Q_i}, \quad (9.14b)$$

leaving finally

$$K = H + \frac{\partial F_1}{\partial t}. \quad (9.14c)$$

Equations (9.14a) are  $n$  relations defining the  $p_i$  as functions of  $q_j$ ,  $Q_j$ , and  $t$ . Assuming they can be inverted, they could then be solved for the  $n$   $Q_i$ 's in terms of  $q_j$ ,  $p_j$ , and  $t$ , thus yielding the first half of the transformation equations (9.4). Once the relations between the  $Q_i$ 's and the old canonical variables  $(q, p)$  have been established, they can be substituted into Eqs. (9.14b) so that they give the  $n$   $P_i$ 's as functions of  $q_j$ ,  $p_j$ , and  $t$ , that is, the second half of the transformation equations (9.4). To complete the story, Eq. (9.14c) provides the connection between the new Hamiltonian,  $K$ , and the old one,  $H$ . We must be careful to read Eq. (9.14c) properly. First  $q$  and  $p$  in  $H$  are expressed as functions of  $Q$  and  $P$  through the inverses of Eqs. (9.4). Then the  $q_i$  in  $\partial F_1/\partial t$  are expressed in terms of  $Q$ ,  $P$  in a similar manner and the two functions are added to yield  $K(Q, P, t)$ .

The procedure described shows how, starting from a given generating function  $F_1$ , the equations of the canonical transformation can be obtained. We can usually reverse the process: Given the equations of transformation (9.4), an appropriate generating function  $F_1$  may be derived. Equations (9.4) are first inverted to express  $p_i$  and  $P_i$  as functions of  $q$ ,  $Q$ , and  $t$ . Equations (9.14a, b) then constitute a coupled set of partial differential equations than can be integrated, in principle, to find  $F_1$  providing the transformation is indeed canonical. Thus,  $F_1$  is always uncertain to within an additive arbitrary function of  $t$  alone (which doesn't affect the equations of transformation), and there may at times be other ambiguities.

It sometimes happens that it is not suitable to describe the canonical transformation by a generating function of the type  $F_1(q, Q, t)$ . For example, the transformation may be such that  $p_i$  cannot be written as functions of  $q$ ,  $Q$ , and  $t$ , but

rather will be functions of  $q$ ,  $P$ , and  $t$ . We would then seek a generating function that is a function of the old coordinates  $q$  and the new momenta  $P$ . Clearly Eq. (9.13) must then be replaced by an equivalent relation involving  $\dot{P}_i$  rather than  $\dot{Q}_i$ . This can be accomplished by writing  $F$  in Eq. (9.11) as

$$F = F_2(q, P, t) - Q_i P_i. \quad (9.15)$$

Substituting this  $F$  in Eq. (9.11) leads to

$$p_i \dot{q}_i - H = -Q_i \dot{P}_i - K + \frac{d}{dt} F_2(q, P, t). \quad (9.16)$$

Again, the total derivative of  $F_2$  is expanded and the coefficients of  $\dot{q}_i$  and  $\dot{P}_i$  collected, leading to the equations

$$p_i = \frac{\partial F_2}{\partial q_i}, \quad (9.17a)$$

$$Q_i = \frac{\partial F_2}{\partial P_i}, \quad (9.17b)$$

with

$$K = H + \frac{\partial F_2}{\partial t}. \quad (9.17c)$$

As before, Eqs. (9.17a) are to be solved for  $P_i$  as functions of  $q_j$ ,  $p_j$ , and  $t$  to correspond to the second half of the transformation equations (9.4). The remaining half of the transformation equations is then provided by Eqs. (9.17b).

The corresponding procedures for the remaining two basic types of generating functions are obvious, and the general results are displayed in Table 9.1.

It is tempting to look upon the four basic types of generating functions as being related to each other through Legendre transformations. For example, the

**TABLE 9.1** Properties of the Four Basic Canonical Transformations

Generating Function	Generating Function Derivatives	Trivial Special Case
$F = F_1(q, Q, t)$	$p_i = \frac{\partial F_1}{\partial q_i}$ $P_i = -\frac{\partial F_1}{\partial Q_i}$	$F_1 = q_i Q_i$ , $Q_i = p_i$ , $P_i = -q_i$
$F = F_2(q, P, t) - Q_i P_i$	$p_i = \frac{\partial F_2}{\partial q_i}$ $Q_i = \frac{\partial F_2}{\partial P_i}$	$F_2 = q_i P_i$ , $Q_i = q_i$ , $P_i = p_i$
$F = F_3(p, Q, t) + q_i p_i$	$q_i = -\frac{\partial F_3}{\partial p_i}$ $P_i = -\frac{\partial F_3}{\partial Q_i}$	$F_3 = p_i Q_i$ , $Q_i = -q_i$ , $P_i = -p_i$
$F = F_4(p, P, t) + q_i p_i - Q_i P_i$	$q_i = -\frac{\partial F_4}{\partial p_i}$ $Q_i = \frac{\partial F_4}{\partial P_i}$	$F_4 = p_i P_i$ , $Q_i = p_i$ , $P_i = -q_i$

transition from  $F_1$  to  $F_2$  is equivalent to going from the variables  $q, Q$  to  $q, P$  with the relation

$$-P_i = \frac{\partial F_1}{\partial Q_i}. \quad (9.18)$$

This is just the form required for a Legendre transformation of the basis variables, as described in Section 8.1, and in analogy to Eq. (8.5) we would set

$$F_2(q, P, t) = F_1(q, Q, t) + P_i Q_i, \quad (9.19)$$

which is equivalent to Eq. (9.15) combined with Eq. (9.12). All the other defining equations for the generating functions can similarly be looked on, in combination with Eq. (9.12) as Legendre transformations from  $F_1$ , with the last entry in Table 9.1 describing a double Legendre transformation. The only drawback to this picture is that it might erroneously lead us to believe that any given canonical transformation can be expressed in terms of the four basic types of Legendre transformations listed in Table 9.1. This is not always possible. Some transformations are just not suitable for description in terms of these or other elementary forms of generating functions, as has been noted above and as will be illustrated in the next section with specific examples. If we try to apply the Legendre transformation process, we are then led to generating functions that are identically zero or are indeterminate. For this reason, we have preferred to define each type of generating function relative to  $F$ , which is some unspecified function of  $2n$  independent coordinates and momenta.

Finally, note that a suitable generating function doesn't have to conform to one of the four basic types for *all* the degrees of freedom of the system. It is possible, and for some canonical transformations necessary, to use a generating function that is a mixture of the four types. To take a simple example, it may be desirable for a particular canonical transformation with two degrees of freedom to be defined by a generating function of the form

$$F'(q_1, p_2, P_1, Q_2, t). \quad (9.20)$$

This generating function would be related to  $F$  in Eq. (9.11) by the equation

$$F = F'(q_1, p_2, P_1, Q_2, t) - Q_1 P_1 + q_2 p_2, \quad (9.21)$$

and the equations of transformation would be obtained from the relations

$$\begin{aligned} P_1 &= \frac{\partial F'}{\partial q_1}, & Q_1 &= \frac{\partial F'}{\partial P_1}, \\ q_2 &= -\frac{\partial F'}{\partial p_2}, & P_2 &= -\frac{\partial F'}{\partial Q_2}, \end{aligned} \quad (9.22)$$

with

$$K = H + \frac{\partial F'}{\partial t}. \quad (9.23)$$

Specific illustrations are given in the next section and in the exercises.

## 9.2 ■ EXAMPLES OF CANONICAL TRANSFORMATIONS

The nature of canonical transformations and the role played by the generating function can best be illustrated by some simple yet important examples. Let us consider, first, a generating function of the second type with the particular form

$$F_2 = q_i P_i \quad (9.24)$$

found in column 3 of Table 9.1. From Eqs. (9.17), the transformation equations are

$$\begin{aligned} p_i &= \frac{\partial F_2}{\partial q_i} = P_i, \\ Q_i &= \frac{\partial F_2}{\partial P_i} = q_i, \\ K &= H. \end{aligned} \quad (9.25)$$

The new and old coordinates are the same; hence  $F_2$  merely generates the *identity transformation* (cf. Table 9.1). We also note, referring to Table 9.1, that the particular generating function  $F_3 = p_i Q_i$  generates an identity transformation with negative signs; that is,  $Q_i = -q_i$ ,  $P_i = -p_i$ .

A more general type of transformation is described by the generating function

$$F_2 = f_i(q_1, \dots, q_n; t)P_i, \quad (9.26)$$

where the  $f_i$  may be any desired set of independent functions. By Eqs. (9.17b), the new coordinates  $Q_i$  are given by

$$Q_i = \frac{\partial F_2}{\partial P_i} = f_i(q_1, \dots, q_n; t) \quad (9.27)$$

Thus, with this generating function the new coordinates depend only upon the old coordinates and the time and do not involve the old momenta. Such a transformation is therefore an example of the class of point transformations defined by Eqs. (9.3). In order to define a point transformation, the functions  $f_i$  must be independent and invertible, so that the  $q_j$  can be expressed in terms of the  $Q_i$ . Since the  $f_i$  are otherwise completely arbitrary, we may conclude that *all point transformations are canonical*. Equation (9.17c) furnishes the new Hamiltonian in terms of the old and of the time derivatives of the  $f_i$  functions.

Note that  $F_2$  as given by Eq. (9.26) is not the only generating function leading to the point transformation specified by the  $f_i$ . Clearly the same point transformation is implicit in the more general form

$$F_2 = f_i(q_1, \dots, q_n; t)P_i + g(q_1, \dots, q_n; t), \quad (9.28)$$

where  $g(q, t)$  is any (differentiable) function of the old coordinates and the time. Equations (9.27), the transformation equations for the coordinates, remain unaltered for this generating function. But the transformation equations of the momenta differ for the two forms. From Eqs. (9.17a), we have

$$p_j = \frac{\partial F_2}{\partial q_j} = \frac{\partial f_i}{\partial q_j} P_i + \frac{\partial g}{\partial q_j}, \quad (9.29)$$

using the form of  $F_2$  given by Eq. (9.28). These equations may be inverted to give  $P$  as a function of  $(q, p)$ , most easily by writing them in matrix notation:

$$\mathbf{p} = \frac{\partial \mathbf{f}}{\partial \mathbf{q}} \mathbf{P} + \frac{\partial g}{\partial \mathbf{q}}. \quad (9.29')$$

Here  $\mathbf{p}$ ,  $\mathbf{P}$ , and  $\partial g / \partial \mathbf{q}$  are  $n$ -elements of single-column matrices, and  $\partial \mathbf{f} / \partial \mathbf{q}$  is a square matrix whose  $ij$ th element is  $\partial f_i / \partial q_j$ . In two dimensions, Eq. (9.29') can be written as

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial q_1} & \frac{\partial f_1}{\partial q_2} \\ \frac{\partial f_2}{\partial q_1} & \frac{\partial f_2}{\partial q_2} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} + \begin{bmatrix} \frac{\partial g}{\partial q_1} \\ \frac{\partial g}{\partial q_2} \end{bmatrix}.$$

It follows that  $\mathbf{P}$  is a linear function of  $\mathbf{p}$  given by

$$\mathbf{P} = \left[ \frac{\partial \mathbf{f}}{\partial \mathbf{q}} \right]^{-1} \left[ \mathbf{p} - \frac{\partial g}{\partial \mathbf{q}} \right]. \quad (9.30)$$

In two dimensions, (9.30) becomes

$$\begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial q_1} & \frac{\partial f_1}{\partial q_2} \\ \frac{\partial f_2}{\partial q_1} & \frac{\partial f_2}{\partial q_2} \end{bmatrix}^{-1} \left[ \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} - \begin{bmatrix} \frac{\partial g}{\partial q_1} \\ \frac{\partial g}{\partial q_2} \end{bmatrix} \right]. \quad (9.31)$$

Thus, the transformation equations (9.27) for  $Q$  are independent of  $g$  and depend only upon the  $f_i(q, t)$ , but the transformation equations (9.29) for  $P$  do depend upon the form of  $g$  and are in general functions of both the old coordinates and momenta. The generating function given by Eq. (9.26) is only a special case of Eq. (9.28) for which  $g = 0$ , with correspondingly specialized transformation equations for  $P$ .

An instructive transformation is provided by the generating function of the first kind,  $F_1(q, Q, t)$ , of the form

$$F_1 = q_k Q_k.$$

The corresponding transformation equations, from (9.14a, b) are

$$p_i = \frac{\partial F_i}{\partial q_i} = Q_i, \quad (9.32a)$$

$$P_i = -\frac{\partial F_i}{\partial Q_i} = -q_i. \quad (9.32b)$$

In effect, the transformation interchanges the momenta and the coordinates; the new coordinates are the old momenta and the new momenta are essentially the old coordinates. Table 9.1 shows that the particular generating function of type  $F_4 = p_i P_i$  produces the same transformation. These simple examples should emphasize the independent status of generalized coordinates and momenta. They are both needed to describe the motion of the system in the Hamiltonian formulation. The distinction between them is basically one of nomenclature. We can shift the names around with at most no more than a change in sign. There is no longer present in the theory any lingering remnant of the concept of  $q_i$  as a spatial coordinate and  $p_i$  as a mass times a velocity. Incidentally, we may see directly from Hamilton's equations,

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad \dot{q}_i = \frac{\partial H}{\partial p_i},$$

that this exchange transformation is canonical. If  $q_i$  is substituted for  $p_i$ , the equations remain in the canonical form only if  $-p_i$  is substituted for  $q_i$ .

A transformation that leaves some of the  $(q, p)$  pairs unchanged, and interchanges the rest (with a sign change), is obviously a canonical transformation of a "mixed" form. Thus, in a system of two degrees of freedom, the transformation

$$\begin{aligned} Q_1 &= q_1, & P_1 &= p_1, \\ Q_2 &= p_2, & P_2 &= -q_2, \end{aligned}$$

is generated by the function

$$F = q_1 P_1 + q_2 Q_2, \quad (9.33)$$

which is a mixture of the  $F_1$  and  $F_2$  types.

### 9.3 ■ THE HARMONIC OSCILLATOR

As a final example, let us consider a canonical transformation that can be used to solve the problem of the simple harmonic oscillator in one dimension. If the force

constant is  $k$ , the Hamiltonian for this problem in terms of the usual coordinates is

$$H = \frac{p^2}{2m} + \frac{kq^2}{2}. \quad (9.34a)$$

Designating the ratio  $k/m$  by  $\omega^2$ ,  $H$  can also be written as

$$H = \frac{1}{2m}(p^2 + m^2\omega^2q^2). \quad (9.34b)$$

This form of the Hamiltonian, as the sum of two squares, suggests a transformation in which  $H$  is cyclic in the new coordinate. If we could find a canonical transformation of the form

$$p = f(P) \cos Q, \quad (9.35a)$$

$$q = \frac{f(P)}{m\omega} \sin Q, \quad (9.35b)$$

then the Hamiltonian as a function of  $Q$  and  $P$  would be simply

$$K = H = \frac{f^2(P)}{2m}(\cos^2 Q + \sin^2 Q) = \frac{f^2(P)}{2m}, \quad (9.36)$$

so that  $Q$  is cyclic. The problem is to find the form of the yet unspecified function  $f(P)$  that makes the transformation canonical. If we use a generating function of the first kind given by

$$F_1 = \frac{m\omega q^2}{2} \cot Q, \quad (9.37)$$

Eqs. (9.14) then provide the equations of transformation,

$$p = \frac{\partial F_1}{\partial q} = m\omega q \cot Q, \quad (9.38a)$$

$$P = -\frac{\partial F_1}{\partial Q} = \frac{m\omega q^2}{2 \sin^2 Q}. \quad (9.38b)$$

Solving for  $q$  and  $p$ , we have\*

$$q = \sqrt{\frac{2P}{m\omega}} \sin Q, \quad (9.39a)$$

\*It can be argued that  $F_1$  does not unambiguously specify the canonical transformation, because in solving Eq. (9.38b) for  $q$  we could have taken the negative square root instead of the positive root as (implied) in Eqs. (9.39). However, the two canonical transformations thus derived from  $F_1$  differ only trivially; a shift in  $\alpha$  by  $\pi$  corresponds to going from one transformation to the other. Nonetheless, it should be kept in mind that the transformations derived from a generating function may at times be double-valued or even have local singularities.

$$p = \sqrt{2pm\omega} \cos Q, \quad (9.39b)$$

and comparison with Eq. (9.35a) evaluates  $f(P)$ :

$$f(P) = \sqrt{2m\omega P}. \quad (9.40)$$

It follows then that the Hamiltonian in the transformed variables is

$$H = \omega P. \quad (9.41)$$

Since the Hamiltonian is cyclic in  $Q$ , the conjugate momentum  $P$  is a constant. It is seen from Eq. (9.41) that  $P$  is in fact equal to the constant energy divided by  $\omega$ :

$$P = \frac{E}{\omega}.$$

The equation of motion for  $Q$  reduces to the simple form

$$\dot{Q} = \frac{\partial H}{\partial P} = \omega,$$

with the immediate solution

$$Q = \omega t + \alpha, \quad (9.42)$$

where  $\alpha$  is a constant of integration fixed by the initial conditions. From Eqs. (9.39), the solutions for  $q$  and  $p$  are

$$q = \sqrt{\frac{2E}{m\omega^2}} \sin(\omega t + \alpha), \quad (9.43a)$$

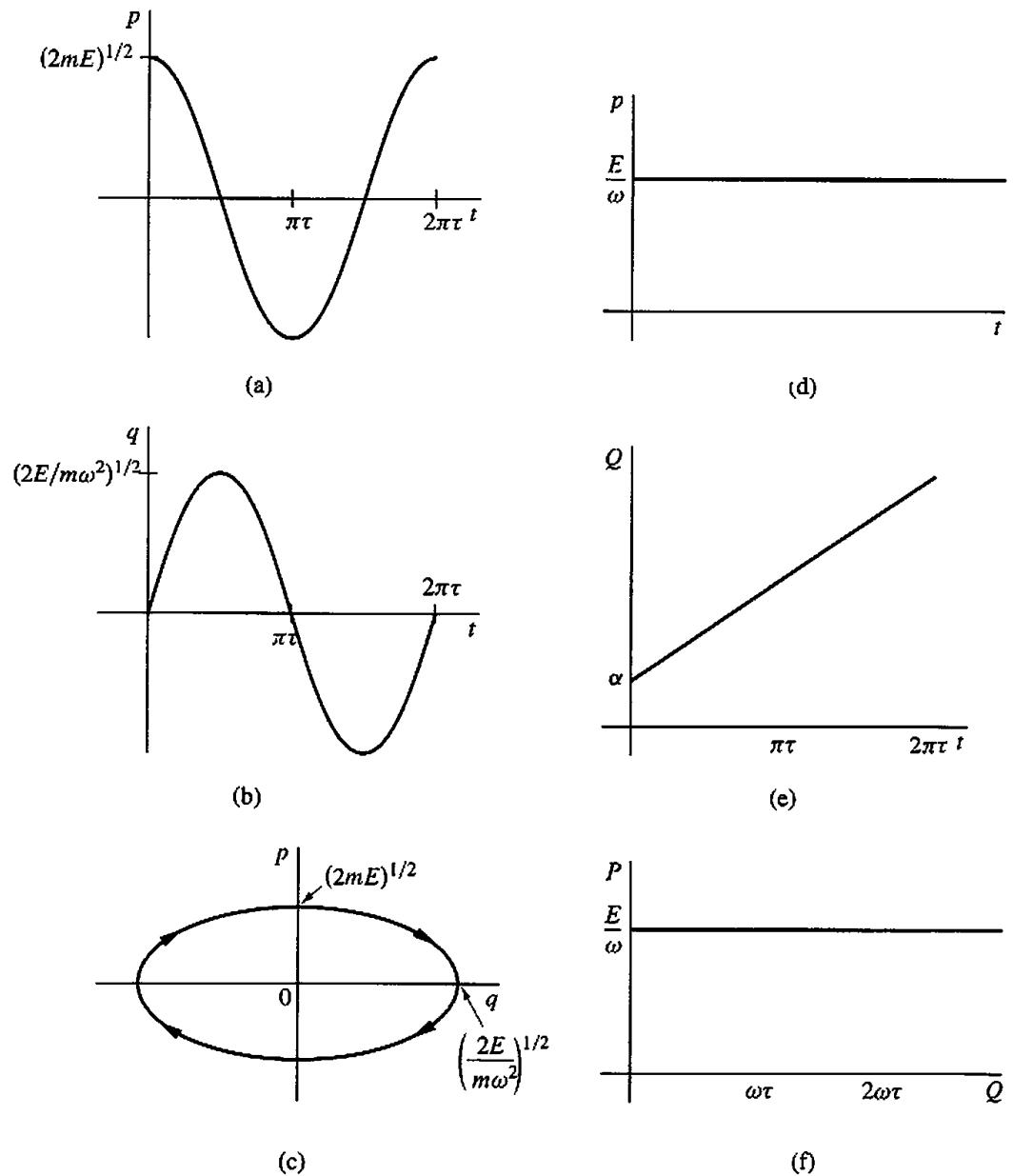
$$p = \sqrt{2mE} \cos(\omega t + \alpha). \quad (9.43b)$$

It is instructive to plot the time dependence of the old and new variables as is shown in Fig. 9.1. We see that  $q$  and  $p$  oscillate (Fig. 9.1a, b) whereas  $Q$  and  $P$  are linear plots (Fig. 9.1d, e). The figure also shows the phase space plots for  $p$  versus  $q$  (Fig. 9.1c) and for  $P$  versus  $Q$  (Fig. 9.1f). Fig. 9.1c is an ellipse with the following semimajor axes (for the  $q$  and  $p$  directions, respectively):

$$a = \sqrt{\frac{2E}{m\omega^2}} \quad \text{and} \quad b = \sqrt{2mE},$$

where  $m$  is the mass of the oscillator,  $\omega$  its frequency, and  $E$  the oscillator's energy. The area,  $A$ , of this ellipse in phase space is

$$A = \pi ab = \frac{2\pi E}{\omega}.$$



**FIGURE 9.1** The harmonic oscillator in two canonical coordinate systems. Drawings (a)–(c) show the  $q, p$  system and (d)–(f) show the  $P, Q$  system.

When we invoke quantum mechanics, we write  $E = \hbar\omega$ , where  $\hbar = h/2\pi$ , and  $h$  is Planck's constant. The coordinate and momentum  $q$  and  $p$  can be normalized as

$$q' = \sqrt{\frac{m\omega^2}{2E}}q \quad \text{and} \quad p' = \frac{p}{\sqrt{2mE}}.$$

to make the phase space plot of  $p'$  versus  $q'$  a circle of area  $\pi$ . This normalized form will be useful in Section 11.1 on chaos.

It would seem that the use of contact transformations to solve the harmonic oscillator problem is similar to “cracking a peanut with a sledge hammer.” We have here however a simple example of how the Hamiltonian can be reduced to a form cyclic in all coordinates by means of canonical transformations. Discussion of general schemes for the solution of mechanical problems by this technique will be reserved for the next chapter. For the present, we shall continue to examine the formal properties of canonical transformations.

#### 9.4 ■ THE SYMPLECTIC APPROACH TO CANONICAL TRANSFORMATIONS

Another method of treating canonical transformations, seemingly unrelated to the generator formalism, can be expressed in terms of the matrix or symplectic formulation of Hamilton’s equations. By way of introduction to this approach, let us consider a restricted canonical transformation, that is, one in which time does not appear in the equations of transformation:

$$\begin{aligned} Q_i &= Q_i(q, p), \\ P_i &= P_i(q, p). \end{aligned} \quad (9.44)$$

We know that the Hamiltonian function does not change in such a transformation. The time derivative of  $Q_i$ , on the basis of Eqs. (9.44), is to be found as

$$\dot{Q}_i = \frac{\partial Q_i}{\partial q_j} \dot{q}_j + \frac{\partial Q_i}{\partial p_j} \dot{p}_j = \frac{\partial Q_i}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial Q_i}{\partial p_j} \frac{\partial H}{\partial q_j}. \quad (9.45)$$

On the other hand, the inverses of Eqs. (9.44),

$$\begin{aligned} q_j &= q_j(Q, P), \\ p_j &= p_j(Q, P), \end{aligned} \quad (9.46)$$

enables us to consider  $H(q, p, t)$  as a function of  $Q$  and  $P$  and to form the partial derivative

$$\frac{\partial H}{\partial P_i} = \frac{\partial H}{\partial p_j} \frac{\partial p_j}{\partial P_i} + \frac{\partial H}{\partial q_j} \frac{\partial q_j}{\partial P_i}. \quad (9.47)$$

Comparing Eqs. (9.45) and (9.47), it can be concluded that

$$\dot{Q}_i = \frac{\partial H}{\partial P_i};$$

that is, the transformation is canonical, only if

$$\left( \frac{\partial Q_i}{\partial q_j} \right)_{q,p} = \left( \frac{\partial p_j}{\partial P_i} \right)_{Q,P}, \quad \left( \frac{\partial Q_i}{\partial p_j} \right)_{q,p} = - \left( \frac{\partial q_j}{\partial P_i} \right)_{Q,P}. \quad (9.48a)$$

The subscripts on the derivatives are to remind us that on the left-hand side of these equations  $Q_i$  is considered as a function of  $(q, p)$  (cf. Eqs. (9.44)), while on the right-hand side the derivatives are for  $q_j$  and  $p_j$  as functions of  $(Q, P)$  (cf. Eqs. (9.46)). A similar comparison of  $\dot{P}_i$  with the partial of  $H$  with respect to  $Q_j$  leads to the conditions

$$\left( \frac{\partial P_i}{\partial q_j} \right)_{q,p} = - \left( \frac{\partial p_j}{\partial Q_i} \right)_{Q,P}, \quad \left( \frac{\partial P_i}{\partial p_j} \right)_{q,p} = \left( \frac{\partial q_j}{\partial Q_i} \right)_{Q,P}. \quad (9.48b)$$

The sets of Eqs. (9.48) together are sometimes known as the “direct conditions” for a (restricted) canonical transformation.

The algebraic manipulation that leads to Eqs. (9.48) can be performed in a compact and elegant manner if we make use of the symplectic notation for the Hamiltonian formulation introduced above at the end of Section 8.1. If  $\eta$  is a column matrix with the  $2n$  elements  $q_i, p_i$ , then Hamilton’s equations can be written, it will be remembered, as Eq. (8.39)

$$\dot{\eta} = J \frac{\partial H}{\partial \eta},$$

where  $J$  is the antisymmetric matrix defined in Eq. (8.38a). Similarly the new set

of equations for the variables in the new canonical transformation the equations of transformation (9.44),

$$\zeta = \zeta(\eta).$$

Analogously to Eq. (9.45) we can seek the equations of motion for the variables by looking at the time derivative of a typical element of  $\zeta$

$$\dot{\zeta}_i = \frac{\partial \zeta_i}{\partial \eta_j} \dot{\eta}_j, \quad i, j = 1, \dots, 2n.$$

In matrix notation, this time derivative can be written as

$$\dot{\zeta} = M \dot{\eta},$$

where  $M$  is the Jacobian matrix of the transformation with elements

$$M_{ij} = \frac{\partial \zeta_i}{\partial \eta_j}.$$

Making use of the equations of motion for  $\eta$ , Eq. (9.50) becomes

$$\dot{\zeta} = M J \frac{\partial H}{\partial \eta}. \quad (9.52)$$

Now, by the inverse transformation  $H$  can be considered as a function of  $\zeta$ , and the derivative with respect to  $\eta_i$  evaluated as

$$\frac{\partial H}{\partial \eta_i} = \frac{\partial H}{\partial \zeta_j} \frac{\partial \zeta_j}{\partial \eta_i},$$

or, in matrix notation\*

$$\frac{\partial H}{\partial \eta} = \tilde{M} \frac{\partial H}{\partial \zeta}. \quad (9.53)$$

The combination of Eqs. (9.52) and (9.53) leads to the form of the equations of motion for any set of variables  $\zeta$  transforming, independently of time, from the canonical set  $\eta$ :

$$\dot{\zeta} = MJ\tilde{M} \frac{\partial H}{\partial \zeta}. \quad (9.54)$$

We have the advantage of knowing from the generator formalism that for a *restricted* canonical transformation the old Hamiltonian expressed in terms of the new variables serves as the new Hamiltonian:

$$\dot{\zeta} = J \frac{\partial H}{\partial \zeta}. \quad (9.54')$$

The transformation, Eq. (9.49), will therefore be canonical if  $M$  satisfies the condition

$$MJ\tilde{M} = J. \quad (9.55)$$

That Eq. (9.55) is also a necessary condition for a restricted canonical transformation is easily shown directly by reversing the order of the steps of the proof. Note that for an extended time-independent canonical transformation, where  $K = \lambda H$ , the condition of Eq. (9.55) would be replaced by

$$MJ\tilde{M} = \lambda J. \quad (9.56)$$

Equation (9.55) may be expressed in various forms. Multiplying from the right by the matrix inverse to  $\tilde{M}$  leads to

$$MJ = J\tilde{M}^{-1}, \quad (9.57)$$

(since the transpose of the inverse is the inverse of the transpose). The elements of the matrix equation (9.57) will be found to be identical with Eqs. (9.48a) and (9.48b). If Eq. (9.57) is multiplied by  $J$  from the left and  $-J$  from the right, then by virtue of Eq. (8.38e) we have

$$JM = \tilde{M}^{-1}J,$$

\*Readers of Section 7.5 will have recognized that Eq. (9.50) is the statement that  $\eta$  transforms contravariantly (as a vector) under the transformation, and Eq. (9.53) says that the partial derivative of  $H$  with respect to the elements of  $\eta$  transforms covariantly (or as a 1-form) (cf. Eqs. (7.50) and (7.54)).

or

$$\tilde{\mathbf{M}}\mathbf{J}\mathbf{M} = \mathbf{J}. \quad (9.58)$$

Equation (9.55), or its equivalent version, Eq. (9.58), is spoken of as the *symplectic condition* for a canonical transformation, and the matrix  $\mathbf{M}$  satisfying the condition is said to be a *symplectic matrix*.

These concepts may become more obvious if we display the details of the  $\mathbf{J}$  and  $\mathbf{M}$  matrices corresponding to the mixed generating function  $F = F_2(q_1, P_1) + F_1(q_2, Q_2)$  of Eq. (9.33). The variables  $\boldsymbol{\eta}$  and  $\boldsymbol{\zeta}$  are column vectors given by

$$\boldsymbol{\eta} = \begin{bmatrix} q_1 \\ q_2 \\ p_1 \\ p_2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\zeta} = \begin{bmatrix} Q_1 \\ Q_2 \\ P_1 \\ P_2 \end{bmatrix}.$$

The transformation  $\dot{\boldsymbol{\zeta}} = \mathbf{M}\dot{\boldsymbol{\eta}}$  (cf. Eq. (9.50)) is made by the following  $\mathbf{M}$  matrix:

$$\begin{bmatrix} \dot{Q}_1 \\ \dot{Q}_2 \\ \dot{P}_1 \\ \dot{P}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} \dot{q}_1 \\ \dot{p}_2 \\ \dot{p}_1 \\ -\dot{q}_2 \end{bmatrix},$$

in agreement with the expressions obtained by differentiating the results of the generating function with respect to time (cf. Column 3, Table 9.1). Hamilton's equations for the transformed variables  $\dot{\boldsymbol{\zeta}} = \mathbf{J}\frac{\partial H}{\partial \boldsymbol{\zeta}}$  (Eq. (9.54')) are expressed as follows independent of the generating function  $F$

$$\begin{bmatrix} \dot{Q}_1 \\ \dot{Q}_2 \\ \dot{P}_1 \\ \dot{P}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\dot{P}_1 \\ -\dot{P}_2 \\ \dot{Q}_1 \\ \dot{Q}_2 \end{bmatrix}$$

where  $-\dot{P}_i = \partial H / \partial \zeta_i$  for  $\zeta_1$  and  $\zeta_2$  and  $\dot{Q}_i = \partial H / \partial \zeta_i$  for  $\zeta_3$  and  $\zeta_4$ . Note that  $\mathbf{M}$  depends on  $F$  whereas  $\mathbf{J}$  does not (cf. Eq. (8.38a)). This formalism is not applicable to all cases. For example, a simple  $\mathbf{M}$  matrix cannot be written for the harmonic oscillator example discussed in Section 9.3.

For a canonical transformation that contains the time as a parameter, the simple derivation given for the symplectic condition no longer holds. Nonetheless, the symplectic condition remains a necessary and sufficient condition for a canonical transformation even if it involves the time. It is possible to prove the general validity of the symplectic conditions for all canonical transformations by straightforward, albeit lengthy, procedures resembling those employed for restricted canonical transformations. Instead we shall take a different tack, one that takes

advantage of the parametric form of the canonical transformations involving time. A canonical transformation of the form

$$\zeta = \zeta(\eta, t) \quad (9.59)$$

evolves continuously as time increases from some initial value  $t_0$ . It is a single-parameter instance of the family of continuous transformations first studied systematically by the mathematician Sophus Lie and as such plays a distinctive role in the transformation theory of classical mechanics.

If the transformation

$$\eta \rightarrow \zeta(t) \quad (9.60a)$$

is canonical, then so obviously is the transformation

$$\eta \rightarrow \zeta(t_0). \quad (9.60b)$$

It follows then from the definition of canonical transformation that the transformation characterized by

$$\zeta(t_0) \rightarrow \zeta(t) \quad (9.60c)$$

is also canonical. Since  $t_0$  in Eq. (9.60b) is a fixed constant, this canonical transformation satisfies the symplectic condition (9.58). If now the transformation of Eq. (9.60c) obeys the symplectic condition, it is easy to show (cf. Derivation 13) that the general transformation Eq. (9.60a) will also.

To demonstrate that the symplectic condition does indeed hold for canonical transformations of the type of Eq. (9.60c), we introduce the notion of an *infinitesimal canonical transformation* (abbreviated I.C.T.), a concept that will prove to be widely useful. As in the case of infinitesimal rotations, such a transformation is one in which the new variables differ from the old only by infinitesimals. Only first-order terms in these infinitesimals are to be retained in all calculations. The transformation equations can then be written as

$$Q_i = q_i + \delta q_i, \quad (9.61a)$$

$$P_i = p_i + \delta p_i, \quad (9.61b)$$

or in matrix form

$$\zeta = \eta + \delta \eta. \quad (9.61c)$$

(Here  $\delta q_i$  and  $\delta p_i$  do *not* represent virtual displacements but are simply the infinitesimal changes in the coordinates and momenta.) An infinitesimal canonical transformation thus differs only infinitesimally from the identity transformation discussed in Section 9.1. In the generator formalism, a suitable generating function for an I.C.T. would therefore be

$$F_2 = q_i P_i + \epsilon G(q, P, t), \quad (9.62)$$

where  $\epsilon$  is some infinitesimal parameter of the transformation, and  $G$  is any (differentiable) function of its  $2n + 1$  arguments. By Eq. (9.17a), the transformation equations for the momenta are to be found from

$$p_J = \frac{\partial F_2}{\partial q_J} = P_J + \epsilon \frac{\partial G}{\partial q_J}$$

or

$$\delta p_J \equiv P_J - p_J = -\epsilon \frac{\partial G}{\partial q_J}. \quad (9.63a)$$

Similarly, by Eq. (9.17b), the transformation equations for  $Q_J$  are determined by the relations

$$Q_J = \frac{\partial F_2}{\partial P_J} = q_J + \epsilon \frac{\partial G}{\partial P_J}.$$

Since the second term is already linear in  $\epsilon$ , and  $P$  differs from  $p$  only by an infinitesimal, it is consistent to first order to replace  $P_J$  in the derivative function by  $p_J$ . We may then consider  $G$  as a function of  $q$ ,  $p$  only (and possibly  $t$ ). Following the usual practice, we will refer to  $G(q, p)$  as the *generating function of the infinitesimal canonical transformation*, although strictly speaking that designation belongs only to  $F$ . The transformation equation for  $Q_J$  can therefore be written as

$$\delta q_J = \epsilon \frac{\partial G}{\partial p_J}. \quad (9.63b)$$

Both transformation equations can be combined into one matrix equation

$$\delta \boldsymbol{\eta} = \epsilon \mathbf{J} \frac{\partial G}{\partial \boldsymbol{\eta}}. \quad (9.63c)$$

An obvious example of an infinitesimal canonical transformation would be the transformation of Eq. (9.60c) when  $t$  differs from  $t_0$  by an infinitesimal  $t$ :

$$\zeta(t_0) \rightarrow \zeta(t_0 + dt), \quad (9.64)$$

with  $dt$  as the infinitesimal parameter  $\epsilon$ . The continuous evolution of the transformation  $\zeta(\eta, t)$  from  $\zeta(\eta, t_0)$  means that the transformation  $\zeta(t_0) \rightarrow \zeta(t)$  can be built up as a succession of such I.C.T.'s in steps of  $dt$ . It will therefore suffice to show that the infinitesimal transformation, Eq. (9.64), satisfies the symplectic condition (9.58). But it follows from the transformation equations (9.63) that the Jacobian matrix of any I.C.T. is a symplectic matrix. By definition the Jacobian matrix (9.51) for an infinitesimal transformation is

$$\mathbf{M} \equiv \frac{\partial \zeta}{\partial \boldsymbol{\eta}} = \mathbf{1} + \frac{\partial \delta \boldsymbol{\eta}}{\partial \boldsymbol{\eta}},$$

or by Eq. (9.63c)

$$\mathbf{M} = \mathbf{1} + \epsilon \mathbf{J} \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}}. \quad (9.65)$$

The second derivative in Eq. (9.65) is a square, symmetric matrix with elements

$$\left( \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}} \right)_{ij} = \frac{\partial^2 G}{\partial \eta_i \partial \eta_j}.$$

Because of the antisymmetrical property of  $\mathbf{J}$ , the transpose of  $\mathbf{M}$  is

$$\tilde{\mathbf{M}} = \mathbf{1} - \epsilon \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}} \mathbf{J}. \quad (9.66)$$

The symplectic condition involves the value of the matrix product

$$\mathbf{M} \mathbf{J} \tilde{\mathbf{M}} = \left( \mathbf{1} + \epsilon \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}} \right) \mathbf{J} \left( \mathbf{1} - \epsilon \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}} \mathbf{J} \right).$$

Consistent to first order in this product is

$$\begin{aligned} \mathbf{M} \mathbf{J} \tilde{\mathbf{M}} &= \mathbf{J} + \epsilon \mathbf{J} \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}} \mathbf{J} - \mathbf{J} \epsilon \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}} \mathbf{J} \\ &= \mathbf{J}, \end{aligned}$$

thus demonstrating that the symplectic condition holds for any infinitesimal canonical transformation. By the chain of reasoning we have spun out, it therefore follows that *any* canonical transformation, whether or not it involves time as a parameter, obeys the symplectic conditions, Eqs. (9.55) and (9.58).

The symplectic approach, for the most part, has been developed independently of the generating function method, except in the treatment of infinitesimal canonical transformations. They are of course connected. We shall sketch later, for example, a proof that the symplectic condition implies the existence of a generating function. But the connection is largely irrelevant. Both are valid ways of looking at canonical transformations, and both encompass all of the needed properties of the transformations. For example, either the symplectic or the generator formalisms can be used to prove that canonical transformations have the four properties that characterize a group (cf. Appendix B).

1. The identity transformation is canonical.
2. If a transformation is canonical, so is its inverse.
3. Two successive canonical transformations (the group “product” operation) define a transformation that is also canonical.
4. The product operation is associative.

We shall therefore be free to use either the generator or the symplectic approach at will, depending on which leads to the simplest treatment at the moment.

## 9.5 ■ POISSON BRACKETS AND OTHER CANONICAL INVARIANTS

The *Poisson bracket* of two functions  $u, v$  with respect to the canonical variables  $(q, p)$  is defined as

$$[u, v]_{q,p} = \frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q_i}. \quad (9.67)$$

In this bilinear expression we have a typical symplectic structure, as in Hamilton's equations, where  $q$  is coupled with  $p$ , and  $p$  with  $-q$ . The Poisson bracket thus lends itself readily to being written in matrix form, where it appears as

$$[u, v]_{\eta} = \widetilde{\frac{\partial u}{\partial \eta}} \frac{\partial v}{\partial \eta}. \quad (9.68)$$

The transpose sign is used on the first matrix on the right-hand side to indicate explicitly that this matrix must be treated as a single-row matrix in the multiplication. On most occasions this specific reminder will not be needed and the transpose sign may be omitted.

Suppose we choose the functions  $u, v$  out of the set of canonical variables  $(q, p)$  themselves. Then it follows trivially from the definition, either as Eq. (9.67) or (9.68), that these Poisson brackets have the values

$$[q_j, q_k]_{q,p} = 0 = [p_j, q_k]_{q,p},$$

and

$$[q_j, p_k]_{q,p} = \delta_{jk} = -[p_j, q_k]_{q,p}. \quad (9.69)$$

We can summarize the relations of Eqs. (9.69) in one equation by introducing a *square matrix Poisson bracket*,  $[\eta, \eta]$ , whose  $lm$  element is  $[\eta_l, \eta_m]$ . Equations (9.69) can then be written as

$$[\eta, \eta]_{\eta} = J. \quad (9.70)$$

Now let us take for  $u, v$  the members of the transformed variables  $(Q, P)$ , or  $\zeta$ , defined in terms of  $(q, p)$  by the transformation equations (9.59). The set of all the Poisson brackets that can be formed out of  $(Q, P)$  comprise the matrix Poisson bracket defined as

$$[\zeta, \zeta]_{\eta} = \widetilde{\frac{\partial \zeta}{\partial \eta}} \frac{\partial \zeta}{\partial \eta}.$$

But we recognize the partial derivatives as defining the square Jacobian matrix of the transformation, so that the Poisson bracket relation is equivalent to

$$[\zeta, \zeta]_{\eta} = \tilde{\mathbf{M}} \mathbf{J} \mathbf{M}. \quad (9.71)$$

If the transformation  $\eta \rightarrow \zeta$  is canonical, then the symplectic condition holds and Eq. (9.71) reduces to (cf. Eq. (9.58))

$$[\zeta, \zeta]_{\eta} = \mathbf{J}, \quad (9.72)$$

and conversely, if Eq. (9.72) is valid, then the transformation is canonical.

Poisson brackets of the canonical variables themselves, such as Eqs. (9.70) or (9.72), are referred to as the *fundamental Poisson brackets*. Since we have from Eq. (9.70) that

$$[\zeta, \zeta]_{\zeta} = \mathbf{J}, \quad (9.73)$$

Eq. (9.72) states that the fundamental Poisson brackets of the  $\zeta$  variables have the same value when evaluated with respect to *any* canonical coordinate set. In other words, the *fundamental Poisson brackets are invariant under canonical transformation*. We have seen from Eq. (9.71) that the invariance is a necessary and sufficient condition for the transformation matrix to be symplectic. The invariance of the fundamental Poisson brackets is thus in all ways equivalent to the symplectic condition for a canonical transformation.

It does not take many more steps to show that *all* Poisson brackets are invariant under canonical transformation. Consider the Poisson bracket of two functions  $u, v$  with respect to the  $\eta$  set of coordinates, Eq. (9.68). In analogy to Eq. (9.53), the partial derivative of  $v$  with respect to  $\eta$  can be expressed in terms of partial derivatives with respect to  $\zeta$  as

$$\frac{\partial v}{\partial \eta} = \tilde{\mathbf{M}} \frac{\partial v}{\partial \zeta}$$

(that is, the partial derivative transforms as a 1-form). In a similar fashion,

$$\widetilde{\frac{\partial u}{\partial \eta}} = \widetilde{\tilde{\mathbf{M}} \frac{\partial u}{\partial \zeta}} = \widetilde{\frac{\partial u}{\partial \zeta}} \mathbf{M}.$$

Hence the Poisson bracket Eq. (9.68) can be written

$$[u, v]_{\eta} = \frac{\widetilde{\partial u}}{\partial \eta} \mathbf{J} \frac{\partial v}{\partial \eta} = \frac{\widetilde{\partial u}}{\partial \zeta} \mathbf{M} \mathbf{J} \tilde{\mathbf{M}} \frac{\partial v}{\partial \zeta}.$$

If the transformation is canonical, the symplectic condition in the form of Eq. (9.55) holds, and we then have

$$[u, v]_{\eta} = \frac{\widetilde{\partial u}}{\partial \zeta} \mathbf{J} \frac{\partial v}{\partial \zeta} \equiv [u, v]_{\zeta}. \quad (9.74)$$

Thus, the Poisson bracket has the same value when evaluated with respect to any canonical set of variables—*all Poisson brackets are canonical invariants*. In writing the symbol for the Poisson bracket, we have so far been careful to indicate by the subscript the set of variables in terms of which the brackets are defined. So long as we use only canonical variables that practice is now seen to be unnecessary, and we shall in general drop the subscript.\*

The hallmark of the canonical transformation is that Hamilton's equations of motion are invariant in form under the transformation. Similarly, the canonical invariance of Poisson brackets implies that equations expressed in terms of Poisson brackets are invariant in form under canonical transformation. As we shall see, we can develop a structure of classical mechanics, paralleling the Hamiltonian formulation, expressed solely in terms of Poisson brackets. Historically this Poisson bracket formulation, which has the same form in all canonical coordinates, was especially useful for carrying out the original transition from classical to quantum mechanics. There is a simple “correspondence principle” that says that the classical Poisson bracket is to be replaced by a suitably defined commutator of the corresponding quantum operators.

The algebraic properties of the Poisson bracket are therefore of considerable interest. We have already used the obvious properties

$$[u, u] = 0, \quad (9.75a)$$

$$[u, v] = -[v, u]. \quad (\text{antisymmetry}) \quad (9.75b)$$

Almost equally obvious are the characteristics

$$[au + bv, w] = a[u, w] + b[v, w], \quad (\text{linearity}) \quad (9.75c)$$

where  $a$  and  $b$  are constants, and

$$[uv, w] = [u, w]v + u[v, w]. \quad (9.75d)$$

One other property is far from obvious, but is very important in defining the nature of the Poisson bracket. It is usually given in the form of *Jacobi's identity*, which states that if  $u$ ,  $v$ , and  $w$  are three functions with continuous second derivatives, then

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0; \quad (9.75e)$$

that is, the sum of the cyclic permutations of the double Poisson bracket of three functions is zero. There seems to be no simple way of proving Jacobi's identity for the Poisson bracket without lengthy algebra. However, it is possible to mitigate the complexity of the manipulations by introducing a special nomenclature. We

\*Note that for a scale transformation, or an extended canonical transformation, where the symplectic condition takes on the form of Eq (9.56), then Poisson brackets do *not* have the same values in all coordinate systems. That is one of the reasons scale transformations are excluded from the class of canonical transformations that are useful to consider.

shall use subscripts on  $u, v, w$  (or functions of them) to denote partial derivatives by the corresponding canonical variable. Thus,

$$u_i \equiv \frac{\partial u}{\partial \eta_i}, \quad \text{and} \quad v_{ij} \equiv \frac{\partial v}{\partial \eta_i \partial \eta_j}.$$

In this notation the Poisson bracket of  $u$  and  $v$  can be expressed as

$$[u, v] = u_i J_{ij} v_j.$$

Here  $J_{ij}$ , as usual, is simply the  $ij$ th element of  $\mathbf{J}$ . In the proof, the only property of  $\mathbf{J}$  that we shall need is its antisymmetry.

Now let us consider the first double Poisson bracket in Eq. (9.75e):

$$[u, [v, w]] = u_i J_{ij} [v, w]_j = u_i J_{ij} (v_k J_{kl} w_l)_j.$$

Because the elements  $J_{kl}$  are constants, the derivative with respect to  $\eta$  doesn't act on them, and we have

$$[u, [v, w]] = u_i J_{ij} (v_k J_{kl} w_l_j + v_{kj} J_{kl} w_l). \quad (9.76)$$

The other double Poisson brackets can be obtained from Eq. (9.76) by cyclic permutation of  $u, v, w$ . There are thus six terms in all, each being a fourfold sum over dummy indices  $i, j, k$ , and  $l$ . Consider the term in Eq. (9.76) involving a second derivative of  $w$ :

$$J_{ij} J_{kl} u_i v_k w_{lj}.$$

The only other second derivative of  $w$  will appear in evaluating the second double Poisson bracket in (Eq. 9.75e):

$$[v, [w, u]] = v_k J_{kl} (w_j J_{ji} u_i)_l.$$

Here the term in the second derivative in  $w$  is

$$J_{ji} J_{kl} u_i v_k w_{jl}.$$

Since the order of differentiation is immaterial,  $w_{lj} = w_{jl}$ , and the sum of the two terms is given by

$$(J_{ij} + J_{ji}) J_{kl} u_i v_k w_{lj} = 0,$$

by virtue of the antisymmetry of  $J$ . The remaining four terms are cyclic permutations and can similarly be divided in two pairs, one involving second derivatives of  $u$  and the other of  $v$ . By the same reasoning, each of these pairs sums to zero, and Jacobi's identity is thus verified.

If the Poisson bracket of  $u, v$  is looked on as defining a "product" operation of the two functions, then Jacobi's identity is the replacement for the associa-

tive law of multiplication. Recall that the ordinary multiplication of arithmetic is associative; that is, the order of a sequence of multiplications is immaterial:

$$a(bc) = (ab)c.$$

Jacobi's identity says that the bracket "product" is not associative and gives the effect of changing the sequence of "multiplications." Brackets that satisfy Eqs. (9.75), together with the expression

$$[u_i, u_j] = \sum_k c_{ij}^k u_k. \quad (9.77)$$

constitute a generally noncommutative algebra called a Lie algebra. For Poisson brackets in three-dimensional space, either the structure constants  $c_{ij}^k$  are all zero or only one term in the right-hand side of Eq. (9.77) exists for any pair of indices. Examples of this will be given later, and a more detailed discussion of Lie algebras is given in Appendix B.

Poisson bracket operation is not the only type of "product" familiar to physicists that satisfies the conditions for a Lie algebra. It will be left to the exercises to show that that vector product of two vectors,

$$\mathbf{v}[\mathbf{A}, \mathbf{B}] \rightarrow \mathbf{A} \times \mathbf{B}, \quad (9.78a)$$

and the commutator of two matrices,

$$\mathbf{M}[\mathbf{A}, \mathbf{B}] \rightarrow \mathbf{AB} - \mathbf{BA}, \quad (9.78b)$$

satisfy the same Lie algebra conditions as the Poisson bracket. It is this last that makes it feasible to replace the classical Poisson bracket by the commutator of the quantum mechanical operators. In other words, the "correspondence principle" can work only because both the Poisson bracket and commutator are representations of a Lie algebra "product."\*

There are other canonical invariants besides the Poisson bracket. One, mainly of historical interest now, is the *Lagrange bracket*, denoted by  $\{u, v\}$ . Suppose  $u$  and  $v$  are two functions out of a set of  $2n$  independent functions of the canonical variables. By inversion, the canonical variables can then be considered as functions of the set of  $2n$  functions. On this basis, the Lagrange bracket of  $u$  and  $v$  with respect to the  $(q, p)$  variables is defined as

\*Of course, we must not mistake the mathematical acceptability of this version of the correspondence principle with its physical necessity. The introduction of the quantum commutation relations was a great act of physical discovery by the pioneers of quantum mechanics. All we show here is that there is a similarity in the mathematical structure of the Poisson bracket formulation of classical mechanics and the commutation relation version of quantum mechanics. The formal correspondence is that

$$[u, v] \rightarrow \frac{1}{i\hbar}(uv - vu)$$

where on the left  $u, v$  are classical functions and on the right they are quantum operators.

$$\{u, v\}_{q,p} = \frac{\partial q_i}{\partial u} \frac{\partial p_i}{\partial v} - \frac{\partial p_i}{\partial u} \frac{\partial q_i}{\partial v}, \quad (9.79)$$

or, in matrix notation,

$$\{u, v\}_\eta = \frac{\tilde{\partial} \eta}{\partial u} \mathbf{J} \frac{\partial \eta}{\partial v}. \quad (9.80)$$

Proof of the canonical invariance of the Lagrange bracket parallels that for the Poisson bracket.

If for  $u$  and  $v$  we take two members of the set of canonical variables, then we obtain the *fundamental Lagrange brackets*:

$$\{q_i, q_j\}_{qp} = 0 = \{p_i, p_j\}_{qp} \quad \{q_i, p_j\}_{qp} = \delta_{ij}, \quad (9.81)$$

or, in matrix notation,

$$\{\eta, \eta\} = \mathbf{J}. \quad (9.82)$$

The Lagrange and Poisson brackets clearly stand in some kind of inverse relationship to each other, but the precise form of this relation is somewhat complicated to express. Let  $u_i$ ,  $i = 1, \dots, 2n$ , be a set of  $2n$  independent functions of the canonical variables, to be represented by a column (or row) matrix  $\mathbf{u}$ . Then  $\{\mathbf{u}, \mathbf{u}\}$  is the  $2n \times 2n$  matrix whose  $ij$ th element is  $\{u_i, u_j\}$ , with a similar description for  $[\mathbf{u}, \mathbf{u}]$ . The reciprocal character of the two brackets manifests itself in the relation

$$\{\mathbf{u}, \mathbf{u}\}[\mathbf{u}, \mathbf{u}] = -\mathbf{1}. \quad (9.83)$$

If for  $u$  we choose the canonical set itself,  $\eta$ , then Eq. (9.83) obviously follows from the fundamental bracket formulas, Eqs. (9.70) and (9.82), and the properties of  $\mathbf{J}$ . The proof for arbitrary  $u$  is not difficult if written in terms of the matrix definitions of the brackets and is reserved for the exercises. While the properties of the Lagrange and Poisson brackets parallel each other in many aspects, note that the Lagrange brackets do *not* obey Jacobi's identity. Lagrange brackets therefore do not qualify as a "product" operation in a Lie algebra.

Another important canonical invariant is the magnitude of a volume element in phase space. A canonical transformation  $\eta \rightarrow \zeta$  transforms the  $2n$ -dimensional phase space with coordinates  $\eta_i$  to another phase space with coordinates  $\zeta_i$ . The volume element

$$(d\eta) = dq_1 dq_2 \dots dq_n dp_1 \dots dp_n$$

transforms to a new volume element

$$(d\zeta) = dQ_1 dQ_2 \dots dQ_n dP_1 \dots dP_n.$$

As is well known, the sizes of the two volume elements are related by the absolute value of the Jacobian determinant  $\|\mathbf{M}\|$ ;

$$(d\zeta) = \|\mathbf{M}\|(d\eta).$$

For example, in the two-dimensional transformation from  $\eta_1 = q, p$  to  $\zeta_1 = Q, P$ , this expression becomes

$$dQ dP = \begin{vmatrix} \frac{\partial q}{\partial Q} & \frac{\partial q}{\partial P} \\ \frac{\partial p}{\partial Q} & \frac{\partial p}{\partial P} \end{vmatrix} dq dp = [q, p]_\zeta dq dp. \quad (9.84)$$

But, by taking the determinant of both sides of the symplectic condition, Eq. (9.58), we have

$$\|\mathbf{M}\|^2 |\mathbf{J}| = |\mathbf{J}|. \quad (9.85)$$

Thus, in a real canonical transformation the Jacobian determinant is  $\pm 1$ , and the absolute value is always unity, proving the canonical invariance of the volume element in phase space. It follows, also, that the volume of any arbitrary region in phase space,

$$J_n = \int \cdots \int (d\eta), \quad (9.86)$$

is a canonical invariant. In our two-dimensional example, the invariant is  $d\eta = dq dp$  and  $J_1 = \int dq dp$ .

The volume integral in Eq. (9.86) is the final member of a sequence of canonical invariants known as the *integral invariants of Poincaré*, comprising integrals over subspaces of phase space of different dimensions. The other members of the sequence cannot be stated as simply as  $J_n$ , and because they are not needed for the further development of the theory, they will not be discussed here.

Finally, the invariance of the fundamental Poisson brackets now enables us to outline a proof that the symplectic condition implies the existence of a generating function, as mentioned at the conclusion of the previous section. To simplify considerations, we shall examine only a system with one degree of freedom; the general method of the proof can be directly extended to systems with many degrees of freedom.\* We suppose that the first of the equations of transformation,

$$Q = Q(q, p), \quad P = P(q, p),$$

\*In the literature, the connection between the symplectic approach and the generator formalism is sometimes referred to as the *Caratheodory theorem*.

is invertable so as to give  $p$  as a function  $q$  and  $Q$ , say

$$p = \phi(q, Q). \quad (9.87)$$

Substitution in the second equation of transformation gives  $P$  as some function of  $q$  and  $Q$ , say

$$P = \psi(q, Q). \quad (9.88)$$

In such a case, we would expect the transformation to be generated by a generating function of the first kind,\*  $F_1$ , with Eqs. (9.87) and (9.88) appearing as

$$p = \frac{\partial F_1(q, Q)}{\partial q}, \quad P = -\frac{\partial F_1}{\partial Q}(q, Q). \quad (9.89)$$

If Eq. (9.89) holds, then it must be true that

$$\frac{\partial \phi}{\partial Q} = -\frac{\partial \psi}{\partial q}. \quad (9.90)$$

Conversely, if we can show that Eq. (9.90) is valid, then there must exist a function  $F_1$  such that  $p$  and  $P$  are given by Eqs. (9.89).

To demonstrate the validity of Eq. (9.90), we try to look on all quantities as functions of  $q$  and  $Q$ . Thus, we of course have the identity

$$\frac{\partial Q}{\partial Q} = 1,$$

but if Eq. (9.87) be substituted in the first transformation equation,

$$Q = Q(q, \phi(q, Q)), \quad (9.91)$$

the partial derivative can also be written

$$\frac{\partial Q}{\partial Q} = \frac{\partial Q}{\partial p} \frac{\partial \phi}{\partial Q},$$

so that we have the relation

$$\frac{\partial Q}{\partial p} \frac{\partial \phi}{\partial Q} = 1. \quad (9.92)$$

In the same spirit we evaluate the Poisson bracket

$$[Q, P] \equiv \frac{\partial Q}{\partial q} \frac{\partial P}{\partial p} - \frac{\partial P}{\partial q} \frac{\partial Q}{\partial p} = 1.$$

\*Of course, if the  $Q$  transformation equation is not invertable, as in the identity transformation, then we would invert the  $P$  equation and be led to a generating function of the second kind.

The derivatives of  $P$  are derivatives of  $\psi$  from Eq. (9.88) considered as a function of  $q$  and  $Q(q, p)$ . Hence, the Poisson bracket can be written

$$[Q, P] = \frac{\partial Q}{\partial q} \frac{\partial \psi}{\partial Q} \frac{\partial Q}{\partial p} - \frac{\partial Q}{\partial p} \left( \frac{\partial \psi}{\partial q} + \frac{\partial \psi}{\partial Q} \frac{\partial Q}{\partial q} \right),$$

or, consolidating terms, as

$$[Q, P] = \frac{\partial \psi}{\partial Q} \left( \frac{\partial Q}{\partial q} \frac{\partial Q}{\partial p} - \frac{\partial Q}{\partial p} \frac{\partial Q}{\partial q} \right) - \frac{\partial Q}{\partial p} \frac{\partial \psi}{\partial q},$$

and therefore

$$1 = - \frac{\partial Q}{\partial p} \frac{\partial \psi}{\partial q}. \quad (9.93)$$

Combining Eqs. (9.92) and (9.93), we have

$$\frac{\partial Q}{\partial p} \frac{\partial \phi}{\partial Q} = - \frac{\partial Q}{\partial p} \frac{\partial \psi}{\partial q}.$$

Since the partial derivative of  $Q$  with respect to  $p$  is the same on both sides of the equation, that is, the other variable being held constant is  $q$  in both cases, and since the derivative doesn't vanish (else the  $Q$  equation could not be inverted), it follows that Eq. (9.90) must be true. Thus, from the value of the fundamental Poisson bracket  $[Q, P]$ , which we have seen is equivalent to the symplectic condition, we are led to the existence of a generating function. The two approaches to canonical transformations, though arrived at independently, are fully equivalent.

## 9.6 ■ EQUATIONS OF MOTION, INFINITESIMAL CANONICAL TRANSFORMATIONS, AND CONSERVATION THEOREMS IN THE POISSON BRACKET FORMULATION

Almost the entire framework of Hamiltonian mechanics can be restated in terms of Poisson brackets. As a result of the canonical invariance of the Poisson brackets, the relations so obtained will also be invariant in form under a canonical transformation. Suppose, for example, we look for the total time derivative of some function of the canonical variables and time,  $u(q, p, t)$ , by use of Hamilton's equations of motion:

$$\frac{du}{dt} = \frac{\partial u}{\partial q_i} \dot{q}_i + \frac{\partial u}{\partial p_i} \dot{p}_i + \frac{\partial u}{\partial t} = \frac{\partial u}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial u}{\partial t},$$

or

$$\frac{du}{dt} = [u, H] + \frac{\partial u}{\partial t}. \quad (9.94)$$

In terms of the symplectic notation, the derivation of Eq. (9.94) would run

$$\frac{du}{dt} = \frac{\partial u}{\partial \eta} \dot{\eta} + \frac{\partial u}{\partial t} = \frac{\partial u}{\partial \eta} J \frac{\partial H}{\partial \eta} + \frac{\partial u}{\partial t},$$

from whence Eq. (9.94) follows, by virtue of (9.68). Equation (9.94) may be looked on as the generalized equation of motion for an arbitrary function  $u$  in the Poisson bracket formulation. It contains Hamilton's equations as a special case when for  $u$  we substitute one of the canonical variables

$$\dot{q}_i = [q_i, H], \quad \dot{p}_i = [p_i, H], \quad (9.95a)$$

or, in symplectic notation,

$$\dot{\eta} = [\eta, H]. \quad (9.95b)$$

That Eq. (9.95b) is identical with Hamilton's equations of motion may be seen directly from the observation that by the definition of the Poisson bracket, Eq. (8.39), we have

$$[\eta, H] = J \frac{\partial H}{\partial \eta}, \quad (9.96)$$

so that Eq. (9.95b) is simply another way of writing Eq. (8.31). Another familiar property may be obtained from Eq. (9.94) by taking  $u$  as  $H$  itself. Equation (9.94) then says that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t},$$

as was obtained previously in Eq. (8.41).

Note that the generalized equation of motion is canonically invariant; it is valid in whatever set of canonical variables  $q, p$  is used to express the function  $u$  or to evaluate the Poisson bracket. However, the Hamiltonian used must be appropriate to the particular set of canonical variables. Upon transforming to another set of variables by a time-dependent canonical transformation, we must also change to the transformed Hamiltonian  $K$ .

If  $u$  is a constant of the motion, then Eq. (9.94) says it must have the property

$$[H, u] = \frac{\partial u}{\partial t}. \quad (9.97)$$

All functions that obey Eq. (9.97) are constants of the motion, and conversely the Poisson bracket of  $H$  with any constant of the motion must be equal to the explicit time derivative of the constant function. We thus have a general test for seeking and identifying the constants of the system. For those constants of the motion not

involving the time explicitly, the test of Eq. (9.97) reduces to requiring that their Poisson brackets with the Hamiltonian vanish, that is,  $[H, u] = 0$ .\*

If two constants of the motion are known, the Jacobi identity provides a possible way for obtaining further constants. Suppose  $u$  and  $v$  are two constants of the motion not explicitly functions of time. Then if  $w$  in Eq. (9.75e) is taken to be  $H$ , the Jacobi identity says

$$[H, [u, v]] = 0;$$

that is, the Poisson bracket of  $u$  and  $v$  is also a constant in time. Even when the conserved quantities depend upon time explicitly, it can be shown with a bit more algebra (cf. Exercise 30) that *the Poisson bracket of any two constants of the motion is also a constant of the motion* (Poisson's theorem). Repeated application of the Jacobi identity in this manner can in principle lead to a complete sequence of constants of the motion. Quite often, however, the process is disappointing. The Poisson bracket of  $u$  and  $v$  frequently turns out to be a trivial function of  $u$  and  $v$  themselves, or even identically zero. Still, the possibility of generating new independent constants of motion by Poisson's theorem should be kept in mind.

The Poisson bracket notation can also be used to reformulate the basic equations of an infinitesimal canonical transformation. As discussed above (Section 9.4), such a transformation is a special case of a transformation that is a continuous function of a parameter, starting from the identity transformation at some initial value of the parameter (which may, for convenience, be set equal to zero). If the parameter is small enough to be treated as a first-order infinitesimal, then the transformed canonical variables differ only infinitesimally from the initial coordinates:

$$\zeta = \eta + \delta\eta \quad (9.98)$$

with the change being given in terms of the generator  $G$  through Eq. (9.63c):

$$\delta\eta = \epsilon J \frac{\partial G(\eta)}{\partial \eta}.$$

Now, by the definition (9.68) of the Poisson bracket, it follows that

$$[\eta, u] = J \frac{\partial u}{\partial \eta} \quad (9.99)$$

(cf. Eq. (9.96)), a relation that remains valid when the Poisson bracket is evaluated in terms of any other canonical variables. If  $u$  is taken to be  $G$ , it is seen that the equations of transformation for an infinitesimal canonical transformation can be

\*In view of the “correspondence principle” between the classical Poisson bracket and the quantum commutator, it is seen that this statement corresponds to the well-known quantum theorem that conserved quantities commute with the Hamiltonian.

written as

$$\delta \eta = \epsilon [\eta, G]. \quad (9.100)$$

Consider now an infinitesimal canonical transformation in which the continuous parameter is  $t$  (as was done in proving the symplectic condition) so that  $\epsilon = dt$ , and let the generating function  $G$  be the Hamiltonian. Then the equations of transformation for this I.C.T. become, by Eq. (9.100),

$$\delta \eta = dt [\eta, H] = \dot{\eta} dt = d\eta. \quad (9.101)$$

These equations state that the transformation changes the coordinates and momenta at the time  $t$  to the values they have at the time  $t + dt$ . Thus, the motion of the system in a time interval  $dt$  can be described by an infinitesimal contact transformation generated by the Hamiltonian. Correspondingly, the system motion in a finite time interval from  $t_0$  to  $t$  is represented by a succession of infinitesimal contact transformations, which, as we have seen, is equivalent to a single finite canonical transformation. Thus, the values of  $q$  and  $p$  at any time  $t$  can be obtained from their initial values by a canonical transformation that is a continuous function of time. According to this view, the motion of a mechanical system corresponds to the continuous evolution or unfolding of a canonical transformation. In a very literal sense, the *Hamiltonian is the generator of the system motion with time*.

Conversely, there must exist a canonical transformation from the values of the coordinates and momenta at any time  $t$  to their constant initial values. Obtaining such a transformation is obviously equivalent to solving the problem of the system motion. At the beginning of the chapter it was pointed out that a mechanical problem could be reduced to finding the canonical transformation for which all momenta are constants of the motion. The present considerations indicate the possibility of an alternative solution by means of the canonical transformation for which *both* the momenta and coordinates are constants of the motion. These two suggestions will be elaborated in the next chapter in order to show how formal solutions may be obtained for any mechanical problem.

Implicit to this discussion has been an altered way of looking at a canonical transformation and the effect it produces. The notion of a canonical transformation was introduced as a change of the coordinates used to characterize phase space. In effect, we switched from one phase space  $\eta$  with coordinates  $(q, p)$  to another,  $\zeta$ , with coordinates  $(Q, P)$ . If the state of the system at a given time was described by a point  $A$  in one system, it could also be described equally well by the transformed point  $A'$  (cf. Fig. 9.2). Any function of the system variables would have the same value for a given system configuration whether it was described by the  $(q, p)$  set or by the  $(Q, P)$  set. In other words, the function would have the same value at  $A'$  as at  $A$ . In analogy to the corresponding description of orthogonal transformations, we may call this the *passive* view of a canonical transformation.

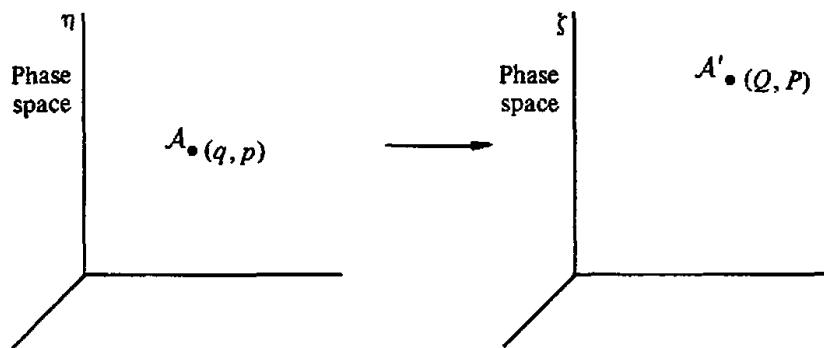


FIGURE 9.2 The passive view of a canonical transformation.

In contrast, we have spoken of the canonical transformation generated by the Hamiltonian as relating the coordinates of one point in phase space to those of another point in the *same* phase space. From this viewpoint, the canonical transformation accomplishes, in the mathematician's language, a mapping of the points of phase space onto themselves. In effect, we have an *active* interpretation of the canonical transformation as "moving" the system point from one position, with coordinates  $(q, p)$ , to another point,  $(Q, P)$ , in phase space (cf. Fig. 9.3). Of course, the canonical transformation in itself cannot move or change the system configuration. What it does is express one configuration of the system in terms of another. With some classes of canonical transformation, the active viewpoint is not helpful. For example, the point transformation from Cartesian coordinates to spherical polar coordinates is a canonical transformation of the passive type, and an "active" interpretation of it would border on the ludicrous.

The active viewpoint is particularly useful for transformations depending continuously on a single parameter. On the active interpretation, the effect of the transformation is to "move" the system point continuously on a curve in phase space as the parameter changes continuously. When the generator of the associated I.C.T. is the Hamiltonian, the curve on which the system point moves is the trajectory of the system in phase space.

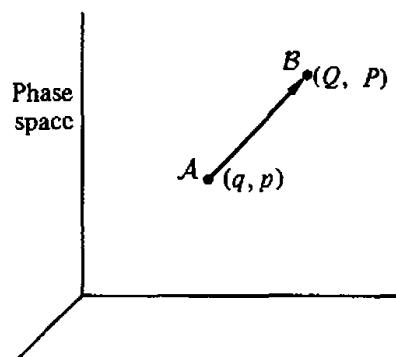


FIGURE 9.3 The active view of a canonical transformation.

If we pose the question, How does a function change under a canonical transformation? the answer depends on whether we should take an active or a passive point of view. From the passive point of view, the function changes in form, or in functional dependence, but it does not change in value. This is because in general the function, call it  $U$ , has a different functional dependence on  $(Q, P)$  than it does on  $(q, p)$ . Its value however remains the same at the corresponding points  $U(q_0, p_0)$  and  $U(Q_0, P_0)$  since  $Q_0 = Q(q_0, p_0)$  and  $P_0 = P(q_0, p_0)$ , so both sets of coordinates refer to the same physical location in phase space but use different coordinates to describe the phase space.

In contrast to this, if we consider the canonical transformation from an active point of view, then we are talking about a translation of the system from point  $\mathcal{A}$  to point  $\mathcal{B}$ , from position  $(q_{\mathcal{A}}, p_{\mathcal{A}})$  to position  $(q_{\mathcal{B}}, p_{\mathcal{B}})$ . From this point of view, the function  $U(q, p)$  does not change its functional dependence upon position and momentum, rather it changes its values as a result of replacing the values  $(q_{\mathcal{A}}, p_{\mathcal{A}})$  by  $(q_{\mathcal{B}}, p_{\mathcal{B}})$  in the function  $U(q, p)$ . There are then two distinct phase spaces, one using  $(q, p)$  and the other using  $(Q, P)$ . The transformation formalism uses the notation  $(q, p)$  for the variables at point  $\mathcal{A}$  and  $(Q, P)$  for the variables at point  $\mathcal{B}$ . This is analogous to a passive rotation in coordinate space corresponding to the rotation of the coordinate axes relative to a stationary object, and an active rotation corresponding to rotating an object relative to a fixed coordinate system.

We shall use the symbol  $\partial$  to denote a change in the value of a function under an "active" infinitesimal canonical transformation:

$$\partial u = u(\mathcal{B}) - u(\mathcal{A}), \quad (9.102)$$

where of course  $\mathcal{A}$  and  $\mathcal{B}$  will be infinitesimally close. Using the matrix notation for the canonical variables, the change in the function value under an I.C.T. would be defined as

$$\partial u = u(\boldsymbol{\eta} + \delta\boldsymbol{\eta}) - u(\boldsymbol{\eta}).$$

Expanding in a Taylor series and retaining terms in first-order infinitesimals, we have, by virtue of Eq. (9.63c),

$$\partial u = \frac{\partial u}{\partial \boldsymbol{\eta}} \delta \boldsymbol{\eta} = \epsilon \frac{\widetilde{\partial u}}{\partial \boldsymbol{\eta}} \mathbf{J} \frac{\partial G}{\partial \boldsymbol{\eta}}.$$

Recalling the definition of the Poisson bracket, Eq. (9.68), we see that the change can be written as

$$\partial u = \epsilon [u, G]. \quad (9.103)$$

An immediate application of Eq. (9.103) is to take for  $u$  one of the phase space coordinates themselves (or the matrix of the coordinates). We then have, by Eq. (9.100),

$$\partial \boldsymbol{\eta} = \epsilon [\boldsymbol{\eta}, G] = \delta \boldsymbol{\eta}.$$

Of course, this result is obvious from the definition of the point  $\mathcal{B}$  in relation to  $\mathcal{A}$ ; the “change” in the coordinates from  $\mathcal{A}$  to  $\mathcal{B}$  is just the infinitesimal difference between the old and new coordinates.

These considerations must be generalized somewhat in talking about the “change in the Hamiltonian.” Recall that the designation “Hamiltonian” does not mean a specific function, the same in all coordinate systems. Rather it refers to that function which in the given phase space defines the canonical equations of motion. Where the canonical transformation depends upon the time, the very meaning of “Hamiltonian” is also transformed. Thus,  $H(\mathcal{A})$  goes over not into  $H(\mathcal{A}')$  but into  $K(\mathcal{A}')$ , and  $H(\mathcal{A})$  will not necessarily have the same value as  $K(\mathcal{A}')$ . In such a case, we shall mean by  $\partial H$  in effect the difference in the value of the Hamiltonian under the two interpretations:

$$\partial H = H(\mathcal{B}) - K(\mathcal{A}'). \quad (9.104)$$

Where the function itself does not change under the canonical transformation the two forms for the change, Eqs. (9.102) and (9.104), are identical since  $u(\mathcal{A}') = u(\mathcal{A})$ . In general,  $K$  is related to  $H$  by the equation

$$K = H + \frac{\partial F}{\partial t},$$

where for an I.C.T. the generating function is given by Eq. (9.62) in terms of  $G$ . Since only  $G$  in that equation can be an explicit function of time, the value of the new Hamiltonian is given by

$$K(\mathcal{A}') = H(\mathcal{A}') + \epsilon \frac{\partial G}{\partial t} = H(\mathcal{A}) + \epsilon \frac{\partial G}{\partial t},$$

and the change in the Hamiltonian is

$$\partial H = H(\mathcal{B}) - H(\mathcal{A}) - \epsilon \frac{\partial G}{\partial t}. \quad (9.105)$$

Following along the path that led from Eq. (9.103), we see that  $\partial H$  is given by

$$\partial H = \epsilon [H, G] - \epsilon \frac{\partial G}{\partial t}. \quad (9.106)$$

From the generalized equation of motion, Eq. (9.106), with  $G$  as  $u$ , it follows finally that the change in  $H$  is

$$\partial H = -\epsilon \frac{dG}{dt}. \quad (9.107)$$

If  $G$  is a constant of the motion, Eq. (9.107) says that it generates an infinitesimal canonical transformation that does not change the value of the Hamiltonian. Equivalently, *the constants of the motion are the generating functions of those infinitesimal canonical transformations that leave the Hamiltonian invariant.* Im-

plied in this conclusion is a connection between the symmetry properties of the system and conserved quantities, a connection that is simplest to see for constants of the motion not explicitly depending upon time. The change in the Hamiltonian under the transformation is then simply the change in the value of the Hamiltonian as the system is moved from configuration  $\mathcal{A}$  to configuration  $\mathcal{B}$ . If the system is symmetrical under the operation that produces this change of configuration, then the Hamiltonian will obviously remain unaffected under the corresponding transformation. To take a simple example, if the system is symmetrical about a given direction, then the Hamiltonian will not change in value if the system as a whole is rotated about that direction. It follows then that the quantity that generates (through an I.C.T.) such a rotation of the system must be conserved. The rotational symmetry of the system implies a particular constant of the motion. This is not the first instance of a connection between constants of the motion and symmetry characteristics. We encountered it previously (Sections 2.6, 8.2) in connection with the conservation of generalized momenta. Here, however, the theorem is more elegant, and more complete, for it embraces all independent constants of the motion and not merely the conserved generalized momenta.

The momentum conservation theorems appear now as a special case of the general statement: If a coordinate  $q_i$  is cyclic, the Hamiltonian is independent of  $q_i$  and will certainly be invariant under an infinitesimal transformation that involves a displacement of  $q_i$  alone. Consider, now, a transformation generated by the generalized momentum conjugate to  $q_i$ :

$$G(q, p) = p_i. \quad (9.108)$$

By Eqs. (9.63a and b), the resultant infinitesimal canonical transformation is

$$\begin{aligned} \delta q_j &= \epsilon \delta_{ij}, \\ \delta p_i &= 0, \end{aligned} \quad (9.109)$$

that is, exactly the required infinitesimal displacement of  $q_i$  and only  $q_i$ . We readily recognize this as the familiar momentum theorem: If a coordinate is cyclic, its conjugate momentum is a constant of the motion. The observation that a displacement of one coordinate alone is generated by the conjugate momentum may be put in a slightly expanded form. If the generating function of an I.C.T. is given by

$$G_l = (\mathbf{J}\boldsymbol{\eta})_l = J_{lr}\eta_r, \quad (9.110)$$

then the equations of transformation as obtained from Eq. (9.63c) appear as

$$\delta\eta_k = \epsilon J_{ks} \frac{\partial G_l}{\partial \eta_s} = \epsilon J_{ks} J_{lr} \delta_{rs} = \epsilon J_{ks} J_{ls}.$$

By virtue of the orthogonality of  $\mathbf{J}$ , these reduce finally to

$$\delta\eta_k = \epsilon \delta_{kl}; \quad (9.111)$$

that is, a displacement of any canonical variables  $\eta_l$  alone is generated in terms of the conjugate variable in the form given by Eq. (9.110). Of course, if  $\eta_l$  is  $q_i$ ,  $G$  from Eq. (9.110) is just  $p_i$ , and if  $\eta_l$  is  $p_i$ ,  $G$  is then  $-q_i$ .

As a specific illustration of these concepts, let us consider again the infinitesimal contact transformation of the dynamical variables that produces a rotation of the system as a whole by an angle  $d\theta$ . The physical significance of the corresponding generating function cannot depend upon the choice of initial canonical coordinates,\* and it is convenient to use for this purpose the Cartesian coordinates of all particles in the system. Nor will there be any loss in generality if the axes are so oriented that the infinitesimal rotation is along the  $z$  axis. For an infinitesimal counterclockwise rotation of each particle, the change in the position vectors is to be found from the infinitesimal rotation matrix of Eq. (4.69). With a rotation only about the  $z$  axis, the changes in the particle coordinates are

$$\delta x_i = -y_i d\theta, \quad \delta y_i = x_i d\theta, \quad \delta z_i = 0. \quad (9.112a)$$

The effect of the transformation on the components of the Cartesian vectors formed by the momenta conjugate to the particle coordinates is similarly given by

$$\delta p_{ix} = -p_{iy} d\theta, \quad \delta p_{iy} = p_{ix} d\theta, \quad \delta p_{iz} = 0. \quad (9.112b)$$

Comparing these transformation equations with Eqs. (9.63a and b), it is seen that the corresponding generating function is

$$G = x_i p_{iy} - y_i p_{ix}, \quad (9.113)$$

with  $d\theta$  as the infinitesimal parameter  $\epsilon$ . For a direct check, note that

$$\begin{aligned} \delta x_i &= d\theta \frac{\partial G}{\partial p_{ix}} = -y_i d\theta, & \delta p_{ix} &= -d\theta \frac{\partial G}{\partial x_i} = -p_{iy} d\theta, \\ \delta y_i &= d\theta \frac{\partial G}{\partial p_{iy}} = x_i d\theta, & \delta p_{iy} &= -d\theta \frac{\partial G}{\partial y_i} = p_{ix} d\theta, \end{aligned}$$

agreeing with Eqs. (9.112). The generating function (9.113) in addition has the physical significance of being the  $z$ -component of the total canonical angular momentum:

$$G = L_z \equiv (r_i \times p_i)_z. \quad (9.114)$$

Since the  $z$  axis was arbitrarily chosen, we can state that the generating function corresponding to an infinitesimal rotation about an axis denoted by the unit vector

\*This can most easily be seen from the canonically invariant Eq. (9.100). The change in the canonical variable  $\eta_l$  remains the same no matter in what set of canonical variables  $G$  is expressed.

$\mathbf{n}$  is

$$G = \mathbf{L} \cdot \mathbf{n}. \quad (9.115)$$

Note that the canonical angular momentum as defined here may differ from the mechanical angular momentum. If the forces on the system are derivable from velocity-dependent potentials, then the canonical momentum vectors  $\mathbf{p}_i$  are not necessarily the same as the linear momentum vectors, and  $\mathbf{L}$  in Eqs. (9.114) and (9.115) may not be the same as the mechanical angular momentum. The result obtained here is therefore a generalization of the conclusion given in Section 2.6 that the momentum conjugate to a rotation coordinate is the corresponding component of the total angular momentum. The proof presented there was restricted to systems with velocity-independent potentials. By virtue of Eqs. (9.108) and (9.109), we can now conclude that the momentum conjugate to a generalized coordinate that measures the rotation of the system as a whole about an axis  $\mathbf{n}$  is the component of the total canonical angular momentum along the same axis. Just as the Hamiltonian is the generator of a displacement of the system in time, so the angular momentum is the generator of the spatial rotations of the system.

It has already been noted that on the “active” interpretation a canonical transformation depending upon a parameter “moves” the system point along a continuous trajectory in phase space. Since the finite transformation can be looked on as the sum of an infinite succession of infinitesimal canonical transformations, each corresponding to an infinitesimal displacement along the curve, it should therefore be possible formally to obtain the finite transformation by integrating the expression for the infinitesimal displacements. We can do this by noting that each point on the trajectory in phase space corresponds to a particular value of the parameter, which we shall call  $\alpha$ , starting from the initial system configuration denoted by  $\alpha = 0$ . If  $u$  is some function of the system configuration, then  $u$  will be a continuous function of  $\alpha$  along the trajectory,  $u(\alpha)$ , with initial value  $u_0 = u(0)$ . (For simplicity, we shall consider  $u$  as not depending explicitly upon time.) Equation (9.103) for the infinitesimal change of  $u$  on the trajectory can be written as

$$\delta u = d\alpha[u, G],$$

or as a differential equation in the variable  $\alpha$ :

$$\frac{du}{d\alpha} = [u, G]. \quad (9.116)$$

We can get  $u(\alpha)$ , and therefore the effect of the finite canonical transformation, by integrating this differential equation. A formal solution may be obtained by expanding  $u(\alpha)$  in a Taylor series about the initial conditions:

$$u(\alpha) = u_0 + \alpha \frac{du}{d\alpha} \Big|_0 + \frac{\alpha^2}{2!} \frac{d^2u}{d\alpha^2} \Big|_0 + \frac{\alpha^3}{3!} \frac{d^3u}{d\alpha^3} \Big|_0 + \dots$$

By Eq. (9.116), we have

$$\frac{du}{d\alpha} \Big|_0 = [u, G]_0,$$

the zero subscript meaning that the value of the Poisson bracket is to be taken at the initial point,  $\alpha = 0$ . Repeated application of Eq. (9.116), taking  $[u, G]$  itself as a function of the system configuration, gives

$$\frac{d^2 u}{d\alpha^2} = [[u, G], G],$$

and the process can be repeated to give the third derivative of  $u$  and so on. The Taylor series for  $u(\alpha)$  thus leads to the formal series solution

$$u(\alpha) = u_0 + \alpha[u, G]_0 + \frac{\alpha^2}{2!} [[u, G], G]_0 + \frac{\alpha^3}{3!} [[[u, G], G], G]_0 + \dots \quad (9.117)$$

If for  $u$  we take any of the canonical variables  $\zeta_i$ , with  $u_0$  the starting set of variables  $\eta_i$ , then Eq. (9.115) is a prescription for finding the transformation equations of the finite canonical transformation generated by  $G$ .

It is not difficult to find specific examples showing that this procedure actually works. Suppose for  $G$  we take  $L_z$ , so that the final canonical transformation should correspond to a finite rotation about the  $z$  axis. The natural parameter to use for  $\alpha$  is the rotation angle. For  $u$ , let us take the  $x$ -coordinate of the  $i$ th particle in the system. Either by direct evaluation of the Poisson brackets or by inference from Eqs. (9.112a), it is easy to see that

$$[X_i, L_z] = -Y_i, \quad [Y_i, L_z] = X_i, \quad (9.118)$$

where capital letters have been used to denote the coordinates after some rotation  $\theta$ , that is, the final coordinate. The initial coordinates, that is, before rotation, are as usual represented by lowercase letters. It follows then that

$$\begin{aligned} [X_i, L_z]_0 &= -y_i, \\ [[X_i, L_z], L_z]_0 &= -[Y_i, L_z]_0 = -x_i, \\ [[[X_i, L_z], L_z], L_z]_0 &= -[X_i, L_z]_0 = y_i, \end{aligned}$$

and so on. The series representation for  $X_i$  thus becomes

$$\begin{aligned} X_i &= x_i - y_i \theta - x_i \frac{\theta^2}{2} + y_i \frac{\theta^3}{3!} + x_i \frac{\theta^4}{4!} - \dots \\ &= x_i \left( 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots \right) - y_i \left( \theta - \frac{\theta^3}{3!} + \dots \right). \end{aligned}$$

The two series will be recognized as the expansion for the cosine and sine, respectively. Hence, the equation for the finite transformation of  $X_i$  is

$$X_i = x_i \cos \theta - y_i \sin \theta,$$

which is exactly what we would expect for the finite rotation of a vector counterclockwise about the  $z$  axis.

For another example, let us consider the situation when  $G = H$  and the parameter is the time. Equation (9.116) then reduces to the equation of motion for  $u$ :

$$\frac{du}{dt} = [u, H],$$

with the formal solution

$$u(t) = u_0 + t[u, H]_0 + \frac{t^2}{2!}[[u, H], H]_0 + \frac{t^3}{3!}[[[u, H], H], H]_0 + \dots \quad (9.119)$$

Here the subscript zero refers to the initial conditions at  $t = 0$ .

Let us apply this prescription to the simple problem of one-dimensional motion with a constant acceleration  $a$ , for which the Hamiltonian is

$$H = \frac{p^2}{2m} - max,$$

with  $u$  as the position coordinate  $x$ . The Poisson brackets needed in Eq. (9.119) are easy to evaluate directly or from the fundamental brackets:

$$[x, H] = \frac{p}{m},$$

$$[[x, H], H] = \frac{1}{m}[p, H] = a.$$

Because this last Poisson bracket is a constant, all higher-order brackets vanish identically and the series terminates, with the complete solution being given by

$$x = x_0 + \frac{p_0 t}{m} + \frac{at^2}{2}.$$

Remembering that  $p_0/m = v_0$ , this will be recognized as the familiar elementary solution to the problem.

It may be felt that what we have done here is a *tour de force*, a mere virtuoso performance. There is force to the objection. We would not propose the formal series solution, Eq. (9.119), as the preferred method for solving realistic problems in mechanics. It is surely one of the most recondite procedures we can conceive of for solving the easiest of freshman physics problems! Nonetheless, the technique provides insights into the structure of classical mechanics as based on canonical transformation theory. The series expansion shows directly that infinitesimal

canonical transformations can generate finite canonical transformations, depending on a parameter, and thus lead to solutions to the equations of motion. Of particular interest for the relation between classical and quantum mechanics is the observation that the series in Eqs. (9.117) or (9.119) bear a family resemblance to the series for an exponential. The nest of Poisson brackets in the  $n$ th term can be considered as the  $n$ th repeated application (from the right!) of the operator  $[ \ , G ]$ , or the  $n$ th power of the operator. Equation (9.119), for example, could symbolically be written as

$$u(t) = ue^{\hat{H}t} \Big|_0. \quad (9.120)$$

The exponential here means no more than its series representations and the symbol  $\hat{H}$  is used to indicate the operator  $[ \ , H ]$ . What we have here is very reminiscent of the Heisenberg picture in quantum mechanics where the  $u(t)$  become time-varying operators, whose time dependence is given in terms of  $\exp[iHt/\hbar]$  in such a manner as to lead to the same equation of motion, Eq. (9.94). (The additional factor  $i/\hbar$  arises out of the correspondence between the classical Poisson bracket and the quantum commutator.) The Poisson bracket formulation of mechanics is thus the classical analog of the Heisenberg picture of quantum mechanics.

## 9.7 ■ THE ANGULAR MOMENTUM POISSON BRACKET RELATIONS

The identification of the canonical angular momentum as the generator of a rigid rotation of the system leads to a number of interesting and important Poisson bracket relations. Equations (9.103) for the change of a function  $u$  under an infinitesimal canonical transformation (on the “active” view) is also valid if  $u$  is taken as the component of a vector along a *fixed* axis in ordinary space. Thus, if  $\mathbf{F}$  is a vector function of the system configuration, then (cf. Eq. (9.116))

$$\partial F_i = d\alpha[F_i, G].$$

Note that the direction along which the component is taken must be fixed, that is, not affected by the canonical transformation. If the direction itself is determined in terms of the system variables, then the transformation changes not only the value of the function but the nature of the function, just as with the Hamiltonian. With this understanding the change in a vector  $\mathbf{F}$  under a rotation of the system about a fixed axis  $\mathbf{n}$ , generated by  $\mathbf{L} \cdot \mathbf{n}$ , can be written in vector notation (cf. Eq. (9.115))

$$\partial \mathbf{F} = d\theta[\mathbf{F}, \mathbf{L} \cdot \mathbf{n}]. \quad (9.121)$$

To put it in other words, Eq. (9.121) implies that the unit vectors  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  that form the basis set for  $\mathbf{F}$  are not themselves rotated by  $\mathbf{L} \cdot \mathbf{n}$ .

The words describing what is meant by Eq. (9.121) must be chosen carefully for another reason. What is spoken of is the rotation of the system under the I.C.T., not necessarily the rotation of the vector  $\mathbf{F}$ . The generator  $\mathbf{L} \cdot \mathbf{n}$  induces a spatial rotation of the system variables, not for example of some external vector such as a magnetic field or the vector of the acceleration of gravity. Under what conditions then does  $\mathbf{L} \cdot \mathbf{n}$  generate a spatial rotation of  $\mathbf{F}$ ? The answer is clear—when  $\mathbf{F}$  is a function only of the system variables ( $q, p$ ) and does not involve any external quantities or vectors not affected by the I.C.T. Only under these conditions does a spatial rotation imply a corresponding rotation of  $\mathbf{F}$ . We shall designate such vectors as *system vectors*. The change in a vector under infinitesimal rotation about an axis  $\mathbf{n}$  has been given several times before (cf. Eq. (2.50) and Eq. (4.75)):

$$d\mathbf{F} = \mathbf{n} d\theta \times \mathbf{F}.$$

For a system vector  $\mathbf{F}$ , the change induced under an I.C.T. generated by  $\mathbf{L} \cdot \mathbf{n}$  can therefore be written as

$$\partial\mathbf{F} = d\theta[\mathbf{F}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} d\theta \times \mathbf{F}. \quad (9.122)$$

Equation (9.122) implies an important Poisson bracket identity obeyed by all system vectors:

$$[\mathbf{F}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} \times \mathbf{F}. \quad (9.123)$$

Note that in Eq. (9.123) there is no longer any reference to a canonical transformation or even to a spatial rotation. It is simply a statement about the value of certain Poisson brackets for a specific class of vectors and, as such, can be verified by direct evaluation in any given case. Suppose, for example, we had a system of an unconstrained particle and used the Cartesian coordinates as the canonical space coordinates. Then the Cartesian vector  $\mathbf{p}$  is certainly a suitable system vector. If  $\mathbf{n}$  is taken as a unit vector in the  $z$  direction, then by direct evaluation we have

$$\begin{aligned} [p_x, xp_y - yp_x] &= -p_y, \\ [p_y, xp_y - yp_x] &= p_x, \\ [p_z, xp_y - yp_x] &= 0. \end{aligned}$$

The right-hand sides of these identities is clearly the same as the components of  $\mathbf{n} \times \mathbf{p}$ , as predicted by Eq. (9.123).

On the other hand, suppose that in the same problem we tried to use for  $\mathbf{F}$  the vector  $\mathbf{A} = \frac{1}{2}(\mathbf{r} \times \mathbf{B})$  where  $\mathbf{B} = Bi$  is a fixed vector along the  $x$  axis. The vector  $\mathbf{A}$  will be recognized as the vector potential corresponding to a uniform magnetic field  $\mathbf{B}$  in the  $x$ -direction. As  $\mathbf{A}$  depends upon a vector external to the system, we would expect it not to fit the characteristics of a system vector and Eq. (9.123) should not hold for it. Indeed, we see that the Poisson brackets involved are here

$$\begin{aligned}[0, xp_y - yp_x] &= 0, \\ \left[ \frac{1}{2}zB, xp_y - yp_x \right] &= 0, \\ \left[ -\frac{1}{2}yB, xp_y - yp_x \right] &= -\frac{1}{2}Bx,\end{aligned}$$

whereas the vector  $\mathbf{n} \times \mathbf{A}$  has instead the components  $(-\frac{1}{2}Bz, 0, 0)$ .

The relation (9.123) may be expressed in various notations. Perhaps the most advantageous is a form using the Levi-Civita density to express the cross product (cf. Eq. (4.77')). The  $i$ th component of Eq. (9.123) for arbitrary  $\mathbf{n}$  then can be written

$$[F_i, L_j n_j] = \epsilon_{ijk} n_j F_k, \quad (9.124)$$

which implies the simple result

$$[F_i, L_j] = \epsilon_{ijk} F_k. \quad (9.125)$$

An alternative statement of Eq. (9.125) is to note that if  $l, m, n$  are three indices in cyclic order, then

$$[F_l, L_m] = F_n, \quad l, m, n \text{ in cyclic order.} \quad (9.125')$$

Another consequence of Eq. (9.123) relates to the dot product of two system vectors:  $\mathbf{F} \cdot \mathbf{G}$ . Being a scalar, such a dot product should be invariant under rotation, and indeed the Poisson bracket of the dot product with  $\mathbf{L} \cdot \mathbf{n}$  is easily shown to vanish:

$$\begin{aligned} [\mathbf{F} \cdot \mathbf{G}, \mathbf{L} \cdot \mathbf{n}] &= \mathbf{F} \cdot [\mathbf{G}, \mathbf{L} \cdot \mathbf{n}] + \mathbf{G} \cdot [\mathbf{F}, \mathbf{L} \cdot \mathbf{n}] \\ &= \mathbf{F} \cdot \mathbf{n} \times \mathbf{G} + \mathbf{G} \cdot \mathbf{n} \times \mathbf{F} \\ &= \mathbf{F} \cdot \mathbf{n} \times \mathbf{G} + \mathbf{F} \cdot \mathbf{G} \times \mathbf{n} \\ &= 0.\end{aligned} \quad (9.126)$$

The magnitude of any system vector therefore has a vanishing Poisson bracket with any component of  $\mathbf{L}$ .

Perhaps the most frequent application of these results arises from taking  $F$  to be the vector  $\mathbf{L}$  itself. We then have

$$[\mathbf{L}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} \times \mathbf{L}, \quad (9.127)$$

$$[L_i, L_j] = \epsilon_{ijk} L_k, \quad (9.128)$$

and

$$[L^2, \mathbf{L} \cdot \mathbf{n}] = 0. \quad (9.129)$$

A number of interesting consequences follow from Eqs. (9.127)

$$[p, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} \times \mathbf{p}$$

$$[p_i, L_j] = \epsilon_{ijk} p_k.$$

If  $L_x$  and  $L_y$  are constants of the motion, Poisson's theorem then states that  $[L_x, L_y] = L_z$  is also a constant of the motion. Thus, if any two components of the angular momentum are constant, the total angular momentum vector is conserved. As a further instance, let us assume that in addition to  $L_x$  and  $L_y$  being conserved there is a Cartesian vector of canonical momentum  $\mathbf{p}$  with  $p_z$  a constant of the motion. Not only is  $L_z$  conserved but we have two further constants of the motion:

$$[p_z, L_x] = p_y$$

and

$$[p_z, L_y] = -p_x,$$

that is, both  $\mathbf{L}$  and  $\mathbf{p}$  are conserved. We have here an instance in which Poisson's theorem does yield new constants of the motion. Note, however, that if  $p_x$ ,  $p_y$ , and  $L_z$  were the given constants of the motion, then their Poisson brackets are

$$[p_x, p_y] = 0,$$

$$[p_x, L_z] = -p_y,$$

$$[p_y, L_z] = p_x.$$

Here no new constants can be obtained from Poisson's theorem.

Recall from the fundamental Poisson brackets, Eqs. (9.69), that the Poisson bracket of any two canonical momenta must always be zero. But, from Eq. (9.128),  $L_i$  does not have a vanishing Poisson bracket with any of the other components of  $\mathbf{L}$ . Thus, while we have described  $\mathbf{L}$  as the total canonical angular momentum by virtue of its definition as  $\mathbf{r}_i \times \mathbf{p}_i$  (summed over all particles), no two components of  $\mathbf{L}$  can simultaneously be canonical variables. However, Eq. (9.129) shows that any one of the components of  $\mathbf{L}$ , and its magnitude  $L$ , can be chosen to be canonical variables at the same time.\*

\*It has been remarked previously that the correspondence between quantum and classical mechanics is such that the quantum mechanical commutator goes over essentially into the classical Poisson bracket as  $\hbar \rightarrow 0$ . Much of the formal structure of quantum mechanics appears as a close copy of the Poisson bracket formulation of classical mechanics. All the results of this section therefore have close quantum analogs. For example, the fact that two components of  $\mathbf{L}$  cannot be simultaneous canonical momenta appears as the well-known statement that  $L_i$  and  $L_j$  cannot have simultaneous eigenvalues. But  $L^2$  and any  $L_i$  can be quantized together. Indeed, most of these relations are known far better in their quantum form than as classical theorems.

## 9.8 ■ SYMMETRY GROUPS OF MECHANICAL SYSTEMS

It has already been pointed out that canonical transformations form a group. Canonical transformations that are analytic functions of continuous parameters form groups that are Lie groups. A Lie group with continuous parameters,  $\theta_i$ , has associated with it a flat vector space whose basis vectors,  $u_i$ , constitute a Lie algebra satisfying the previously given condition on the Poisson bracket

$$[u_i, u_j] = \sum_k c_{ij}^k u_k. \quad (9.77)$$

The elements,  $Q(\theta_i)$ , of the associated Lie group are related to the elements of the Lie algebra by

$$Q(\theta_i) = \exp\left(\frac{i}{2} \sum \theta_i u_i\right). \quad (9.130)$$

The definitions of Lie groups and Lie algebras are considered in more detail in Appendix B.

In Chapter 4 of the first two editions of this text, an extensive discussion was given of the Pauli matrix representation of the rotational group in three dimensions where the Pauli matrices that form the basis,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are both hermitian (the matrix is equal to its own transpose complex conjugate) and unitary (the transpose complex conjugate of the matrix is the inverse). These matrices have the properties\*

$$[\sigma_i, \sigma_j] = 2i\sigma_k$$

for  $i$ ,  $j$ , and  $k$  a cyclic permutation of  $x$ ,  $y$ , and  $z$ . The structure constants are thus  $c_{ij}^k = 2i\epsilon_{ijk}$  and  $\sigma_i^2 = 1$ , the unit  $2 \times 2$  matrix. The Euler angles can be used as the parameters that generate the group elements. For a rotation in the  $y$ - $z$  plane we have, for example,

$$Q(\theta) = 1 \cos \frac{\theta}{2} + i\sigma_x \sin \frac{\theta}{2} = \begin{pmatrix} \cos \frac{\theta}{2} & i \sin \frac{\theta}{2} \\ i \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix}.$$

\*Some physicists define a Lie algebra with the expression  $[u_l, u_j] = i \sum_k c_{lj}^k u_k$  instead of Eq. (9.77). This makes the structure constants in the following discussion real. Many mathematicians omit the  $i = \sqrt{-1}$  in the definition. The present text follows the latter convention.

In this formalism, vectors are represented by  $2 \times 2$  matrices of the form

$$V_{(x,y,z)} = \begin{pmatrix} V_z & V_x - iV_y \\ V_x + iV_y & -V_z \end{pmatrix},$$

and a rotation is performed by a similarity transformation

$$V_{(x',y',z')} = Q(\theta) V_{(x,y,z)} Q^\dagger(\theta),$$

where  $Q^\dagger$  is the adjoint, or complex conjugate transpose of the matrix  $Q$ .

The  $2 \times 2$  matrices  $Q$  are unitary with determinant +1, so they constitute a representation of the special unitary group in two dimensions, SU(2). The set of unitary  $2 \times 2$  matrices with determinant either +1 or -1 has twice as many elements (both infinite in number), which form the full unitary group U(2) in two dimensions. This group of  $2 \times 2$  rotation matrices has the same properties as the group of the associated infinitesimal canonical transformations (I.C.T.). It is customary to work primarily with the I.C.T.'s as they are easier to handle. The Lie groups of I.C.T.'s whose generators are the constants of the motion of the system are known as the *symmetry groups* of the system for, as we have seen, such transformations leave the Hamiltonian invariant. Finding the symmetry groups of a system goes a long way toward solving the problem of its classical motion and is even closer to a solution of the quantum-mechanical problem.

A system with spherical symmetry is invariant under rotation about any axis, so it can be represented by the group SU(2) as discussed above. Of more practical use is the set of the usual  $3 \times 3$  rotation matrices with determinant +1, which represent the special rotation group in three dimensions  $R(3) \equiv SO(3)$ . The vector  $\mathbf{L}$  is conserved in such a system in accord with our identification of the components of  $\mathbf{L}$  as the generators of spatial rotations. For the group of transformations generated by  $L_i$ , Eq. (9.128) shows that the structure constants are  $c_{ij}^k = \epsilon_{ijk}$ , and it is this relationship that stamps the group as being the rotation group in three dimensions. Thus, the matrix generators  $\mathbf{M}_i$  of infinitesimal rotations, Eqs. (4.79), have been seen to obey the commutation relations, Eq. (4.80),

$$[\mathbf{M}_i, \mathbf{M}_j] = \epsilon_{ijk} \mathbf{M}_k, \quad (4.80)$$

that is, with the same structure constants as for  $L_i$ . The quantities  $L_i$  and  $\mathbf{M}_i$  are different physically; the brackets in Eqs. (9.125) and (4.80) refer to different operations (although they share the same significant algebraic properties). But the identity of the structure constants for  $L_i$  and  $\mathbf{M}_i$  (cf. Eqs. (9.128) and (4.80)) shows that they have the same group structure, that of SO(3).

For the bound Kepler problem, we have seen (Section 3.9) that there exists in addition to  $\mathbf{L}$  another conserved vector quantity,  $\mathbf{A}$ , the Laplace-Runge-Lenz vector defined by Eq. (3.82)

$$\mathbf{A} = \mathbf{p} \times \mathbf{L} - \frac{mkr}{r}. \quad (3.82)$$

The Poisson bracket relations of the components of  $\mathbf{A}$  with themselves and with the components of  $\mathbf{L}$  can be obtained in a straightforward manner. Since  $\mathbf{A}$  clearly qualifies as a system vector, we immediately have the bracket relations

$$[A_i, L_j] = \epsilon_{ijk} A_k. \quad (9.131)$$

The Poisson brackets of the components of  $\mathbf{A}$  among themselves cannot be obtained by any such simple stratagem, but after a fair amount of tedious manipulation it is found that\*

$$[A_1, A_2] = - \left( p^2 - \frac{2mk}{r} \right) L_3. \quad (9.132)$$

The quantity on the right in the parentheses will be recognized as  $2mH$ , which has the conserved value  $2mE$ . If we therefore introduce a new constant vector  $\mathbf{D}$  defined as

$$\mathbf{D} = \frac{\mathbf{A}}{\sqrt{-2mE}} \equiv \frac{\mathbf{A}}{\sqrt{2m|E|}} \quad (9.133)$$

(note that  $E$  is negative for bound motion!), then the components of  $\mathbf{D}$  satisfy the Poisson bracket relation

$$[D_1, D_2] = L_3.$$

By cyclically permuting the indices, the complete set of Poisson brackets follows immediately. Thus, the components of  $\mathbf{L}$  and  $\mathbf{D}$  together form a Lie algebra for the bound Kepler problem, with structure constants to be obtained from the identities.

$$[L_i, L_j] = \epsilon_{ijk} L_k, \quad (9.128)$$

$$[D_i, L_j] = \epsilon_{ijk} D_k, \quad (9.134)$$

and

$$[D_i, D_j] = \epsilon_{ijk} L_k. \quad (9.135)$$

An examination of the fundamental matrices for rotation will show that the symmetry group for the bound Kepler problem is to be identified with the group of four-dimensional real proper rotations, called the special orthogonal group of dimension 4, which is usually designated as  $SO(4)$  or  $R(4)$ . Such a transformation preserves the value of the scalar quadratic form  $x_\mu x_\mu$ , where all the  $x_\mu$  are real. An orthogonal transformation in four dimensions has 10 conditions on the 16 el-

\*Some reduction in the length of the derivation is obtained by identifying  $\mathbf{p} \times \mathbf{L}$  as a system vector  $\mathbf{C}$ , and first evaluating the Poisson brackets  $[C_1, (\mathbf{p} \times \mathbf{L})_2]$  and  $[C_1, \mathbf{r}/r]$  making use of the fundamental Poisson brackets and Eqs. (9.125) to the utmost.

ements of the matrix with determinant  $\pm 1$ , so only 6 are independent. By looking on the infinitesimal transformation as being made up of a sequence of rotations in the various planes, we can easily obtain the corresponding six generators. Three of them are rotations in the three distinct  $x_i-x_j$  planes and so correspond to the  $\mathbf{M}_i$  generators of Eqs. (4.79), except that there are added zeros in the zeroth row and column. The remaining three generate infinitesimal rotations in the  $x_0-x_1$  planes. Thus, the generator matrix for an infinitesimal rotation in the  $x_0-x_1$  plane would be

$$\mathbf{N}_1 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (9.136)$$

with  $\mathbf{N}_2$  and  $\mathbf{N}_3$  given in corresponding fashion. Direct matrix multiplication shows that these six matrices satisfy the commutator (or Lie bracket) relations

$$\begin{aligned} [\mathbf{M}_i, \mathbf{M}_j] &= \epsilon_{ijk} \mathbf{M}_k, \\ [\mathbf{N}_i, \mathbf{M}_j] &= \epsilon_{ijk} \mathbf{N}_k, \\ [\mathbf{N}_i, \mathbf{N}_j] &= \epsilon_{ijk} \mathbf{M}_k, \end{aligned}$$

with structure constants  $c_{ij}^k = \epsilon_{ijk}$ . Since these are the same as the Poisson bracket relations, Eqs. (9.128), (9.134), and (9.135), the identification of the symmetry group of the bound Kepler problem with  $R(4)$  is thus proven.

Note that for the Kepler problem with positive energy (that is, scattering)  $\mathbf{A}$  is still a constant of the motion,\* but the appropriate reduced real vector, instead of  $\mathbf{D}$ , is  $\mathbf{C}$  defined as

$$\mathbf{C} = \frac{\mathbf{A}}{\sqrt{2mE}}, \quad (9.137)$$

and the Poisson bracket relations for  $\mathbf{L}$  and  $\mathbf{C}$  are now

$$\begin{aligned} [L_i, L_j] &= \epsilon_{ijk} L_k, \\ [C_i, L_j] &= \epsilon_{ijk} C_k, \\ [C_i, C_j] &= -\epsilon_{ijk} L_k. \end{aligned} \quad (9.138)$$

These structure constants are the same as for the restricted Lorentz group, which must therefore be the symmetry group for the positive energy Kepler problem—in nonrelativistic mechanics. We must not read any kinship of physical ideas into this happenstance. The Kepler problem does *not* contain in it the seed of the basic conceptions of special relativity; it is purely a problem of nonrelativistic Newtonian mechanics. That the symmetry group may involve a space of higher dimension than ordinary space is connected with the fact that the symmetry we seek here

\*The arguments of Section 3.9 are independent of the sign of either  $E$  or the force constant  $k$ .

is one in the six-dimensional phase space. The symmetry group consists of the canonical transformations in this space that leave the Hamiltonian unchanged. It should not be surprising therefore that the group can be interpreted in terms of transformations of spaces of more than three dimensions.

The two-dimensional isotropic harmonic oscillator is another mechanical system for which a symmetry group is easily identified. In Cartesian coordinates, the Hamiltonian for this system may be written as

$$H = \frac{1}{2m}(p_x^2 + m^2\omega^2x^2) + \frac{1}{2m}(p_y^2 + m^2\omega^2y^2). \quad (9.139)$$

As it doesn't depend on time explicitly, the Hamiltonian is constant and is equal to the total energy of the system. The  $z$  axis is an axis of symmetry for the system, and hence the angular momentum along that axis (which is in fact the total angular momentum) is also a constant of motion:

$$L = xp_y - yp_x. \quad (9.140)$$

Further constants of the motion exist for this problem that can be written as components of a symmetrical two-dimensional tensor  $\mathbf{A}$  defined as

$$A_{ij} = \frac{1}{2m}(p_i p_j + m^2\omega^2 x_i x_j). \quad (9.141)$$

Of the three distinct elements of the tensor, the diagonal terms may be identified as the energies associated with the separate one-dimensional motions along the  $x$  and  $y$  axes, respectively. Physically, as there is no coupling between the two motions, the two energies must separately be constant. A little more formally, it is obvious from the way in which  $H$  has been written in Eq. (9.139) that  $A_{11}$  and  $A_{22}$  each have a vanishing Poisson bracket with  $H$ . The off-diagonal element of  $\mathbf{A}$ ,

$$A_{12} = A_{21} = \frac{1}{2m}(p_x p_y + m^2\omega^2 x y), \quad (9.142)$$

is a little more difficult to recognize. That it is a constant of the motion may easily be seen by evaluating the Poisson bracket with  $H$ . In relation to the separate  $x$  and  $y$  motions,  $A_{11}$  and  $A_{22}$  are related to the amplitudes of the oscillations, whereas  $A_{12}$  is determined by the phase difference between the two vibrations. Thus, the solutions for the motion can be written as

$$x = \sqrt{\frac{2A_{11}}{m\omega^2}} \sin(\omega t + \theta_1),$$

$$y = \sqrt{\frac{2A_{22}}{m\omega^2}} \sin(\omega t + \theta_2),$$

and it then follows from Eq. (9.142) that

$$A_{12} = \sqrt{A_{11}A_{22}} \cos(\theta_2 - \theta_1). \quad (9.143)$$

The trace of the  $\mathbf{A}$  tensor is the total energy of the harmonic oscillator. Out of the elements of the matrix, we can form two other distinct constants of the motion, which it is convenient to write in the form

$$S_1 = \frac{A_{12} + A_{21}}{2\omega} = \frac{1}{2m\omega}(p_x p_y + m^2\omega^2 xy), \quad (9.144)$$

$$S_2 = \frac{A_{22} - A_{11}}{2\omega} = \frac{1}{4m\omega} [p_y^2 - p_x^2 + m^2\omega^2(y^2 - x^2)]. \quad (9.145)$$

To these we may add a third constant of the motion from Eq. (9.140):

$$S_3 = \frac{L}{2} = \frac{1}{2}(xp_y - yp_x). \quad (9.146)$$

The quantities  $S_i$  plus the total energy  $H$  form four algebraic constants of the motion not involving time explicitly. It is clear that not all of them can be independent, because in a system of two degrees of freedom there can at most be only three such constants. We know that the orbit for the isotropic harmonic oscillator is an ellipse and three constants of the motion are needed to describe the parameters of the orbit in the plane—say, the semimajor axis, the eccentricity, and the orientation of the ellipse. The fourth constant of motion relates to the passage of the particle through a specific point at a given time and would therefore be explicitly time dependent. Hence, there must exist a single relation connecting  $S_i$  and  $H$ . By direct evaluation it is easy to show that\*

$$S_1^2 + S_2^2 + S_3^2 = \frac{H^2}{4\omega^2}. \quad (9.147)$$

By straight forward manipulation of the Poisson brackets, we can verify that the three  $S_i$  quantities satisfy the relations

$$[S_i, S_j] = \epsilon_{ijk} S_k. \quad (9.148)$$

These are the same relations as for the three-dimensional angular momentum vector, or for the generators of rotation in a three-dimensional space. The group of transformations generated by  $S_i$  may therefore be identified with R(3) or SO(3). Actually, there is some ambiguity in the identification.

\*An equivalent form of the condition Eq. (9.147) is that the determinant of  $\mathbf{A}$  is  $L^2\omega^2/4$ . It will be recalled that similarly in the case of the Kepler problem, the components of the new vector constant of motion  $\mathbf{A}$  were not all independent of the other constants of the motion. There exist indeed two relations linking  $\mathbf{A}$ ,  $\mathbf{L}$ , and  $H$ , Eqs. (3.83) and (3.87).

There is a homomorphism (in this case, a 2 to 1 mapping) between the orthogonal unimodular group  $\text{SO}(3)$  also called the rotation group  $R(3)$  in three dimensions and the unitary unimodular group\*  $SU(2)$  in two dimensions. It turns out that  $SU(2)$  is here more appropriate. To glimpse at the circumstances justifying this choice, note that Eq. (9.147) suggests there is a three-dimensional space, each point of which corresponds to a particular set of orbital parameters. For a given system energy, Eq. (9.147) says the orbit “points” in this space lie on a sphere. The constants  $S_i$  generate three-dimensional rotations on this sphere; that is, they change one orbit into another orbit having the same energy. It may be shown that  $S_1$  generates a transformation that changes the eccentricity of the orbit and that for any given final eccentricity we can find *two* transformations leading to it. It is this double-valued quality of the transformation that indicates  $SU(2)$  rather than  $\text{SO}(3)$  is the correct symmetry group for the two-dimensional harmonic oscillator.

For higher dimensions, the structure constants of the Lie algebras of the  $\text{SO}(n)$  rotation groups and the  $SU(n)$  unitary groups are no longer identical, and a clear-cut separation between the two can be made. For the three-dimensional isotropic harmonic oscillator, there is again a tensor constant of the motion defined by Eq. (9.141), except that the indices now run from 1 to 3. The distinct components of this tensor, together with the components of  $\mathbf{L}$  now satisfy Poisson bracket relations with the rather complicated structure constants that belong to  $SU(3)$ . Indeed, it is possible to show that for the  $n$ -dimensional isotropic harmonic oscillator the symmetry group is  $SU(n)$ .

It has previously been pointed out in Section 3.9 that there exists a connection between the existence of additional algebraic constants of the motion—and therefore of higher-symmetry groups—and degeneracy in the motions of the system. In the case of the Kepler and isotropic harmonic oscillator problems, the additional constants of the motion are related to parameters of the orbit. Unless the orbit is closed, that is, the motion is confined to a single curve, we can hardly talk of such orbital parameters. Only when the various components of the motion have commensurate periods will the orbit be closed. The classic example is the two-dimensional anisotropic oscillator. When the frequencies in the  $x$  and  $y$  directions are rational fractions of each other, the particle traverses a closed Lissajous figure. But if the frequencies are incommensurate, the motion of the particle is space-filling or ergodic, eventually coming as close as desired to any specific point in the rectangle defined by the energies of motion in the two directions (ergotic hypothesis). Attempts at finding complicated (and perhaps complex) symmetry groups for incommensurate systems, applicable to all problems of the same number of degrees of freedom, have not yet proved fruitful. We shall have occasion in Section 13.7 to consider further the relation between symmetry and invariance when we discuss Noether’s theorem which gives a formal proof of the relation between invariance and conserved quantities.

\*A matrix is unitary if its inverse is its transpose complex conjugate, and a unimodular matrix is one whose determinant is +1.

## 9.9 ■ LIOUVILLE'S THEOREM

As a final application of the Poisson bracket formalism, we shall briefly discuss a fundamental theorem of statistical mechanics known as Liouville's theorem. While the exact motion of any system is completely determined in classical mechanics by the initial conditions, it is often impracticable to calculate an exact solution for complex systems. It would be obviously hopeless, for example, to calculate completely the motion of some  $10^{23}$  molecules in a volume of gas. In addition, the initial conditions are often only incompletely known. We may be able to state that at time  $t_0$  a given mass of gas has a certain energy, but we cannot determine the initial coordinates and velocities of each molecule. Statistical mechanics therefore makes no attempt to obtain a complete solution for systems containing many particles. Its aim, instead, is to make predictions about certain average properties by examining the motion of a large number of identical systems. The values of the desired quantities are then computed by forming averages over all the systems in the *ensemble*. All the members of the ensemble are as like the actual systems as our imperfect knowledge permits, but they may have any of the initial conditions that are consistent with this incomplete information. Since each system is represented by a single point in phase space, the ensemble of systems corresponds to a swarm of points in phase space. Liouville's theorem states that the density of systems in the neighborhood of some given system in phase space remains constant in time.

The density,  $D$ , as defined above can vary with time through two separate mechanisms. Since it is the density in the neighborhood of a given system point, there will be an *implicit* dependence as the coordinates of the system ( $q_i, p_i$ ) vary with time, and the system point wanders through phase space. There may also be an explicit dependence upon time. The density may still vary with time even when evaluated at a fixed point in phase space. By Eq. (9.94), the total time derivative of  $D$ , due to both types of variation with time, can be written as

$$\frac{dD}{dt} = [D, H] + \frac{\partial D}{\partial t}, \quad (9.149)$$

where the Poisson bracket arises from the implicit dependence, and the last term from the explicit dependence.

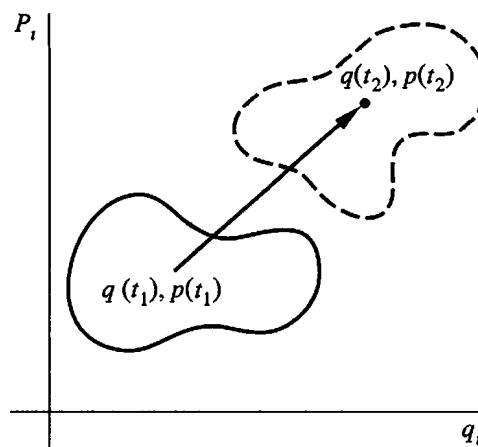
The ensemble of system points moving through phase space behaves much like a fluid in a multidimensional space, and there are numerous similarities between our discussion of the ensemble and the well-known notions of fluid dynamics. In Eq. (9.149), the total derivative is a derivative of the density as we follow the motion of a particular bit of the ensemble "fluid" in time. It is sometimes referred to as the *material* or *hydrodynamic* derivative. On the other hand, the partial derivative is at fixed ( $q, p$ ); it is as if we station ourselves at a particular spot in phase space and measure the time variation of the density as the ensemble of system points flows by us. These two derivatives correspond to two viewpoints frequently used in considering fluid flow. The partial derivative at a fixed point in phase space is in line with the Eulerian viewpoint that looks on the coordinates solely as iden-

tifying a point in space. The total derivative fits in with the Lagrangian picture in which individual particles are followed in time; the coordinates in effect rather identify a particle than a point in space. Basically, our consideration of phase space has been more like the Lagrangian viewpoint; the collection of quantities  $(q, p)$  identifies a system and its changing configuration with time.

Consider an infinitesimal volume in phase space surrounding a given system point, with the boundary of the volume formed by some surface of neighboring system points at the time  $t = 0$ . Note that the surface of the volume is one-dimension less than the volume. In the course of time, the system points defining the volume move about in phase space, and the volume contained by them will take on different shapes as time progresses. The dashed curve in Fig. 9.4 indicates the evolution of the infinitesimal volume with time. It is clear that the number of systems within the volume remains constant, for a system initially inside can never get out. If some system point were to cross the border, it would occupy at some time the same position in phase space as one of the system points defining the boundary surface. Since the subsequent motion of a system is uniquely determined by its location in phase space at a particular time, the two systems would travel together from there on. Hence, the system can never leave the volume. By the same token, a system initially outside can never enter the volume.

It has been shown that on the active picture of a canonical transformation, the motion of a system point in time is simply the evolution of a canonical transformation generated by the Hamiltonian. The canonical variables  $(q, p)$  at time  $t_2$ , as shown in Fig. 9.4, are related to the variables at time  $t_1$  by a particular canonical transformation. The change in the infinitesimal volume element about the system point over the time interval is given by the same canonical transformation. Now, Poincaré's integral invariant, Eq. (9.86), says that a volume element in phase space is invariant under a canonical transformation. Therefore, the size of the volume element about the system point cannot vary with time.

Thus, both the number of systems in the infinitesimal region,  $dN$ , and the volume,  $dV$ , are constants, and consequently the density



**FIGURE 9.4** Motion of a volume in two-dimensional phase space.

$$D = \frac{dN}{dV}$$

must also be constant in time, that is,

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

which proves Liouville's theorem. An alternative statement of the theorem follows from Eq. (9.149) as

$$\frac{\partial D}{\partial t} = -[D, H]. \quad (9.150)$$

When the ensemble of systems is in statistical equilibrium, the number of systems in a given state must be constant in time, which is to say that the density of system points at a given spot in phase space does not change with time. The variation of  $D$  with time at a fixed point corresponds to the partial derivative with respect to  $t$ , which therefore must vanish in statistical equilibrium. By Eq. (9.150), it follows that the equilibrium condition can be expressed as

$$[D, H] = 0.$$

We can ensure equilibrium therefore by choosing the density  $D$  to be a function of those constants of the motion of the system not involving time explicitly, for then the Poisson bracket with  $H$  must vanish. Thus, for conservative systems  $D$  can be any function of the energy, and the equilibrium condition is automatically satisfied. The characteristics of the ensemble will be determined by the choice of function for  $D$ . As an example, one well-known ensemble, the *microcanonical* ensemble, occurs if  $D$  is constant for systems having a given narrow energy range and zero outside the range.

The considerations have been presented here to illustrate the usefulness of the Poisson bracket formulation in classical statistical mechanics. Further discussion of these points would carry us far outside our field.

## DERIVATIONS

- One of the attempts at combining the two sets of Hamilton's equations into one tries to take  $q$  and  $p$  as forming a complex quantity. Show directly from Hamilton's equations of motion that for a system of one degree of freedom the transformation

$$Q = q + ip, \quad P = Q^*$$

is not canonical if the Hamiltonian is left unaltered. Can you find another set of coordinates  $Q'$ ,  $P'$  that are related to  $Q$ ,  $P$  by a change of scale only, and that are canonical?

2. Show that the transformation for a system of one degree of freedom,

$$Q = q \cos \alpha - p \sin \alpha,$$

$$P = q \sin \alpha + p \cos \alpha,$$

satisfies the symplectic condition for any value of the parameter  $\alpha$ . Find a generating function for the transformation. What is the physical significance of the transformation for  $\alpha = 0$ ? For  $\alpha = \pi/2$ ? Does your generating function work for both of these cases.

3. In Section 8.4 some of the problems of treating time as one of the canonical variables are discussed. If we are able to sidestep these difficulties, show that the equations of transformation in which  $t$  is considered a canonical variable reduce to Eqs. (9.14) if in fact the transformation does not affect the time scale.

4. Show directly that the transformation

$$Q = \log\left(\frac{1}{q} \sin p\right), \quad P = q \cot p$$

is canonical.

5. Show directly that for a system of one degree of freedom the transformation

$$Q = \arctan \frac{\alpha q}{p}, \quad P = \frac{\alpha q^2}{2} \left(1 + \frac{p^2}{\alpha^2 q^2}\right)$$

is canonical, where  $\alpha$  is an arbitrary constant of suitable dimensions.

6. The transformation equations between two sets of coordinates are

$$\begin{aligned} Q &= \log(1 + q^{1/2} \cos p), \\ P &= 2(1 + q^{1/2} \cos p)q^{1/2} \sin p. \end{aligned}$$

- (a) Show directly from these transformation equations that  $Q, P$  are canonical variables if  $q$  and  $p$  are.

- (b) Show that the function that generates this transformation is

$$F_3 = -(e^Q - 1)^2 \tan p.$$

7. (a) If each of the four types of generating functions exist for a given canonical transformation, use the Legendre transformation to derive relations between them.  
(b) Find a generating function of the  $F_4$  type for the identity transformation and of the  $F_3$  type for the exchange transformation.  
(c) For an orthogonal point transformation of  $q$  in a system of  $n$  degrees of freedom, show that the new momenta are likewise given by the orthogonal transformation of an  $n$ -dimensional vector whose components are the old momenta plus a gradient in configuration space.

8. Prove directly that the transformation

$$Q_1 = q_1, \quad P_1 = p_1 - 2p_2,$$

$$Q_2 = p_2, \quad P_2 = -2q_i - q_2$$

is canonical and find a generating function.

- 9. (a)** For a single particle show directly (that is, by direct evaluation of the Poisson brackets), that if  $u$  is a scalar function only of  $r^2$ ,  $p^2$ , and  $\mathbf{r} \cdot \mathbf{p}$ , then

$$[u, \mathbf{L}] = 0.$$

- (b)** Similarly show directly that if  $\mathbf{F}$  is a vector function,

$$\mathbf{F} = u\mathbf{r} + v\mathbf{p} + w(\mathbf{r} \times \mathbf{p}),$$

where  $u$ ,  $v$ , and  $w$  are scalar functions of the same type as in part (a), then

$$[F_i, L_j] = \epsilon_{ijk} F_k.$$

- 10.** Find under what conditions

$$Q = \frac{\alpha p}{x}, \quad P = \beta x^2,$$

where  $\alpha$  and  $\beta$  are constants, represents a canonical transformation for a system of one degree of freedom, and obtain a suitable generating function. Apply the transformation to the solution of the linear harmonic oscillator.

- 11.** Determine whether the transformation

$$Q_1 = q_1 q_2, \quad P_1 = \frac{p_1 - p_2}{q_2 - q_1} + 1,$$

$$Q_2 = q_1 + q_2, \quad P_2 = \frac{q_2 p_2 - q_1 p_1}{q_2 - q_1} - (q_2 + q_1)$$

is canonical.

- 12.** Show that the direct conditions for a canonical condition are given immediately by the symplectic condition expressed in the form

$$\mathbf{J}\mathbf{M} = \mathbf{M}^{-1}\mathbf{J}.$$

- 13.** The set of restricted canonical transformations has a group-property. Verify this statement once using the invariance of Hamilton's principle under canonical transformation (cf. Eq. (9.11)), and again using the symplectic condition.

- 14.** Prove that the transformation

$$Q_1 = q_1^2, \quad Q_2 = q_2 \sec p_2,$$

$$P_1 = \frac{p_1 \cos p_2 - 2q_2}{2q_1 \cos p_2}, \quad P_2 = \sin p_2 - 2q_1$$

is canonical, by any method you choose. Find a suitable generating function that will lead to this transformation.

- 15. (a)** Using the fundamental Poisson brackets find the values of  $\alpha$  and  $\beta$  for which the equations

$$Q = q^\alpha \cos \beta p, \quad P = q^\alpha \sin \beta p$$

represent a canonical transformation.

- (b)** For what values of  $\alpha$  and  $\beta$  do these equations represent an *extended* canonical transformation? Find a generating function of the  $F_3$  form for the transformation.  
**(c)** On the basis of part (b), can the transformation equations be modified so that they describe a canonical transformation for all values of  $\beta$ ?

- 16.** For a symmetric rigid body, obtain formulas for evaluating the Poisson brackets

$$[\dot{\phi}, f(\theta, \phi, \psi)], \quad [\dot{\psi}, f(\theta, \phi, \psi)]$$

where  $\theta$ ,  $\phi$ , and  $\psi$  are the Euler angles, and  $f$  is any arbitrary function of the Euler angles.

- 17.** Show that the Jacobi identity is satisfied if the Poisson bracket sign stands for the commutator of two square matrices:

$$[A, B] = AB - BA.$$

Show also that for the same representation of the Poisson bracket that

$$[A, BC] = [A, B]C + B[A, C].$$

- 18.** Prove Eq. (9.83) using the symplectic matrix notation for the Lagrange and Poisson brackets.

- 19.** Verify the analog of the Jacobi identity for Lagrange brackets,

$$\frac{\partial \{u, v\}}{\partial w} + \frac{\partial \{v, w\}}{\partial u} + \frac{\partial \{w, u\}}{\partial v} = 0,$$

where  $u$ ,  $v$ , and  $w$  are three functions in terms of which the  $(q, p)$  set can be specified.

- 20. (a)** Verify that the components of the two-dimensional matrix  $A$ , defined by Eq. (9.141), are constants of the motion for the two-dimensional isotropic harmonic oscillator problem.

- (b)** Verify that the quantities  $S_i$ ,  $i = 1, 2, 3$ , defined by Eqs. (9.144), (9.145), (9.146), have the properties stated in Eqs. (9.147) and (9.148).

## EXERCISES

- 21. (a)** For a one-dimensional system with the Hamiltonian

$$H = \frac{p^2}{2} - \frac{1}{2q^2},$$

show that there is a constant of the motion

$$D = \frac{pq}{2} - Ht.$$

- (b) As a generalization of part (a), for motion in a plane with the Hamiltonian

$$H = |\mathbf{p}|^n - ar^{-n},$$

where  $\mathbf{p}$  is the vector of the momenta conjugate to the Cartesian coordinates, show that there is a constant of the motion

$$D = \frac{\mathbf{p} \cdot \mathbf{r}}{n} - Ht.$$

- (c) The transformation  $Q = \lambda q$ ,  $p = \lambda P$  is obviously canonical. However, the same transformation with  $t$  time dilatation,  $Q = \lambda q$ ,  $p = \lambda P$ ,  $t' = \lambda^2 t$ , is not. Show that, however, the equations of motion for  $q$  and  $p$  for the Hamiltonian in part (a) are invariant under this transformation. The constant of the motion  $D$  is said to be associated with this invariance.

22. For the point transformation in a system of two degrees of freedom,

$$Q_1 = q_1^2, \quad Q_2 = q_1 + q_2,$$

find the most general transformation equations for  $P_1$  and  $P_2$  consistent with the overall transformation being canonical. Show that with a particular choice for  $P_1$  and  $P_2$  the Hamiltonian

$$H = \left( \frac{p_1 - p_2}{2q_1} \right)^2 + p_2 + (q_1 + q_2)^2$$

can be transformed to one in which both  $Q_1$  and  $Q_2$  are ignorable. By this means solve the problem and obtain expressions for  $q_1$ ,  $q_2$ ,  $p_1$ , and  $p_2$  as functions of time and their initial values.

23. By any method you choose, show that the following transformation is canonical:

$$x = \frac{1}{\alpha} (\sqrt{2P_1} \sin Q_1 + P_2), \quad p_x = \frac{\alpha}{2} (\sqrt{2P_1} \cos Q_1 - Q_2),$$

$$y = \frac{1}{\alpha} (\sqrt{2P_1} \cos Q_1 + Q_2), \quad p_y = -\frac{\alpha}{2} (\sqrt{2P_1} \sin Q_1 - P_2),$$

where  $\alpha$  is some fixed parameter.

Apply this transformation to the problem of a particle of charge  $q$  moving in a plane that is perpendicular to a constant magnetic field  $\mathbf{B}$ . Express the Hamiltonian for this problem in the  $(Q_i, P_i)$  coordinates letting the parameter  $\alpha$  take the form

$$\alpha^2 = \frac{qB}{c}.$$

From this Hamiltonian, obtain the motion of the particle as a function of time.

**24. (a)** Show that the transformation

$$Q = p + iaq, \quad P = \frac{p - iaq}{2ia}$$

is canonical and find a generating function.

**(b)** Use the transformation to solve the linear harmonic oscillator problem.

**25. (a)** The Hamiltonian for a system has the form

$$H = \frac{1}{2} \left( \frac{1}{q^2} + p^2 q^4 \right).$$

Find the equation of motion for  $q$ .

**(b)** Find a canonical transformation that reduces  $H$  to the form of a harmonic oscillator. Show that the solution for the transformed variables is such that the equation of motion found in part (a) is satisfied.

**26.** A system of  $n$  particles moves in a plane under the influence of interaction forces derived from potential terms depending only upon the scalar distances between particles.

**(a)** Using plane polar coordinates for each particle (relative to a common origin), identify the form of the Hamiltonian for the system.

**(b)** Find a generating function for the canonical transformation that corresponds to a transformation to coordinates rotating in the plane counterclockwise with a uniform angular rate  $\omega$  (the same for all particles). What are the transformation equations for the momenta?

**(c)** What is the new Hamiltonian? What physical significance can you give to the difference between the old and the new Hamiltonians?

**27. (a)** In the problem of small oscillations about steady motion, show that at the point of steady motion all the Hamiltonian variables  $\eta$  are constant. If the values for steady motion are  $\eta_0$  so that  $\eta = \eta_0 + \xi$ , show that to the lowest nonvanishing approximation the effective Hamiltonian for small oscillation can be expressed as

$$H(\eta_0, \xi) = \frac{1}{2} \xi S \xi,$$

where  $S$  is a square matrix with components that are functions of  $\eta_0$  only.

**(b)** Assuming all frequencies of small oscillation are distinct, let  $M$  be a square  $2n \times 2n$  matrix formed by the components of a possible set of eigenvectors (for both positive and negative frequencies). Only the directions of the eigenvectors are fixed, not their magnitudes. Show that it is possible to apply conditions to the eigenvectors (in effect fixing their magnitudes) that make  $M$  the Jacobian matrix of a canonical transformation.

**(c)** Show that the canonical transformation so found transforms the effective Hamiltonian to the form

$$H = i\omega_j q_j p_j,$$

where  $\omega_j$  is the magnitude of the normal frequencies. What are the equations of motion in this set of canonical coordinates?

(d) Finally, show that

$$F_2 = q_j P_j + \frac{i}{2} \frac{P_j^2}{\omega_j} - \frac{i}{4} \omega_j q_j^2$$

leads to a canonical transformation that decomposes  $H$  into the Hamiltonians for a set of uncoupled linear harmonic oscillators that oscillate in the normal modes.

28. A charged particle moves in space with a constant magnetic field  $\mathbf{B}$  such that the vector potential,  $\mathbf{A}$ , is

$$\mathbf{A} = \frac{1}{2}(\mathbf{B} \times \mathbf{r})$$

(a) If  $v_j$  are the Cartesian components of the velocity of the particle, evaluate the Poisson brackets

$$[v_i, v_j], \quad i \neq j = 1, 2, 3.$$

(b) If  $p_i$  is the canonical momentum conjugate to  $x_i$ , also evaluate the Poisson brackets

$$\begin{aligned} [x_i, v_j], \quad [p_i, v_j], \\ [x_i, \dot{p}_j], \quad [p_i, \dot{p}_j]. \end{aligned}$$

29. The semimajor axis  $a$  of the elliptical Kepler orbit and the eccentricity  $e$  are functions of first integrals of the motion, and therefore of the canonical variables. Similarly, the mean anomaly

$$\phi \equiv \omega(t - T) = \psi - e \sin \psi$$

is a function of  $r$ ,  $\theta$ , and the conjugate momenta. Here  $T$  is the time of periapsis passage and is a constant of the motion. Evaluate the Poisson brackets that can be formed of  $a$ ,  $e$ ,  $\phi$ ,  $\omega$ , and  $T$ . There are in fact only nine nonvanishing distinct Poisson brackets out of these quantities.

30. (a) Prove that the Poisson bracket of two constants of the motion is itself a constant of the motion even when the constants depend upon time explicitly.  
 (b) Show that if the Hamiltonian and a quantity  $F$  are constants of the motion, then the  $n$ th partial derivative of  $F$  with respect to  $t$  must also be a constant of the motion.  
 (c) As an illustration of this result, consider the uniform motion of a free particle of mass  $m$ . The Hamiltonian is certainly conserved, and there exists a constant of the motion

$$F = x - \frac{pt}{m}.$$

Show by direct computation that the partial derivative of  $F$  with  $t$ , which is a constant of the motion, agrees with  $[H, F]$ .

31. Show by the use of Poisson brackets that for a one-dimensional harmonic oscillator there is a constant of the motion  $u$  defined as

$$u(q, p, t) = \ln(p + im\omega q) - i\omega t, \quad \omega = \sqrt{\frac{k}{m}}.$$

What is the physical significance of this constant of the motion?

32. A system of two degrees of freedom is described by the Hamiltonian

$$H = q_1 p_1 - q_2 p_2 - aq_1^2 + bq_2^2.$$

Show that

$$F_1 = \frac{p_1 - aq_1}{q_2} \quad \text{and} \quad F_2 = q_1 q_2$$

are constants of the motion. Are there any other independent algebraic constants of the motion? Can any be constructed from Jacobi's identity?

33. Set up the magnetic monopole described in Exercise 28 (Chapter 3) in Hamiltonian formulation (you may want to use spherical polar coordinates). By means of the Poisson bracket formulation, show that the quantity  $D$  defined in that exercise is conserved.
34. Obtain the motion in time of a linear harmonic oscillator by means of the formal solution for the Poisson bracket version of the equation of motion as derived from Eq. (9.116). Assume that at time  $t = 0$  the initial values are  $x_0$  and  $p_0$ .
35. A particle moves in one dimension under a potential

$$V = \frac{mk}{x^2}.$$

Find  $x$  as a function of time, by using the symbolic solution of the Poisson bracket form for the equation of motion for the quantity  $y = x^2$ . Initial conditions are that at  $t = 0$ ,  $x = x_0$ , and  $v = 0$ .

36. (a) Using the theorem concerning Poisson brackets of vector functions and components of the angular momentum, show that if  $\mathbf{F}$  and  $\mathbf{G}$  are two vector functions of the coordinates and momenta only, then

$$[\mathbf{F} \cdot \mathbf{L}, \mathbf{G} \cdot \mathbf{L}] = \mathbf{L} \cdot (\mathbf{G} \times \mathbf{F}) + L_i L_j [F_i, G_j].$$

- (b) Let  $\mathbf{L}$  be the total angular momentum of a rigid body with one point fixed and let  $L_\mu$  be its component along a set of Cartesian axes fixed in the rigid body. By means of part (a) find a general expansion for

$$[L_\mu, L_\nu], \quad \mu, \nu = 1, 2, 3.$$

(Hint: Choose for  $\mathbf{F}$  and  $\mathbf{G}$  unit vectors along the  $\mu$  and  $\nu$  axes.)

- (c) From the Poisson bracket equations of motion for  $L_\mu$  derive Euler's equations of motion for a rigid body.

37. Set up the problem of the spherical pendulum in the Hamiltonian formulation, using spherical polar coordinates for the  $q_i$ . Evaluate directly in terms of these canonical variables the following Poisson brackets:

$$[L_x, L_y], \quad [L_y, L_z], \quad [L_z, L_x],$$

showing that they have the values predicted by Eq. (9.128). Why is it that  $p_\theta$  and  $p_\psi$  can be used as canonical momenta, although they are perpendicular components of the angular momentum?

- 38.** In Section 9.7, it is shown that if any two components of the angular momentum are conserved, then the total angular momentum is conserved. If two of the components are identically zero, the third must be conserved. From this it would appear to follow that in any motion confined to a plane, so that the components of the angular momentum in the plane are zero, the total angular momentum is constant. There appear to be a number of obvious contradictions to this prediction; for example, the angular momentum of an oscillating spring in a watch, or the angular momentum of a plane disk rolling down an inclined plane all in the same vertical plane. Discuss the force of these objections and whether the statement of the theorem requires any restrictions.

- 39. (a)** Show from the Poisson bracket condition for conserved quantities that the Laplace–Runge–Lenz vector  $\mathbf{A}$ ,

$$\mathbf{A} = \mathbf{p} \times \mathbf{L} - \frac{mkr}{r},$$

is a constant of the motion for the Kepler problem.

- (b)** Verify the Poisson bracket relations for the components of  $\mathbf{A}$  as given by Eq. (9.131).

- 40.** Consider a system that consists of a rigid body in three-space with one point fixed. Using cylindrical coordinates find the canonical transformation corresponding to new axes rotating about the  $z$ -axis with an arbitrary time-dependent angular velocity. Verify that your proposed solution is canonical.
- 41.** We start with a time independent Hamiltonian  $H_o(q, p)$  and impose an external oscillating field making the Hamiltonian

$$H = H_o(q, p) - \varepsilon \sin \omega t$$

where  $\varepsilon$  and  $\omega$  are given constants.

- (a)** How are the canonical equations modified?  
**(b)** Find a canonical transformation that restores the canonical form of the equations of motion and determine the “new” Hamiltonian.  
**(c)** Give a possible physical interpretation of the imposed field.

# CHAPTER

# 10

## Hamilton–Jacobi Theory and Action-Angle Variables

It has already been mentioned that canonical transformations may be used to provide a general procedure for solving mechanical problems. Two methods have been suggested. If the Hamiltonian is conserved, then a solution could be obtained by transforming to new canonical coordinates that are all cyclic, thereby providing new equations of motion with trivial solutions. An alternative technique is to seek a canonical transformation from the coordinates and momenta,  $(q, p)$ , at the time  $t$ , to a new set of constant quantities, which may be the  $2n$  initial values,  $(q_0, p_0)$ , at  $t = 0$ . With such a transformation, the equations of transformation relating the old and new canonical variables are exactly the desired solution of the mechanical problem:

$$q = q(q_0, p_0, t),$$

$$p = p(q_0, p_0, t).$$

They give the coordinates and momenta as a function of their initial values and the time. This last procedure is the more general one, especially as it is applicable, in principle at least, even when the Hamiltonian involves the time. We shall therefore begin our discussion by considering how such a transformation may be found.

### 10.1 ■ THE HAMILTON–JACOBI EQUATION FOR HAMILTON’S PRINCIPAL FUNCTION

We can automatically ensure that the new variables are constant in time by requiring that the transformed Hamiltonian,  $K$ , shall be identically zero, for then the equations of motion are

$$\begin{aligned}\frac{\partial K}{\partial P_i} &= \dot{Q}_i = 0, \\ -\frac{\partial K}{\partial Q_i} &= \dot{P}_i = 0.\end{aligned}\tag{10.1}$$

As we have seen,  $K$  must be related to the old Hamiltonian and to the generating function by the equation

$$K = H + \frac{\partial F}{\partial t},$$

and hence will be zero if  $F$  satisfies the equation

$$H(q, p, t) + \frac{\partial F}{\partial t} = 0. \quad (10.2)$$

It is convenient to take  $F$  as a function of the old coordinates  $q_i$ , the new constant momenta  $P_i$ , and the time; in the notation of the previous chapter we would designate the generating function as  $F_2(q, P, t)$ . To write the Hamiltonian in Eq. (10.2) as a function of the same variables, use may be made of the equations of transformation (cf. Eq. (9.17a)),

$$p_i = \frac{\partial F_2}{\partial q_i},$$

so that Eq. (10.2) becomes

$$H\left(q_1, \dots, q_n; \frac{\partial F_2}{\partial q_1}, \dots, \frac{\partial F_2}{\partial q_n}; t\right) + \frac{\partial F_2}{\partial t} = 0. \quad (10.3)$$

Equation (10.3), known as the *Hamilton–Jacobi equation*, constitutes a partial differential equation in  $(n + 1)$  variables,  $q_1, \dots, q_n; t$ , for the desired generating function. It is customary to denote the solution  $F_2$  of Eq. (10.3) by  $S$  and to call it *Hamilton's principal function*.

Of course, the integration of Eq. (10.3) only provides the dependence on the old coordinates and time; it would not appear to tell how the new momenta are contained in  $S$ . Indeed, the new momenta have not yet been specified except that we know they must be constants. However, the nature of the solution indicates how the new  $P_i$ 's are to be selected.

Mathematically Eq. (10.3) has the form of a first-order partial differential equation in  $n + 1$  variables. Suppose there exists a solution to Eq. (10.3) of the form

$$F_2 \equiv S = S(q_1, \dots, q_n; \alpha_1, \dots, \alpha_{n+1}; t), \quad (10.4)$$

where the quantities  $\alpha_1, \dots, \alpha_{n+1}$  are  $n + 1$  *independent* constants of integration. Such solutions are known as *complete solutions* of the first-order partial differential equation.\* One of the constants of integration, however, is in fact irrelevant to the solution, for it will be noted that  $S$  itself does not appear in Eq. (10.3); only its partial derivatives with respect to  $q$  or  $t$  are involved. Hence, if  $S$  is some solution of the differential equation, then  $S + \alpha$ , where  $\alpha$  is any constant, must also be a solution. One of the  $n + 1$  constants of integration in Eq. (10.4) must therefore appear only as an additive constant tacked on to  $S$ . But by the same token, an additive constant has no importance in a generating function, since only partial derivatives of the generating function occur in the transformation equations.

\*Equation (10.4) is not the only type of solution possible for Eq. (10.3). The most general form of the solution involves one or more arbitrary functions rather than arbitrary constants. Nor is there necessarily a unique solution of the form (10.4). There may be several complete solutions for the given equation. But all that is important for the subsequent argument is that there exist a complete solution.

Hence, for our purposes a complete solution to Eq. (10.3) can be written in the form

$$S = S(q_1, \dots, q_n; \alpha_1, \dots, \alpha_n; t), \quad (10.5)$$

where none of the  $n$  independent constants is solely additive. In this mathematical garb,  $S$  tallies exactly with the desired form for an  $F_2$  type of generating function, for Eq. (10.5) presents  $S$  as a function of  $N$  coordinates, the time  $t$ , and  $n$  independent quantities  $\alpha_i$ . We are therefore at liberty to take the  $n$  constants of integration to be the new (constant) momenta:

$$P_i = \alpha_i. \quad (10.6)$$

Such a choice does not contradict the original assertion that the new momenta are connected with the initial values of  $q$  and  $p$  at time  $t_0$ . The  $n$  transformation equations (9.17a) can now be written as

$$p_i = \frac{\partial S(q, \alpha, t)}{\partial q_i}, \quad (10.7)$$

where  $q, \alpha$  stand for the complete set of quantities. At the time  $t_0$ , these constitute  $n$  equations relating the  $n \alpha$ 's with the initial  $q$  and  $p$  values, thus enabling us to evaluate the constants of integration in terms of the specific initial conditions of the problem. The other half of the equations of transformation, which provide the new constant coordinates, appear as

$$Q_i = \beta_i = \frac{\partial S(q, \alpha, t)}{\partial \alpha_i}. \quad (10.8)$$

The constant  $\beta$ 's can be similarly obtained from the initial conditions, simply by calculating the value of the right side of Eq. (10.8) at  $t = t_0$  with the known initial values of  $q_i$ . Equations (10.8) can then be "turned inside out" to furnish  $q_j$  in terms of  $\alpha, \beta$ , and  $t$ :

$$q_j = q_j(\alpha, \beta, t), \quad (10.9)$$

which solves the problem of giving the coordinates as functions of time and the initial conditions.\* After the differentiation in Eqs. (10.7) has been performed,

\*As a mathematical point, it may be questioned whether the process of "turning inside out" is feasible for Eqs. (10.7) and (10.8), that is, whether they can be solved for  $\alpha_i$  and  $q_i$ , respectively. The question hinges on whether the equations in each set are independent, for otherwise they are obviously not sufficient to determine the  $n$  independent quantities  $\alpha_i$  or  $q_i$  as the case may be. To simplify the notation, let  $S_\alpha$  symbolize members of the set of partial derivatives of  $S$  with respect to  $\alpha_i$ , so that Eq. (10.8) is represented by  $\beta = S_\alpha$ . That the derivatives  $S_\alpha$  in (10.8) form independent functions of the  $q$ 's follows directly from the nature of a complete solution to the Hamilton-Jacobi equation; indeed this is what we mean by saying the  $n$  constants of integration are independent. Consequently, the Jacobian of  $S_\alpha$  with respect to  $q_i$  cannot vanish. Since the order of differentiation is immaterial, this is equivalent to saying that the Jacobian of  $S_q$  with respect to  $\alpha_i$  cannot vanish, which proves the independence of Eqs. (10.7).

Eqs. (10.9) may be substituted for the  $q$ 's, thus giving the momenta  $p_i$  as functions of the  $\alpha$ ,  $\beta$ , and  $t$ :

$$p_i = p_i(\alpha, \beta, t). \quad (10.10)$$

Equations (10.9) and (10.10) thus constitute the desired complete solution of Hamilton's equations of motion.

Hamilton's principal function is thus the generator of a canonical transformation to constant coordinates and momenta; *when solving the Hamilton–Jacobi equation, we are at the same time obtaining a solution to the mechanical problem*. Mathematically speaking, we have established an equivalence between the  $2n$  canonical equations of motion, which are first-order differential equations, to the first-order partial differential Hamilton–Jacobi equation. This correspondence is not restricted to equations governed by the Hamiltonian; indeed, the general theory of first-order partial differential equations is largely concerned with the properties of the equivalent set of first-order ordinary differential equations. Essentially, the connection can be traced to the fact that both the partial differential equation and its canonical equations stem from a common variational principle, in this case Hamilton's modified principle.

To a certain extent, the choice of the  $\alpha_i$ 's as the new momenta is arbitrary. We could just as well choose any  $n$  quantities,  $\gamma_i$ , which are independent functions of the  $\alpha_i$  constants of integration:

$$\gamma_i = \gamma_i(\alpha_1, \dots, \alpha_n). \quad (10.11)$$

By means of these defining relations, Hamilton's principal function can be written as a function of  $q_i$ ,  $\gamma_i$ , and  $t$ , and the rest of the derivation then goes through unchanged. It often proves convenient to take some particular set of  $\gamma_i$ 's as the new momenta, rather than the constants of integration that appear naturally in integrating the Hamilton–Jacobi equation.

Further insight into the physical significance of Hamilton's principal function  $S$  is furnished by an examination of its total time derivative, which can be computed from the formula

$$\frac{dS}{dt} = \frac{\partial S}{\partial q_i} \dot{q}_i + \frac{\partial S}{\partial t},$$

since the  $P_i$ 's are constant in time. By Eqs. (10.7) and (10.3), this relation can also be written

$$\frac{dS}{dt} = p_i \dot{q}_i - H = L, \quad (10.12)$$

so that Hamilton's principal function differs at most from the indefinite time integral of the Lagrangian only by a constant:

$$S = \int L dt + \text{constant}. \quad (10.13)$$

Now, Hamilton's principle is a statement about the definite integral of  $L$ , and from it we obtained the solution of the problem via the Lagrange equations. Here the same action integral, in an indefinite form, furnishes another way of solving the problem. In actual calculations, the result expressed by Eq. (10.13) is of no help, because we cannot integrate the Lagrangian with respect to time until  $q_i$  and  $p_i$  are known as functions of time, that is, until the problem is solved.

When the Hamiltonian does not depend explicitly upon the time, Hamilton's principle function can be written in the form

$$S(q, \alpha, t) = W(q, \alpha) - at, \quad (10.14)$$

where  $W(q, \alpha)$  is called *Hamilton's characteristic function*. The physical significance of  $W$  can be understood by writing its total time derivative

$$\frac{dW}{dt} = \frac{\partial W}{\partial q_i} \dot{q}_i.$$

Comparing this expression to the results of substituting Eq. (10.14) into Eq. (10.7), it is clear that

$$p_i = \frac{\partial W}{\partial q_i}, \quad (10.15)$$

and hence,

$$\frac{dW}{dt} = p_i \dot{q}_i. \quad (10.16)$$

This can be integrated to give

$$W = \int p_i \dot{q}_i dt = \int p_i dq_i, \quad (10.17)$$

which is just the abbreviated action defined by Eq. (8.80).

## 10.2 ■ THE HARMONIC OSCILLATOR PROBLEM AS AN EXAMPLE OF THE HAMILTON–JACOBI METHOD

To illustrate the Hamilton–Jacobi technique for solving the motion of mechanical systems, we shall work out in detail the simple problem of a one-dimensional harmonic oscillator. The Hamiltonian is

$$H = \frac{1}{2m}(p^2 + m^2\omega^2q^2) \equiv E, \quad (10.18)$$

where

$$\omega = \sqrt{\frac{k}{m}}, \quad (10.19)$$

$k$  being the force constant. We obtain the Hamilton–Jacobi equations for  $S$  by setting  $p$  equal to  $\partial S/\partial q$  and substituting in the Hamiltonian; the requirement that the new Hamiltonian vanishes becomes

$$\frac{1}{2m} \left[ \left( \frac{\partial S}{\partial q} \right)^2 + m^2 \omega^2 q^2 \right] + \frac{\partial S}{\partial t} = 0. \quad (10.20)$$

Since the explicit dependence of  $S$  on  $t$  is present only in the last term, Eq. (10.14) can be used to eliminate the time from the Hamilton–Jacobi equation (10.20)

$$\frac{1}{2m} \left[ \left( \frac{\partial W}{\partial q} \right)^2 + m^2 \omega^2 q^2 \right] = \alpha. \quad (10.21)$$

The integration constant  $\alpha$  is thus to be identified with the total energy  $E$ . This can also be recognized directly from Eq. (10.14) and the relation (cf. Eq. (10.3))

$$\frac{\partial S}{\partial t} + H = 0,$$

which then reduces to

$$H = \alpha.$$

Equation (10.21) can be integrated immediately to

$$W = \sqrt{2m\alpha} \int dq \sqrt{1 - \frac{m\omega^2 q^2}{2\alpha}}, \quad (10.22)$$

so that

$$S = \sqrt{2m\alpha} \int dq \sqrt{1 - \frac{m\omega^2 q^2}{2\alpha}} - \alpha t. \quad (10.23)$$

While the integration involved in Eq. (10.23) is not particularly difficult, there is no reason to carry it out at this stage, for what is desired is not  $S$  but its partial derivatives. The solution for  $q$  arises out of the transformation equation (10.8):

$$\beta' = \frac{\partial S}{\partial \alpha} = \sqrt{\frac{m}{2\alpha}} \int \frac{dq}{\sqrt{1 - \frac{m\omega^2 q^2}{2\alpha}}} - t,$$

which can be integrated without trouble to give

$$t + \beta' = \frac{1}{\omega} \arcsin q \sqrt{\frac{m\omega^2}{2\alpha}}. \quad (10.24)$$

Equation (10.24) can be immediately “turned inside out” to furnish  $q$  as a function of  $t$  and the two constants of integration  $\alpha$  and  $\beta = \beta' \omega$ :

$$q = \sqrt{\frac{2\alpha}{m\omega^2}} \sin(\omega t + \beta), \quad (10.25)$$

which is the familiar solution for a harmonic oscillator. Formally, the solution for the momentum comes from the transformation equation (10.7), which, using Eq. (10.22), can be written

$$p = \frac{\partial S}{\partial q} = \frac{\partial W}{\partial q} = \sqrt{2m\alpha - m^2\omega^2q^2}. \quad (10.26)$$

In conjunction with the solution for  $q$ , Eq. (10.25), this becomes

$$p = \sqrt{2m\alpha(1 - \sin^2(\omega t + \beta))},$$

or

$$p = \sqrt{2m\alpha} \cos(\omega t + \beta) \quad (10.27)$$

Of course, this result checks with the simple identification of  $p$  as  $m\dot{q}$ .

To complete the story, the constants  $\alpha$  and  $\beta$  must be connected with the initial conditions  $q_0$  and  $p_0$  at time  $t = 0$ . By squaring Eqs. (10.25) and (10.27), it is clearly seen that  $\alpha$  is given in terms of  $q_0$  and  $p_0$  by the equation

$$2m\alpha = p_0^2 + m^2\omega^2q_0^2. \quad (10.28)$$

The same result follows immediately of course from the previous identification of  $\alpha$  as the conserved total energy  $E$ . Finally, the phase constant  $\beta$  is related to  $q_0$  and  $p_0$  by

$$\tan \beta = m\omega \frac{q_0}{p_0}. \quad (10.29)$$

The choice  $q_0 = 0$  and hence  $\beta = 0$  corresponds to starting the motion with the oscillator at its equilibrium position  $q = 0$ .

Thus, Hamilton's principle function is the generator of a canonical transformation to a new coordinate that measures the phase angle of the oscillation and to a new canonical momentum identified as the total energy.

If the solution for  $q$  is substituted into Eq. (10.23), Hamilton's principal function can be written as

$$S = 2\alpha \int \cos^2(\omega t + \beta) dt - \alpha t = 2\alpha \int (\cos^2(\omega t + \beta) - \frac{1}{2}) dt. \quad (10.30)$$

Now, the Lagrangian is

$$\begin{aligned} L &= \frac{1}{2m}(p^2 - m^2\omega^2q^2) \\ &= \alpha(\cos^2(\omega t + \beta) - \sin^2(\omega t + \beta)) \\ &= 2\alpha(\cos^2(\omega t + \beta) - \frac{1}{2}), \end{aligned}$$

so that  $S$  is the time integral of the Lagrangian, in agreement with the general relation (10.13). Note that the identity could not be proved until *after* the solution to the problem had been obtained.

As another illustration for the Hamilton–Jacobi method, it is instructive to consider the two-dimensional anisotropic harmonic oscillator. If we let  $m$  be the mass of the oscillating body and  $k_x$  and  $k_y$  be the spring constants in the  $x$ - and  $y$ -directions, respectively, the Hamiltonian is

$$E = \frac{1}{2m}(p_x^2 + p_y^2 + m^2\omega_x^2x^2 + m^2\omega_y^2y^2),$$

where

$$\omega_x = \sqrt{\frac{k_x}{m}} \quad \text{and} \quad \omega_y = \sqrt{\frac{k_y}{m}}.$$

Since the coordinates and momenta separate into two distinct sets, the principal function can be written as a sum of the characteristic function for each pair. Assuming that we solve the  $y$ -functional dependency first, this means

$$S(x, y, \alpha, \alpha_y, t) = F_x(x, \alpha) + F_y(y, \alpha_y) - \alpha t, \quad (10.31)$$

and the Hamilton–Jacobi equation assumes the form

$$\frac{1}{2m} \left[ \left( \frac{\partial W}{\partial x} \right)^2 + m^2\omega_x^2x^2 + \left( \frac{\partial W}{\partial y} \right)^2 + m^2\omega_y^2y^2 \right] = \alpha \quad (10.32)$$

in analogy with Eq. (10.18). Since the variables are separated, the  $y$ -part of the Eq. (10.32) must be equal to a constant, which we call  $\alpha_y$ , so

$$\frac{1}{2m} \left( \frac{\partial W}{\partial y} \right)^2 + \frac{1}{2}m\omega_y^2y^2 = \alpha_y, \quad (10.33)$$

and we replace the  $y$ -term in (10.32) with  $\alpha_y$  from (10.33), yielding

$$\frac{1}{2m} \left( \frac{\partial W}{\partial x} \right)^2 + \frac{1}{2}m\omega_x^2x^2 = \alpha_x, \quad (10.34)$$

where we write  $\alpha - \alpha_y = \alpha_x$  showing the symmetry of Eqs. (10.33) and (10.34).

Each equation has a solution analogous to Eqs. (10.25) and (10.27), so

$$\begin{aligned} x &= \sqrt{\frac{2\alpha_x}{m\omega_x^2}} \sin(\omega_x t + \beta_x) \\ p_x &= \sqrt{2m\alpha_x} \cos(\omega_x t + \beta_x) \\ y &= \sqrt{\frac{2\alpha_y}{m\omega_y^2}} \sin(\omega_y t + \beta_y) \\ p_y &= \sqrt{2m\alpha_y} \cos(\omega_y t + \beta_y), \end{aligned} \tag{10.35}$$

where the  $\beta_i$ 's are phase constants and the total energy is given by

$$E = \alpha_x + \alpha_y = \alpha.$$

As a third example of Hamilton–Jacobi theory, we again consider the two-dimensional harmonic oscillator; only we will assume the oscillator is isotropic, so

$$k_x = k_y = k \quad \text{and} \quad \omega_x = \omega_y = \omega,$$

and use polar coordinates to write

$$\begin{aligned} x &= r \cos \theta & r &= \sqrt{x^2 + y^2} \\ y &= r \sin \theta & \theta &= \tan^{-1} \frac{y}{x} \\ p_x &= m\dot{x} & p_r &= m\dot{r} \\ p_y &= m\dot{y} & p_\theta &= mr^2\dot{\theta}. \end{aligned} \tag{10.36}$$

The Hamiltonian now written as

$$E = \frac{1}{2m} \left( p_r^2 + \frac{p_\theta^2}{r^2} + m^2 \omega^2 r^2 \right) \tag{10.37}$$

is cyclic in the angular coordinate  $\theta$ . The principle function can then be written as

$$\begin{aligned} S(r, \theta, \alpha, \alpha_\theta) &= W_r(r, \alpha) + W_\theta(\theta, \alpha_\theta) - \alpha t \\ &= W_r(r, \alpha) + \theta\alpha_\theta - \alpha t, \end{aligned} \tag{10.38}$$

where, as we show later, a cyclic coordinate  $q_i$  always has the characteristic function component  $W_{q_i} = q_i \alpha_i$ . The canonical momentum  $p_\theta$  associated with the cyclic coordinate,  $\theta$ , is calculated from the generating function

$$p_\theta = \frac{\partial F_\theta}{\partial \theta} = \alpha_\theta$$

has its expected constant value.

When this  $p_\theta$  is substituted into Eqs. (10.37) and (10.38),  $W_r(r, \alpha)$  satisfies

$$\frac{1}{2m} \left( \frac{\partial W_r}{\partial r} \right)^2 + \frac{\alpha_\theta^2}{2mr^2} + \frac{1}{2} m\omega^2 r^2 = \alpha. \quad (10.39)$$

Rather than solving this equation directly for  $W_r$ , we shall write the Cartesian coordinate solution for these conditions as

$$\begin{aligned} x &= \sqrt{\frac{2\alpha}{m\omega^2}} \sin(\omega t + \beta) & p_x &= \sqrt{2m\alpha} \cos(\omega t + \beta) \\ y &= \sqrt{\frac{2\alpha}{m\omega^2}} \sin \omega t & p_y &= \sqrt{2m\alpha} \cos \omega t, \end{aligned} \quad (10.35')$$

and use these to get the polar counterparts,

$$r = \sqrt{\frac{2\alpha}{m\omega^2}} \sqrt{\sin^2 \omega t + \sin^2(\omega t + \beta)}, \quad p_r = m\dot{r},$$

and

$$\theta = \tan^{-1} \left[ \frac{\sin \omega t}{\sin(\omega t + \beta)} \right], \quad p_\theta = mr^2 \dot{\theta}.$$

There are two limiting cases. The linear case is when  $\beta = 0$ , for which

$$r = \sqrt{\frac{4\alpha}{m\omega^2}} \sin \omega t, \quad p_r = \sqrt{2m\alpha} \cos \omega t,$$

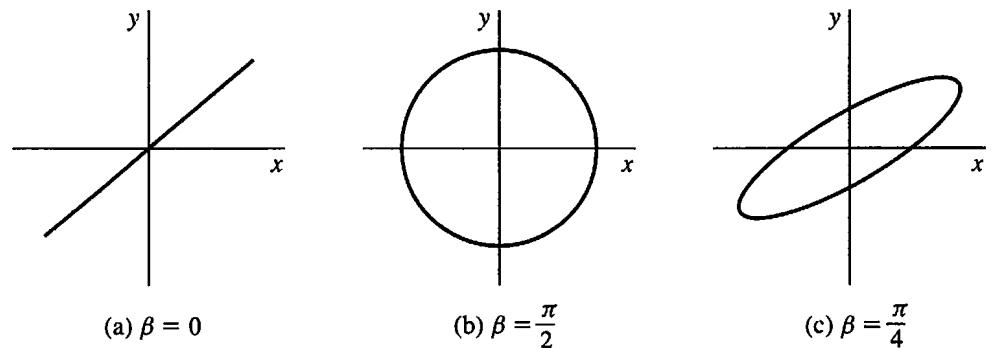
and

$$\theta = \frac{\pi}{4}, \quad p_\theta = 0.$$

The motion in an  $x$ - $y$  plot will be an oscillation along a diagonal line as shown in Fig. 10.1a. The other limiting case is when  $\beta = \pi/2$ , for which

$$\begin{aligned} r &= r_0 = \sqrt{\frac{2\alpha}{m\omega^2}}, & p_r &= 0 \\ \theta &= \omega t, & p_\theta &= mr_0^2 \omega. \end{aligned} \quad (10.42)$$

The motion in an  $x$ - $y$  plot for this limiting case is a circle of radius  $r_0$  as is shown in Figure 10.1b. For other values of  $\beta$  ( $0 < \beta < \pi/2$ ), the orbit in coordinate space is an ellipse. The case for  $\beta = \pi/4$  is shown in Fig. 10.1c. The plots shown in Fig. 10.1 are further examples of *Lissajous figures*.



**FIGURE 10.1** The two limiting cases (a) and (b) for the harmonic oscillator and an intermediate example (c).

## 10.3 ■ THE HAMILTON–JACOBI EQUATION FOR HAMILTON'S CHARACTERISTIC FUNCTION

It was possible to integrate the Hamilton–Jacobi equation for the simple harmonic oscillator primarily because  $S$  could be separated into two parts, one involving  $q$  only and the other only time. Such a separation of variables using Hamilton’s characteristic function  $W(q, \alpha)$  (Eq. (10.14)) is always possible *whenever the old Hamiltonian does not involve time explicitly*. This provides us with the restricted Hamilton–Jacobi equation

$$H\left(q_i, \frac{\partial W}{\partial q_i}\right) = \alpha_1, \quad (10.43)$$

which no longer involves the time. One of the constants of integration, namely  $\alpha_1$ , is thus equal to the constant value of  $H$ . (Normally  $H$  will be the energy, but remember that this need not always be the case, cf. Section 8.2.)

The time-independent function, Hamilton's characteristic function  $W$ , appears here merely as a part of the generating function  $S$  when  $H$  is constant. It can also be shown that  $W$  separately generates its own contact transformation with properties quite different from that generated by  $S$ . Let us consider a canonical transformation in which the new momenta are all constants of the motion  $\alpha_i$ , and where  $\alpha_1$  in particular is the constant of motion  $H$ . If the generating function for this transformation be denoted by  $W(q, P)$ , then the equations of transformation are

$$p_i = \frac{\partial W}{\partial q_i}, \quad Q_i = \frac{\partial W}{\partial P_i} = \frac{\partial W}{\partial \alpha_i}. \quad (10.44)$$

While these equations resemble Eqs. (10.7) and (10.8) respectively for Hamilton's principal function  $S$ , the condition now determining  $W$  is that  $H$  is the new canonical momentum  $\alpha_1$ :

$$H(q_i, p_i) = \alpha_1.$$

Using Eqs. (10.44), this requirement becomes the partial differential equation:

$$H \left( q_i, \frac{\partial W}{\partial q_i} \right) = \alpha_1,$$

which is seen to be identical with Eq. (10.43). Since  $W$  does not involve the time, the new and old Hamiltonians are equal, and it follows that  $K = \alpha_1$ .

Hamilton's characteristic function  $W$  thus generates a canonical transformation in which all the new coordinates are cyclic. It was noted in the introduction to this chapter that when  $H$  is a constant of the motion, a transformation of this nature in effect solves the mechanical problem involved, for the integration of the new equations of motion is then trivial. The canonical equations for  $P_i$ , in fact, merely repeat the statement that the momenta conjugate to the cyclic coordinates are all constant:

$$\dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0, \quad P_i = \alpha_i. \quad (10.45)$$

Because the new Hamiltonian depends upon only one of the momenta  $\alpha_i$ , the equations of motion for  $\dot{Q}_i$  are

$$\begin{aligned} \dot{Q}_i &= \frac{\partial K}{\partial \alpha_i} = 1, & i = 1, \\ &= 0, & i \neq 1, \end{aligned}$$

with the immediate solutions

$$\begin{aligned} Q_1 &= t + \beta_i \equiv \frac{\partial W}{\partial \alpha_1}, \\ Q_i &= \beta_i \equiv \frac{\partial W}{\partial \alpha_i} \quad i \neq 1. \end{aligned} \quad (10.46)$$

The only coordinate that is not simply a constant of the motion is  $Q_1$ , which is equal to the time plus a constant. We have here another instance of the conjugate relationship between the time as a coordinate and the Hamiltonian as its conjugate momentum.

The dependence of  $W$  on the old coordinates  $q_i$  is determined by the partial differential equation (10.43), which, like Eq. (10.3), is also referred to as the Hamilton–Jacobi equation. There will now be  $n$  constants of integration in a complete solution, but again one of them must be merely an additive constant. The  $n - 1$  remaining independent constants,  $\alpha_2, \dots, \alpha_n$ , together with  $\alpha_1$  may then be taken as the new constant canonical momenta. When evaluated at  $t_0$  the first half of Eqs. (10.44) serve to relate the  $n$  constants  $\alpha_i$  with the initial values of  $q_i$  and  $p_i$ . Finally, Eqs. (10.45) and (10.46) can be solved for the  $q_i$  as a function of  $\alpha_i$ ,  $\beta_i$ , and the time  $t$ , thus completing the solution of the problem. It will be noted

that  $(n - 1)$  of the Eqs. (10.46) do not involve the time at all. One of the  $q_i$ 's can be chosen as an independent variable, and the remaining coordinates can then be expressed in terms of it by solving only these time-independent equations. We are thus led directly to the *orbit equations* of the motion. In central force motion, for example, this technique would furnish  $r$  as a function of  $\theta$ , without the need for separately finding  $r$  and  $\theta$  as functions of time.

It is not always necessary to take  $\alpha_1$  and the constants of integration in  $W$  as the new constant canonical momenta. Occasionally it is desirable rather to use some particular set of  $n$  independent functions of the  $\alpha_i$ 's as the transformed momenta. Designating these constants by  $\gamma_i$  the characteristic function  $W$  can then be expressed in terms of  $q_i$  and  $\gamma_i$  as the independent variables. The Hamiltonian will in general depend upon more than one of the  $\gamma_i$ 's and the equations of motion for  $\dot{Q}_i$  become

$$\dot{Q}_i = \frac{\partial K}{\partial \gamma_i} = v_i,$$

where the  $v_i$ 's are functions of  $\gamma_i$ . In this case, all the new coordinates are linear functions of time:

$$Q_i = v_i t + \beta_i. \quad (10.47)$$

The form of  $W$  cannot be found a priori without obtaining a complete integral of the Hamilton–Jacobi equation. The procedures involved in solving a mechanical problem by either Hamilton's principal or characteristic function may now be summarized in the following tabular form:

The two methods of solution are applicable when the Hamiltonian

is any general function of $q, p, t$ :	is conserved: $H(q, p, t) = \text{constant.}$
--	--

We seek canonical transformations to new variables such that

all the coordinates and momenta $Q_i, P_i$ are constants of the motion.	all the momenta $P_i$ are constants.
--	--------------------------------------

To meet these requirements it is sufficient to demand that the new Hamiltonian

shall vanish identically: $K = 0.$	shall be cyclic in all the coordinates: $K = H(P_i) = \alpha_1.$
---------------------------------------	---

Under these conditions, the new equations of motion become

$\dot{Q}_i = \frac{\partial K}{\partial P_i} = 0,$	$\dot{Q}_i = \frac{\partial K}{\partial P_i} = v_i,$
$\dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0,$	$\dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0,$

with the immediate solutions

$$\begin{array}{ll} Q_i = \beta_i, & Q_i = v_i t + \beta_i \\ P_i = \gamma_i, & P_i = \gamma_i \end{array}$$

which satisfy the stipulated requirements.

The generating function producing the desired transformation is Hamilton's

<b>Principal Function:</b> $S(q, P, t),$	<b>Characteristic Function:</b> $W(q, P),$
---	---

satisfying the Hamilton–Jacobi partial differential equation:

$$H\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0. \quad | \quad H\left(q, \frac{\partial W}{\partial q}\right) - \alpha_1 = 0.$$

A complete solution to the equation contains

$n$ nontrivial constants of integration $\alpha_1, \dots, \alpha_n$ .	$n - 1$ nontrivial constants of integration, which together with $\alpha_1$ form a set of $n$ independent constants $\alpha_1, \dots, \alpha_n$ .
---	---

The new constant momenta,  $P_i = \gamma_i$ , can be chosen as any  $n$  independent functions of the  $n$  constants of integration:

$$P_i = \gamma_i(\alpha_1, \dots, \alpha_n), \quad \quad | \quad \quad P_i = \gamma_i(\alpha_1, \dots, \alpha_n),$$

so that the complete solutions to the Hamilton–Jacobi equation may be considered as functions of the new momenta:

$$S = S(q_i, \gamma_i, t). \quad \quad \quad | \quad \quad \quad W = W(q_i, \gamma_i).$$

In particular, the  $\gamma_i$ 's may be chosen to be the  $\alpha_i$ 's themselves. One-half of the transformations equations.

$$p_i = \frac{\partial S}{\partial q_i}, \quad | \quad p_i = \frac{\partial W}{\partial q_i},$$

are fulfilled automatically, since they have been used in constructing the Hamilton–Jacobi equation. The other half,

$$Q_i = \frac{\partial S}{\partial \gamma_i} = \beta_i, \quad | \quad Q_i = \frac{\partial W}{\partial \gamma_i} = v_i(\gamma_j)t + \beta_i,$$

can be solved for  $q_i$  in terms of  $t$  and the  $2n$  constants  $\beta_i, \gamma_i$ . The solution to the problem is then completed by evaluating these  $2n$  constants in terms of the initial values,  $(q_{i0}, p_{i0})$ , of the coordinates and momenta.

When the Hamiltonian does not involve time explicitly, both methods are suitable, and the generating functions are then related to each other according to the formula

$$S(q, P, t) = W(q, P) - \alpha_1 t.$$

#### 10.4 ■ SEPARATION OF VARIABLES IN THE HAMILTON–JACOBI EQUATION

It might appear from the preceding section that little practical advantage has been gained through the introduction of the Hamilton–Jacobi procedure. Instead of solving the  $2n$  ordinary differential equations that make up the canonical equations of motion, we now must solve the partial differential Hamilton–Jacobi equation, and partial differential equations can be notoriously complicated to solve. Under certain conditions, however, it is possible to separate the variables in the Hamilton–Jacobi equation, and the solution can then always be reduced to quadratures. In practice, the Hamilton–Jacobi technique becomes a useful computational tool only when such a separation can be effected.

A coordinate  $q_j$  is said to be separable in the Hamilton–Jacobi equation when (say) Hamilton’s principal function can be split into two additive parts, one of which depends only on the coordinate  $q_j$  and the other is entirely independent of  $q_j$ . Thus, if  $q_1$  is taken as a separable coordinate, then the Hamiltonian must be such that one can write

$$\begin{aligned} S(q_1, \dots, q_n; \alpha_1, \dots, \alpha_n; t) &= S_1(q_1; \alpha_1, \dots, \alpha_n; t) \\ &\quad + S'(q_2, \dots, q_n; \alpha_1, \dots, \alpha_n; t), \end{aligned} \quad (10.48)$$

and the Hamilton–Jacobi equation can be split into two equations—one separately for  $S_1$  and the other for  $S'$ . Similarly the Hamilton–Jacobi equation is described as *completely separable* (or simply, *separable*) if all the coordinates in the problem are separable. A solution for Hamilton’s principal function of the form

$$S = \sum_i S_i(q_i; \alpha_1, \dots, \alpha_n; t) \quad (10.49)$$

will then split the Hamilton–Jacobi equation into  $n$  equations of the type

$$H_i \left( q_j; \frac{\partial S_j}{\partial q_j}; \alpha_1, \dots, \alpha_n; t \right) + \frac{\partial S_j}{\partial t} = 0. \quad (10.50)$$

If the Hamiltonian does not explicitly depend upon the time, then, for each  $S_i$  we have

$$S_i(q_j; \alpha_1, \dots, \alpha_n; t) = W_i(q_j; \alpha_1, \dots, \alpha_n; t) - \alpha_i t, \quad (10.51)$$

which provide  $n$  restricted Hamilton–Jacobi equations,

$$H_i \left( q_i; \frac{\partial W_i}{\partial q_i}; \alpha_1, \dots, \alpha_n \right) = \alpha_i. \quad (10.52)$$

(No summation in Eqs. (10.50) to (10.52)!)

The functions  $H_i$  in Eqs. (10.50) and (10.52) may or may not be Hamiltonians, and the  $\alpha_i$  may be an energy, an angular momentum squared, or some other quantity depending on the nature of  $q_i$ . We shall show this by example in the Kepler problem in the next section.

The constants  $\alpha_i$  are referred to now as the *separation constants*. Each of the Eqs. (10.52) involves only one of the coordinates  $q_i$  and the corresponding partial derivative of  $W_i$  with respect to  $q_i$ . They are therefore a set of ordinary differential equations of a particularly simple form. Since the equations are only of first order, it is always possible to reduce them to quadratures; we have only to solve for the partial derivative of  $W_i$  with respect to  $q_i$  and then integrate over  $q_i$ . In practice, each  $H_i$  will only contain one or at most a few of the  $\alpha$ 's. There will also be cases where a subset of  $r$  variables can be separated in this fashion, leaving  $n - r$  variables, which will not separate. We shall also examine this eventuality in the next section.

It is possible to find examples in which the Hamilton–Jacobi equation can be solved without separating the time variable (cf. Exercise 8). Nonetheless, almost all useful applications of the Hamilton–Jacobi method involve Hamiltonians not explicitly dependent upon time, for which  $t$  is therefore a separable variable. The subsequent discussion on separability is thus restricted to such systems where  $H$  is a constant of motion, and Hamilton's characteristic function  $W$  will be used exclusively.

## 10.5 ■ IGNORABLE COORDINATES AND THE KEPLER PROBLEM

We can easily show that any cyclic or ignorable coordinate is separable. Suppose that the cyclic coordinate is  $q_1$ ; the conjugate momentum  $p_1$  is a constant, say  $\gamma$ . The Hamilton–Jacobi equation for  $W$  is then

$$H \left( q_2, \dots, q_n; \gamma; \frac{\partial W}{\partial q_2}, \dots, \frac{\partial W}{\partial q_n} \right) = \alpha_1. \quad (10.53)$$

If we try a separated solution of the form

$$W = W_1(q_1, \alpha) + W'(q_2, \dots, q_n; \alpha), \quad (10.54)$$

then it is obvious that Eq. (10.53) involves only the separate function  $W'$ , while  $W_1$  is the solution of the equation

$$p_1 = \gamma = \frac{\partial W_1}{\partial q_1}. \quad (10.55)$$

The constant  $\gamma$  is thus the separation constant, and the obvious solution for  $W_1$  (to within a trivial additive constant) is

$$W_1 = \gamma q_1,$$

and  $W$  is given by

$$W = W' + \gamma q_1. \quad (10.56)$$

There is an obvious resemblance between Eq. (10.56) and the form  $S$  assumes when  $H$  is not an explicit function of time, Eq. (10.43). Indeed, both equations can be considered as arising under similar circumstances. We have seen that  $t$  may be considered in some sense as a generalized coordinate with  $-H$  as its canonical momentum (cf. Eq. (8.58)). If  $H$  is conserved, then  $t$  may be treated as a cyclic coordinate.

If  $S$  of the  $n$  coordinates are noncyclic (that is, they appear explicitly in the Hamiltonian), then the Hamiltonian is of the form  $H(q_1, \dots, q_s; \alpha_1, \dots, \alpha_n; t)$ . The characteristic function can then be written as

$$W(q_1, \dots, q_s; \alpha_1, \dots, \alpha_n) = \sum_{i=1}^s W_i(q_i; \alpha_1, \dots, \alpha_n) + \sum_{i=s+1}^n q_i \alpha_i, \quad (10.56')$$

and there are  $s$  Hamilton–Jacobi equations to be solved:

$$H\left(q_1; \frac{\partial W_1}{\partial q_1}; \alpha_2, \dots, \alpha_n\right) = \alpha_1. \quad (10.57)$$

Since these are ordinary first-order differential equations in the independent variable  $q_1$ , they can be immediately reduced to quadratures, and the complete solutions for  $W$  can be obtained.

In general, a coordinate  $q_j$  can be separated if  $q_j$  and the conjugate momentum  $p_j$  can be segregated in the Hamiltonian into some function  $f(q_j, p_j)$  that does not contain any of the other variables. If we then seek a trial solution of the form

$$W = W_j(q_j, \alpha) + W'(q_i, \alpha),$$

where  $q_i$  represents the set of all  $q$ 's *except*  $q_j$ , then the Hamilton–Jacobi equation appears as

$$H\left(q_i, \frac{\partial W'}{\partial q_i}, f\left(q_j, \frac{\partial W_j}{\partial q_j}\right)\right) = \alpha_1. \quad (10.58)$$

In principle, at least, Eq. (10.58) can be inverted so as to solve for  $f$ :

$$f\left(q_j, \frac{\partial W_j}{\partial q_j}\right) = g\left(q_i, \frac{\partial W'}{\partial q_i}, \alpha_1\right). \quad (10.59)$$

The argument used previously in connection with Eq. (10.51) holds here in slightly varied guise;  $f$  is not a function of any of the  $q$ 's except  $q_j$ ;  $g$  on the other hand is independent of  $q_j$ . Hence, Eq. (10.59) can hold only if both sides are equal to the same constant, independent of all  $q$ 's:

$$\begin{aligned} f\left(q_j, \frac{\partial W_j}{\partial q_j}\right) &= \alpha_j, \\ g\left(q_i, \frac{\partial W'}{\partial q_i}\right) &= \alpha_j, \end{aligned} \quad (10.60)$$

and the separation of the variable has been accomplished.

Note that the separability of the Hamilton–Jacobi equation depends not only on the physical problem involved but also on the choice of the system of generalized coordinates employed. Thus, the one-body central force problem is separable in polar coordinates, but not in Cartesian coordinates. For some problems, it is not possible to completely separate the Hamilton–Jacobi equation, the famous three-body problem being one illustration. On the other hand, in many of the basic problems of mechanics and atomic physics, separation is possible in more than one set of coordinates. In general, it is feasible to solve the Hamilton–Jacobi equation in closed form only when the variables are completely separable. Considerable ingenuity has therefore been devoted to finding the separable systems of coordinates appropriate to each problem.

No simple criterion can be given to indicate what coordinate systems lead to separable Hamilton–Jacobi equations for any particular problem. In the case of orthogonal coordinate systems, the so-called Staeckel conditions have proved useful. They provide necessary and sufficient conditions for separability under certain circumstances. A proof of the sufficiency of the conditions and references will be found in Appendix D of the second edition of this text.

The Staeckel conditions for the separation of the Hamilton–Jacobi equations are:

1. The Hamiltonian is conserved.
2. The Lagrangian is no more than a quadratic function of the generalized velocities, so the Hamiltonian takes the form:

$$H = \frac{1}{2}(\tilde{\mathbf{p}} - \tilde{\mathbf{a}})\mathbf{T}^{-1}(\mathbf{p} - \mathbf{a}) + V(q). \quad (8.27)$$

3. The vector  $\mathbf{a}$  has elements  $a_i$  that are functions only of the corresponding coordinate, that is  $a_i = a_i(q_i)$ .
4. the potential function can be written as a sum of the form

$$V(q) = \sum_i \frac{V_i(q_i)}{T_{ii}}. \quad (10.61)$$

5. Consider the matrix  $\phi^{-1}$ , with an inverse  $\phi$  whose elements are

$$\delta_{ij}\phi_{ij}^{-1} = \frac{1}{T_{ii}}. \quad (\text{no summation on } i) \quad (10.62)$$

where

$$\left( \frac{\partial W_i}{\partial q_i} - a_i \right) = 2\delta_{ik}\phi_{kj}\gamma_j$$

with  $\gamma$  a constant unspecified vector. If the diagonal elements of both  $\phi$  and  $\phi^{-1}$  depend only upon the associated coordinate, that is,  $\phi^{-1}_{ii}$  and  $\phi_{ii}$  are constants or a function of  $q_i$  only, then provided 1–4 are true, the Hamiltonian–Jacobi equations separate.

Since we have assumed that the generalized coordinates  $q_i$  form an orthogonal coordinate system, the matrix  $T$  (introduced in Section 8.1) is diagonal. It follows that the inverse matrix  $T^{-1}$  is also diagonal and, if we are dealing with a particle in an external force field, the diagonal elements are:

$$\phi_{ii}^{-1} = \frac{1}{T_{ii}} = \frac{1}{m}, \quad (\text{no summation}) \quad (10.63)$$

so the fifth Stackel condition is satisfied.

If the Staeckel conditions are satisfied, then Hamilton's characteristic function is completely separable:

$$W(q) = \sum_i W_i(q_i),$$

with the  $W_i$  satisfying equations of the form

$$\left( \frac{\partial W_i}{\partial q_i} - a_i \right)^2 = -2V_i(q_i) + 2\phi_{ij}\gamma_j, \quad (10.64)$$

where  $\gamma_j$  are constants of integration (and there is summation only over the index  $j$ ).

While these conditions appear mysterious and complicated, their application usually is fairly straightforward. As an illustration of some of the ideas developed here about separability, the Hamilton–Jacobi equation for a particle moving in a central force will be discussed in polar coordinates. The problem will then be generalized to arbitrary potential laws, to furnish an application of the Staeckel conditions.

Let us first consider the central force problem in terms of the polar coordinates  $(r, \psi)$  in the plane of the orbit. The motion then involves only two degrees of freedom and the Hamiltonian has the form

$$H = \frac{1}{2m} \left( p_r^2 + \frac{p_\psi^2}{r^2} \right) + V(r), \quad (10.65)$$

which is cyclic in  $\psi$ . Consequently, Hamilton's characteristic function appears as

$$W = W_1(r) + \alpha_\psi \psi, \quad (10.66)$$

where  $\alpha_\psi$  is the constant angular momentum  $p_\psi$  conjugate to  $\psi$ . The Hamilton–Jacobi equation then becomes

$$\left( \frac{\partial W_1}{\partial r} \right)^2 + \frac{\alpha_\psi^2}{r^2} + 2mV(r) = 2m\alpha_1, \quad (10.67)$$

where  $\alpha_1$  is the constant identified physically as the total energy of the system. Solving Eq. (10.66) for the partial derivative of  $W_1$  we obtain

$$\frac{\partial W_1}{\partial r} = \sqrt{2m(\alpha_1 - V) - \frac{\alpha_\psi^2}{r^2}},$$

so that  $W$  is

$$W = \int dr \sqrt{2m(\alpha_1 - V) - \frac{\alpha_\psi^2}{r^2}} + \alpha_\psi \psi. \quad (10.68)$$

With this form for the characteristic function, the transformation equations (10.46) appear as

$$t + \beta_1 = \frac{\partial W}{\partial \alpha_1} = \int \frac{m dr}{\sqrt{2m(\alpha_1 - V) - \frac{\alpha_\psi^2}{r^2}}}, \quad (10.69a)$$

and

$$\beta_2 = \frac{\partial W}{\partial \alpha_\psi} = - \int \frac{\alpha_\psi dr}{r^2 \sqrt{2m(\alpha_1 - V) - \frac{\alpha_\psi^2}{r^2}}} + \psi. \quad (10.69b)$$

Equation (10.69a) furnishes  $r$  as a function of  $t$  and agrees with the corresponding solution, Eq. (3.18), found in Chapter 3, with  $\alpha_1$  and  $\alpha_\psi$  written explicitly as  $E$  and  $l$ , respectively. It has been remarked previously that the remaining transformation equations for  $Q_i$ , here only Eq. (10.69b), should provide the orbit equation. If the variable of integration in Eq. (10.69b) is changed to  $u = 1/r$ , the equation reduces to

$$\psi = \beta_2 - \int \frac{du}{\sqrt{\frac{2m}{\alpha_\psi^2}(\alpha_1 - V) - u^2}}$$

which agrees with Eq. (3.37) previously found for the orbit, identifying  $\psi$  as  $\theta$  and  $\beta_2$  as  $\theta_0$ .

As a further example of separation of variables, we shall examine the same central force problem, but in spherical polar coordinates, that is, ignoring our a priori knowledge that the orbit lies in a plane. The appropriate Hamiltonian has been shown to be (cf. Eq. (8.29)):

$$H = \frac{1}{2m} \left( p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2 \theta} \right) + V(r). \quad (10.70)$$

If the variables in the corresponding Hamilton–Jacobi equation are separable, then Hamilton’s characteristic function must have the form

$$W = W_r(r) + W_\theta(\theta) + W_\phi(\phi). \quad (10.71)$$

The coordinate  $\phi$  is cyclic in the Hamiltonian and hence

$$W_\phi = \alpha_\phi \phi \quad (10.72)$$

where  $\alpha_\phi$  is a constant of integration. In terms of this form for  $W$ , the Hamilton–Jacobi equation reduces to

$$\left( \frac{\partial W_r}{\partial r} \right)^2 + \frac{1}{r^2} \left[ \left( \frac{\partial W_\theta}{\partial \theta} \right)^2 + \frac{\alpha_\phi^2}{\sin^2 \theta} \right] + 2mV(r) = 2mE, \quad (10.73)$$

where we have explicitly identified the constant Hamiltonian with the total energy  $E$ . Note that all dependence on  $\theta$ , and on  $\theta$  alone, has been segregated into the expression within the square brackets. The Hamilton–Jacobi equation then conforms to the appearance of Eq. (10.58), and following the argument given there we see that the quantity in the square brackets must be a constant:

$$\left( \frac{\partial W_\theta}{\partial \theta} \right)^2 + \frac{\alpha_\phi^2}{\sin^2 \theta} = \alpha_\theta^2. \quad (10.74)$$

Finally the dependence of  $W$  on  $r$  is given by the remainder of the Hamilton–Jacobi equation:

$$\left( \frac{\partial W_r}{\partial r} \right)^2 + \frac{\alpha_\theta^2}{r^2} = 2m(E - V(r)). \quad (10.75)$$

The variables in the Hamilton–Jacobi equation are thus completely separated. Equations (10.74) and (10.75) may be easily reduced to quadratures providing at least a formal solution for  $W_\theta(\theta)$  and  $W_r(r)$ , respectively.

Note that the constants of integration  $\alpha_\phi$ ,  $\alpha_\theta$ ,  $\alpha_1$  all have directly recognizable physical meanings. The quantity  $\alpha_\phi$  is of course the constant value of the angular momentum about the polar axis (cf. Eq. (10.44)):

$$\alpha_\phi = p_\phi = \frac{\partial W_\phi}{\partial \phi}. \quad (10.76)$$

To identify  $\alpha_\theta$  we use Eq. (10.44) to rewrite Eq. (10.74) as

$$p_\theta^2 + \frac{p_\theta^2}{\sin^2 \theta} = \alpha_\theta^2, \quad (10.74')$$

so that the Hamiltonian, Eq. (10.70) appears as

$$H = \frac{1}{2m} \left( p_r^2 + \frac{\alpha_\theta^2}{r^2} \right) + V(r). \quad (10.70')$$

Comparison with Eq. (10.65) for the Hamiltonian as expressed in terms of polar coordinates in the plane of the orbit shows that  $\alpha_\theta$  is the same as  $p_\psi$ , the magnitude of the total angular momentum:

$$\alpha_\theta = p_\psi \equiv l. \quad (10.77)$$

Lastly,  $\alpha_1$  is of course the total energy  $E$ . Indeed, the three differential equations for the component parts of  $W$  can be looked on as statements of conservation theorems. Equation (10.75) says the  $z$ -component of the angular momentum vector,  $\mathbf{L}$ , is conserved, while Eq. (10.74) states the conservation of the magnitude,  $l$ , of the angular momentum. And Eq. (10.75) is a form of the energy conservation theorem.

In this simple example, some of the power and elegance of the Hamilton–Jacobi method begins to be apparent. A few short steps suffice to obtain the dependence of  $r$  on  $t$  and the orbit equation, Eqs. (10.69a and b), results derived earlier only with considerable labor. The conserved quantities of the central force problem also appear automatically. Separation of variables for the purely central force problem can also be performed in other coordinate systems, for example, parabolic coordinates, and the conserved quantities appear there in forms appropriate to the particular coordinates.

Finally, we can employ the Staeckel conditions to find the most general form of a scalar potential  $V$  for a single particle for which the Hamilton–Jacobi equation is separable in spherical polar coordinates. The matrix  $\phi$  of the Staeckel conditions depends only on the coordinate system and not on the potential. Since the Hamilton–Jacobi equation is separable in spherical polar coordinates for at least one potential, that is, the central force potential, it follows that the matrix  $\phi$  does exist. The specific form of  $\phi$  is not needed to answer our question. Further, since a by hypothesis is zero, all we need do is apply Eq. (10.62) to find the most general separable form of  $V$ . From the kinetic energy (Eq. 8.28'), the diagonal elements of  $T$  are

$$T_{rr} = m, \quad T_{\theta\theta} = mr^2, \quad T_{\phi\phi} = mr^2 \sin^2 \theta.$$

By Eq. (10.62) it follows that the desired potential must have the form

$$V(q) = V_r(r) + \frac{V_\theta(\theta)}{r^2} + \frac{V_\phi(\phi)}{r^2 \sin^2 \theta}. \quad (10.78)$$

It is easy to verify directly that with this potential the Hamilton–Jacobi equation is still completely separable in spherical polar coordinates.

### 10.6 ■ ACTION-ANGLE VARIABLES IN SYSTEMS OF ONE DEGREE OF FREEDOM

Of especial importance in many branches of physics are systems in which the motion is periodic. Very often we are interested not so much in the details of the orbit as in the frequencies of the motion. An elegant and powerful method of handling such systems is provided by a variation of the Hamilton–Jacobi procedure. In this technique, the integration constants  $\alpha_i$  appearing directly in the solution of the Hamilton–Jacobi equation are not themselves chosen to be the new momenta. Instead, we use suitably defined constants  $J_i$ , which form a set of  $n$  independent functions of the  $\alpha_i$ 's, and which are known as the *action variables*.

For simplicity, we shall first consider in this section systems of one degree of freedom. It is assumed the system is conservative so that the Hamiltonian can be written as

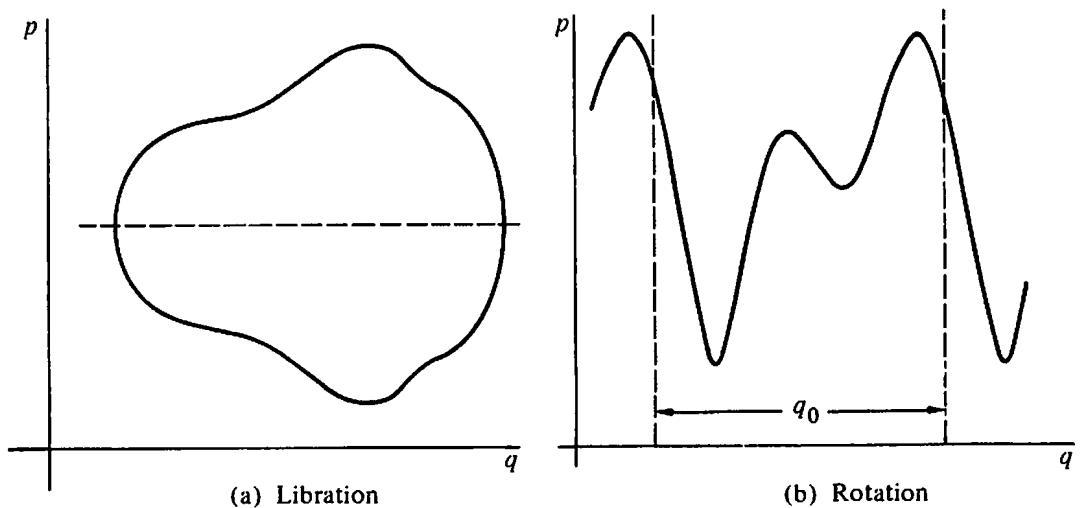
$$H(q, p) = \alpha_1.$$

Solving for the momentum, we have that

$$p = p(q, \alpha_1), \quad (10.79)$$

which can be looked on as the equation of the orbit traced out by the system point in the two-dimensional phase space,  $p, q$  when the Hamiltonian has the constant value  $\alpha_1$ . What is meant by the term “periodic motion” is determined by the characteristics of the phase space orbit. Two types of periodic motion may be distinguished:

1. In the first type, the orbit is *closed*, as shown in Fig. 10.2(a), and the system point retraces its steps periodically. Both  $q$  and  $p$  are then periodic functions of the time with the same frequency. Periodic motion of this nature will be found when the initial position lies between two zeros of the kinetic energy. It is often designated by the astronomical name *libration*, although to a physicist it is more likely to call to mind the common oscillatory systems, such as the one-dimensional harmonic oscillator.
2. In the second type of periodic motion, the orbit in phase space is such that  $p$  is some periodic function of  $q$ , with period  $q_0$ , as illustrated in Fig. 10.2(b). Equivalently, this kind of motion implies that when  $a$  is increased by  $q_0$ , the configuration of the system remains essentially unchanged. The most familiar example is that of a rigid body constrained to rotate about a given axis, with  $q$  as the angle of rotation. Increasing  $q$  by  $2\pi$  then produces no essential change in the state of the system. Indeed, the position coordinate in this type of periodicity is invariably an angle of rotation, and the motion



**FIGURE 10.2** Orbit of the system point in phase space for periodic motion of one-dimensional systems.

will be referred to simply as *rotation*, in contrast to libration. The values of  $q$  are no longer bounded but can increase indefinitely.

It may serve to clarify these ideas to note that both types of periodicity may occur in the same physical system. The classic example is the simple pendulum where  $q$  is the angle of deflection  $\theta$ . If the length of the pendulum is  $l$  and the potential energy is taken as zero at the point of suspension, then the constant energy of the system is given by

$$E = \frac{p_\theta^2}{2ml^2} - mgl \cos \theta. \quad (10.80)$$

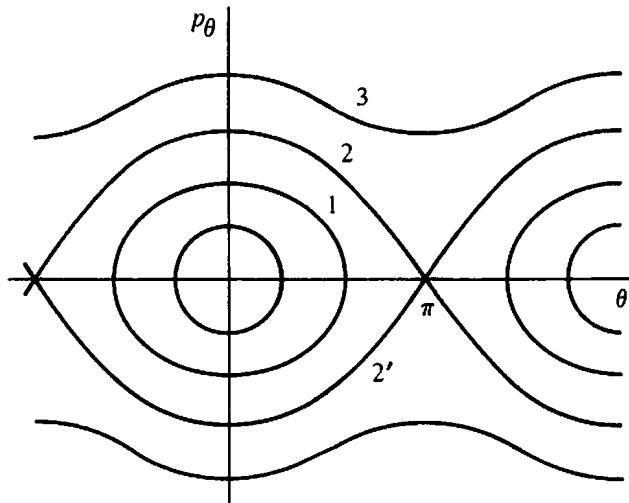
Solving Eq. (10.64) for  $p_\theta$ , the equation of the path of the system point in phase space is

$$p_\theta = \pm \sqrt{2ml^2(E + mgl \cos \theta)}. \quad (10.81)$$

If  $E$  is less than  $mgl$ , then physical motion of the system can only occur for  $|\theta|$  less than a bound,  $\theta'$ , defined by the equation

$$\cos \theta' = -\frac{E}{mgl}.$$

Under these conditions, the pendulum oscillates between  $-\theta'$  and  $+\theta'$ , which is a periodic motion of the libration type. The system point then traverses some such path in phase space as the curve 1 of Fig. 10.3. However, if  $E > mgl$ , all values of  $\theta$  correspond to physical motion and  $\theta$  can increase without limit to produce a periodic motion of the rotation type. What happens physically in this case is that the pendulum has so much energy that it can swing through the vertical position



**FIGURE 10.3** Phase space orbits for the simple pendulum.

$\theta = \pi$  and therefore continues rotating. Curve 3 in Fig. 10.3 corresponds to the rotation motion of the pendulum. The limiting case when  $E = mgl$  is illustrated by curves 2 and 2' in Fig. 10.3. At this energy, the pendulum arrives at  $\theta = \pi$ , the vertical position, with zero kinetic energy, that is,  $p_\theta = 0$ . It is then in unstable equilibrium and could in principle remain there indefinitely. However, if there is the slightest perturbation, it could continue its motion either along curve 2 or switch to curve 2'—it could fall down either way. The point  $\theta = \pi$ ,  $p_\theta = 0$  is a saddle point of the Hamiltonian function  $H = E(p_\theta, \theta)$  and there are two paths of constant  $E$  in phase space that intersect at the saddle point. We have here an instance of what is called a *bifurcation*, a phenomenon that will be discussed extensively in the next chapter. (See also Section 6.6.)

For either type of periodic motion, we can introduce a new variable  $J$  designed to replace  $\alpha_1$  as the transformed (constant) momentum. The so-called action variable  $J$  is defined as (cf. Eq. (8.80))

$$J = \oint p \, dq, \quad (10.82)$$

where the integration is to be carried over a complete period of libration or of rotation, as the case may be. (The designation as action variable stems from the resemblance of Eq. (10.82) to the abbreviated action of Section 8.6. Note that  $J$  always has the dimensions of an angular momentum.) From Eq. (10.79), it follows that  $J$  is always some function of  $\alpha_1$  alone:

$$\alpha_1 \equiv H = H(J). \quad (10.83)$$

Hence, Hamilton's characteristic function can be written as

$$W = W(q, J). \quad (10.84)$$

The generalized coordinate conjugate to  $J$ , known as the *angle variable*  $w$ , is defined by the transformation equation:

$$w = \frac{\partial W}{\partial J}. \quad (10.85)$$

Correspondingly, the equation of motion for  $w$  is

$$\dot{w} = \frac{\partial H(J)}{\partial J} = v(J), \quad (10.86)$$

where  $v$  is a constant function of  $J$  only. Equation (10.86) has the immediate solution

$$w = vt + \beta, \quad (10.87)$$

so that  $w$  is a linear function of time, exactly as in Eq. (10.47).

So far the action-angle variables appear as no more than a particular set of the general class of transformed coordinates to which the Hamilton-Jacobi equation leads. Equation (10.85) could be solved for  $q$  as a function of  $w$  and  $J$ , which, in combination with Eq. (10.87), would give the desired solution for  $q$  as a function of time. But when employed in this fashion the variables have no significant advantage over any other set of coordinates generated by  $W$ . Their particular merit rises rather from the physical interpretation that can be given to  $v$ . Consider the change in  $w$  as  $q$  goes through a complete cycle of libration or rotation, as given by

$$\Delta w = \oint \frac{\partial w}{\partial q} dq. \quad (10.88)$$

By Eq. (10.85), this can also be written

$$\Delta w = \oint \frac{\partial^2 W}{\partial q \partial J} dq. \quad (10.89)$$

Because  $J$  is a constant, the derivative with respect to  $J$  can be taken outside the integral sign:

$$\Delta w = \frac{d}{dJ} \oint \frac{\partial W}{\partial q} dq = \frac{d}{dJ} \oint p dq = 1, \quad (10.90)$$

where the last step follows from the definition for  $J$ , Eq. (10.82).

Equation (10.90) states that  $w$  changes by unity as  $q$  goes through a complete period. But from Eq. (10.87), it follows that if  $\tau$  is the period for a complete cycle of  $q$ , then

$$\Delta w = 1 = v\tau.$$

Hence, the constant  $v$  can be identified as the reciprocal of the period,

$$v = \frac{1}{\tau}, \quad (10.91)$$

and is therefore *the frequency associated with the periodic motion of  $q$* . The use of action-angle variables thus provides a powerful technique for obtaining the frequency of periodic motion *without finding a complete solution to the motion of the system*. If it is known a priori that a system of one degree of freedom is periodic according to the definitions given above, then the frequency can be found once  $H$  is determined as a function of  $J$ . The derivative of  $H$  with respect to  $J$ , by Eq. (10.86), then directly gives the frequency  $v$  of the motion. The designation of  $w$  as an angle variable becomes obvious from the identification of  $v$  in Eq. (10.87) as a frequency. Since  $J$  has the dimensions of an angular momentum, the coordinate  $w$  conjugate to it is an angle.\*

As an illustration of the application of action-angle variables to find frequencies, let us again consider the familiar linear harmonic oscillator problem. From Eqs. (10.26) and the defining equation (10.82), the constant action variable  $J$  is given by

$$J = \oint p dq = \oint \sqrt{2m\alpha - m^2\omega^2 q^2} dq, \quad (10.92)$$

where  $\alpha$  is the constant total energy and  $\omega^2 = k/m$ . The substitution (10.25)

$$q = \sqrt{\frac{2\alpha}{m\omega^2}} \sin \theta$$

reduces the integral to

$$J = \frac{2\alpha}{\omega} \int_0^{2\pi} \cos^2 \theta d\theta, \quad (10.93)$$

where the limits are such as to correspond to a complete cycle in  $q$ . This integrates to

$$J = \frac{2\pi\alpha}{\omega}$$

or, solving for  $\alpha$ ,

$$\alpha \equiv H = \frac{J\omega}{2\pi}. \quad (10.94)$$

The frequency of oscillation is therefore

\*For some applications the action variable is defined in the literature of celestial mechanics as  $(2\pi)^{-1}$  times the value given in Eq. (10.82). By Eq. (10.90), the corresponding angle variable is  $2\pi$  times our definition and in place of  $v$  we have  $\omega$ , the angular frequency. However, we shall stick throughout to the familiar definitions used in physics, as given above.

$$\frac{\partial H}{\partial J} = v = \frac{\omega}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{k}{m}}, \quad (10.95)$$

which is the customary formula for the frequency of a linear harmonic oscillator. Although it is entirely unnecessary for obtaining the frequencies, it is nevertheless instructive (and useful for future applications) to write the solutions, Eqs. (10.25) and (10.27), in terms of  $J$  and  $w$ . It will be recognized first that the combination  $(\omega t + \beta)$  is by Eqs. (10.95) and (10.87) the same as  $2\pi w$ , with the constant of integration suitably redefined. Hence, the solutions for  $q$ , Eq. (10.25), and  $p$ , Eq. (10.27), take on the form

$$q = \sqrt{\frac{J}{\pi m \omega}} \sin 2\pi w, \quad (10.96)$$

$$p = \sqrt{\frac{m J \omega}{\pi}} \cos 2\pi w. \quad (10.97)$$

Note that Eqs. (10.96) and (10.97) can also be looked on as the transformation equations from the  $(w, J)$  set of canonical variables to the  $(q, p)$  canonical set.

## 10.7 ■ ACTION-ANGLE VARIABLES FOR COMPLETELY SEPARABLE SYSTEMS\*

Action-angle variables can also be introduced for certain types of motion of systems with many degrees of freedom, providing there exists one or more sets of coordinates in which the Hamilton–Jacobi equation is completely separable. As before, only conservative systems will be considered, so that Hamilton's characteristic function will be used. Complete separability means that the equations of canonical transformation have the form

$$p_i = \frac{\partial W_i(q_i; \alpha_1, \dots, \alpha_n)}{\partial q_i}, \quad (10.98)$$

which provides each  $p_i$  as a function of the  $q_i$  and the  $n$  integration constants  $\alpha_j$ :

$$p_i = p_i(q_i; \alpha_1, \dots, \alpha_n). \quad (10.99)$$

Equation (10.99) is the counterpart of Eq. (10.79), which applied to systems of one degree of freedom. It will be recognized that Eq. (10.99) here represents the orbit equation of the projection of the system point on the  $(p_i, q_i)$  plane in phase space. We can define action-angle variables for the system when the orbit equations for *all* of the  $(q_i, p_i)$  pairs describe either closed orbits (libration, as in Fig. 10.2(a)) or periodic functions of  $q_i$  (rotation, as in Fig. 10.2(b)).

Note that this characterization of the motion does not mean that each  $q_i$  and  $p_i$  will necessarily be periodic functions of the time, that is, that they repeat their

\*Unless otherwise stated, the summation convention will *not* be used in this section.

values at fixed time intervals. Even when each of the separated  $(q_i, p_i)$  sets are indeed periodic in this sense, the overall system motion need not be periodic. Thus, in a three-dimensional harmonic oscillator the frequencies of motion along the three Cartesian axes may all be different. In such an example, it is clear the complete motion of the particle may not be periodic. If the separate frequencies are not rational fractions of each other, the particle will not traverse a closed curve in space but will describe an open “Lissajous figure.” Such motion will be described as *multiply periodic*. It is the advantage of the action-angle variables that they lead to an evaluation of all the frequencies involved in multiply periodic motion without requiring a complete solution of the motion.

In analogy to Eq. (10.82), the action variables  $J_i$  are defined in terms of line integrals over complete periods of the orbit in the  $(q_i, p_i)$  plane:

$$J_i = \oint p_i dq_i. \quad (10.100)$$

If one of the separation coordinates is cyclic, its conjugate momentum is constant. The corresponding orbit in the  $q_i, p_i$  plane of phase space is then a horizontal straight line, which would not appear to be in the nature of a periodic motion. Actually the motion can be considered as a limiting case of the rotation type of periodicity, in which  $q_i$  may be assigned any arbitrary period. Since the coordinate in a rotation periodicity is invariably an angle, such a cyclic  $q_i$  always has a natural period of  $2\pi$ . Accordingly, the integral in the definition of the action variable corresponding to a cyclic angle coordinate is to be evaluated from 0 to  $2\pi$ , and hence

$$J_i = 2\pi p_i \quad (10.101)$$

for all cyclic variables.

By Eq. (10.98),  $J_i$  can also be written as

$$J = \oint \frac{\partial W_i(q_i; \alpha_1, \dots, \alpha_n)}{\partial q_i} dq_i. \quad (10.102)$$

Since  $q_i$  is here merely a variable of integration, each action variable  $J_i$  is a function only of the  $n$  constants of integration appearing in the solution of the Hamilton–Jacobi equation. Further, it follows from the independence of the separate variable pairs  $(q_i, p_i)$  that the  $J_i$ ’s form  $n$  independent functions of the  $\alpha_i$ ’s and hence are suitable for use as a set of new constant momenta. Expressing the  $\alpha_i$ ’s as functions of the action variables, the characteristic function  $W$  can be written in the form

$$W = W(q_1, \dots, q_n; J_1, \dots, J_n) = \sum_j W_j(q_j; J_1, \dots, J_n),$$

while the Hamiltonian appears as a function of the  $J_i$ ’s only:

$$H = \alpha_1 = H(J_1, \dots, J_n). \quad (10.103)$$

As in the system of one degree of freedom, we can define conjugate angle variables  $w_i$  by the equations of transformation that here appear as

$$w_i = \frac{\partial W}{\partial J_i} = \sum_{j=1}^n \frac{\partial W_j(q_j; J_1, \dots, J_n)}{\partial J_i}. \quad (10.104)$$

Note in general  $w_i$  could be a function of several or all of the  $q_i$ ; that is,  $w_i = w_i(q_1, \dots, q_n; J_1, \dots, J_n)$ . The  $w_i$ 's satisfy equations of motion given by

$$\dot{w}_i = \frac{\partial H(J_1, \dots, J_n)}{\partial J_i} = v_i(J_1, \dots, J_n). \quad (10.105)$$

Because the  $v_i$ 's are constants, functions of the action variables only, the angle variables are all linear functions of time

$$w_i = v_i t + \beta_i. \quad (10.106)$$

Note that in general the separate  $w_i$ 's increase in time at different rates.

The constants  $v_i$  can be identified with the frequencies of the multiply periodic motion, but the argument to demonstrate the relation is more subtle than for periodic systems of one degree of freedom. The transformation equations to the  $(w, J)$  set of variables implies that each  $q_j$  (and  $p_j$ ) is a function of the constants  $J_i$  and the variables  $w_i$ . What we want to find is what sort of mathematical function the  $q$ 's are of the  $w$ 's. To do this, we examine the change in a particular  $w_i$  when each of the variables  $q_j$  is taken through an integral number,  $m_j$ , of cycles of libration or rotation. In carrying out this purely mathematical procedure, we are clearly *not* following the motion of the system in time. It is as if the flow of time were suspended and each of the  $q$ 's were moved, manually as it were, independently through a number of cycles of their motion. In effect, we are dealing with analogues of the virtual displacements of Chapter 1, and accordingly the infinitesimal change in  $w_i$  as the  $q_j$ 's are changed infinitesimally will be denoted by  $\delta w_i$  and is given by

$$\delta w_i = \sum_j \frac{\partial w_i}{\partial q_j} dq_j = \sum_j \frac{\partial^2 W}{\partial J_i \partial q_j} dq_j,$$

where use has been made of Eq. (10.104). The derivative with respect to  $q_i$  vanishes except for the  $W_j$  constituent of  $W$ , so that by Eq. (10.98)  $\delta w_i$  reduces to

$$\delta w_i = \frac{\partial}{\partial J_i} \sum_j p_j(q_j, J) dq_j. \quad (10.107)$$

Equation (10.107) represents  $\delta w_i$  as the sum of independent contributions each involving the  $q_j$  motion. The total change in  $w_i$  as a result of the specified ma-

neuver is therefore

$$\Delta w_i = \sum_j \frac{\partial}{\partial J_i} \oint_{m_j} p_j(q_j, J) dq_j. \quad (10.108)$$

the differential operator with respect to  $J_i$  can be kept outside the integral signs because throughout the cyclic motion of  $q_i$  all the  $J$ 's are of course constant. Below each integral sign, the symbol  $m_j$  indicates the integration is over  $m_j$  cycles of  $q_j$ . But each of the integrals is, by the definition of the action variables, exactly  $m_j J_j$ . Since the  $J$ 's are independent, it follows that

$$\Delta w_i = m_i. \quad (10.109)$$

Further, note that if any  $q_j$  does not go through a complete number of cycles, then in the integration over  $q_j$  there will be a remainder of an integral over a fraction of a cycle and  $\Delta w_i$  will not have an integral value. If the sets of  $w$ 's and  $m$ 's are treated as vectors  $\mathbf{w}$  and  $\mathbf{m}$ , respectively, Eq. (10.109) can be written as

$$\Delta \mathbf{w} = \mathbf{m}. \quad (10.109')$$

Suppose, first, that the separable motions are all of the libration type so that each  $q_j$ , as well as  $p_j$ , returns to its initial value on completion of a complete cycle. The result described by Eq. (10.109') could now be expressed somewhat as follows:  $\boldsymbol{\eta}$  (the vector of  $q$ 's and  $p$ 's) is such a function of  $\mathbf{w}$  that a change  $\Delta \boldsymbol{\eta} = 0$  corresponds to a change  $\Delta \mathbf{w} = \mathbf{m}$ , a vector of integer values. Since the number of cycles in the chosen motions of  $q_j$  are arbitrary,  $\mathbf{m}$  can be taken as zero except for  $m_i = 1$ , and all the components of  $\boldsymbol{\eta}$  remain unchanged or return to their original values. Hence, in the most general case the components of  $\boldsymbol{\eta}$  must be periodic functions of *each*  $w_i$  with period unity; that is, the  $q$ 's and  $p$ 's are multiply periodic functions of the  $w$ 's with unit periods. Such a multiply periodic function can always be represented by a multiple Fourier expansion, which for  $q_k$ , say, would appear as

$$q_k = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \dots, \dots, \sum_{j_n=-\infty}^{\infty} a_{j_1, \dots, j_n}^{(k)} \cdot e^{2\pi i(j_1 w_1 + j_2 w_2 + j_3 w_3 + \dots + j_n w_n)}, \quad (\text{libration}) \quad (10.110)$$

where the  $j$ 's are  $n$  integer indices running from  $-\infty$  to  $\infty$ . By treating the set of  $j$ 's also as a vector in the same  $n$ -dimensional space with  $\mathbf{w}$ , the expansion can be written more compactly as

$$q_k = \sum_{\mathbf{j}} a_{\mathbf{j}}^{(k)} e^{2\pi i \mathbf{j} \cdot \mathbf{w}}, \quad (\text{libration}). \quad (10.110')$$

If we similarly write Eq. (10.109') as a vector equation,

$$\mathbf{w} = \mathbf{v}t + \boldsymbol{\beta}, \quad (10.106')$$

then the time dependence of  $q_k$  appears in the form

$$q_k(t) = \sum_j a_j^{(k)} e^{2\pi j \cdot (vt + \beta)}, \quad (\text{libration}). \quad (10.111)$$

Note that in general  $q_k(t)$  is *not* a periodic function of  $t$ . Unless the various  $v_i$ 's are commensurate (that is, rational multiples of each other),  $q_k$  will not repeat its values at regular intervals of time. Considered as a function of  $t$ ,  $q_k$  is designated as a *quasi-periodic* function. Finally it should be remembered that the coefficients  $a_j^{(k)}$  can be found by the standard procedure for Fourier coefficients; that is, they are given by the multiple integral over the unit cell in  $w$  space:

$$a_j^{(k)} = \int_0^1, \dots, \int_0^1 q_k(w) e^{-2\pi i j \cdot w} (dw). \quad (10.112)$$

Here  $(dw)$  stands for the volume element in the  $n$ -dimensional space of the  $w_i$ 's.

When the motion is in the nature of a rotation, then in a complete cycle of the separated variable pair  $(q_k, p_k)$  the coordinate  $q_k$  does not return to its original value, but instead increases by the value of its period  $q_{0k}$ . Such a rotation coordinate is therefore not itself even multiply periodic. However, during the cycle we have seen that  $w_k$  increases by unity. Hence, the function  $q_k - w_k q_{0k}$  does return to its initial value and, like the librational coordinates, is a multiply periodic function of all the  $w$ 's with unit periods. We can therefore expand the function in a multiple Fourier series analogous to Eq. (10.110)

$$q_k - w_k q_{0k} = \sum_j a_j^{(k)} e^{2\pi i j \cdot w}, \quad (\text{rotation}) \quad (10.113)$$

or

$$q_k = q_{0k}(v_k t + \beta_k) + \sum_j a_j^{(k)} e^{2\pi i j \cdot (vt + \beta)}, \quad (\text{rotation}). \quad (10.114)$$

Thus, it is always possible to derive a multiply periodic function from a rotation coordinate, which can then be handled exactly like a librational coordinate. To simplify the further discussion, we shall therefore confine ourselves primarily to the librational type of motion.

The separable momentum coordinates,  $p_k$ , are by the nature of the assumed motion also multiply periodic functions of the  $w$ 's and can be expanded in a multiple Fourier series similar to Eq. (10.110). It follows then that any function of the several variable pairs  $(q_k, p_k)$  will also be multiply periodic functions of the  $w$ 's and can be written in the form

$$f(q, p) = \sum_j b_j e^{2\pi i j \cdot w} = \sum_j b_j e^{2\pi i j \cdot (vt + \beta)}. \quad (10.115)$$

For example, where the Cartesian coordinate of particles in the system are not themselves the separation coordinates, they can still be written as functions of time in the fashion of Eq. (10.115).

While Eqs. (10.110) and (10.111) represent the most general type of motion consistent with the assumed nature of the problem, not all systems will exhibit this full generality. In particular, for most problems simple enough to be used as illustrations of the application of action-angle variables, Eq. (10.104) simplifies to

$$w_i = \frac{\partial w_i}{\partial J_i}(q_i; J_1, \dots, J_n) \quad (10.116)$$

and each separation coordinate  $q_i$  is a function only of its corresponding  $w_k$ . When this happens,  $q_k$  is then a periodic function of  $w_k$  (and therefore of time), and the multiple Fourier series reduces to a single Fourier series:

$$q_k = \sum_j a_j^{(k)} e^{2\pi i j w_k} = \sum_j a_j^{(k)} e^{2\pi i j (v_k t + \beta_k)}. \quad (10.117)$$

In the language of Chapter 6, in such problems the  $q_k$ 's are in effect the normal coordinates of the system. However, even when the motion in the  $q$ 's can be so simplified, it frequently happens that functions of all the  $q$ 's, such as Cartesian coordinates, remain multiply periodic functions of the  $w$ 's and must be represented as in Eq. (10.115). If the various frequencies  $v_k$  are incommensurate, then such functions are not periodic functions of time. The motion of a two-dimensional anisotropic harmonic oscillator provides a convenient and familiar example of these considerations.

Suppose that in a particular set of Cartesian coordinates the Hamiltonian is given by

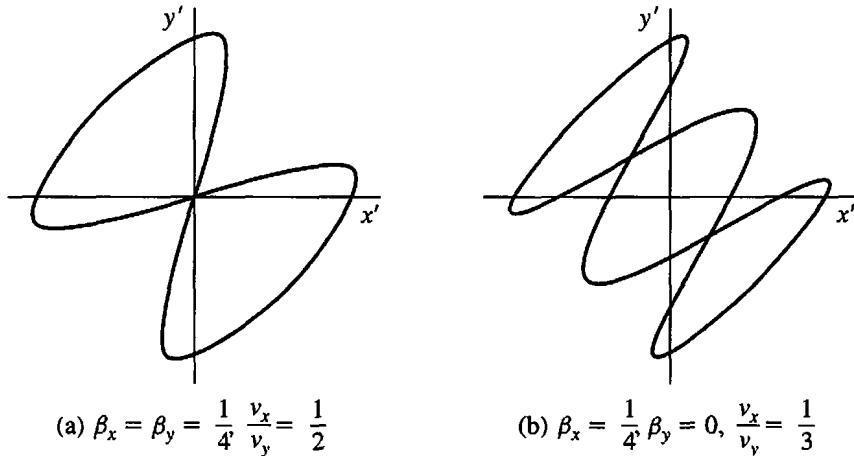
$$H = \frac{1}{2m}[(p_x^2 + 4\pi^2 m^2 v_x^2 x^2) + (p_y^2 + 4\pi^2 m^2 v_y^2 y^2)].$$

These Cartesian coordinates are therefore suitable separation variables, and each will exhibit simple harmonic motion with frequencies  $v_x$  and  $v_y$ , respectively. Thus, the solutions for  $x$  and  $y$  are particularly simple forms of the single Fourier expansions of Eq. (10.117). Suppose now that the coordinates are rotated  $45^\circ$  about the  $z$  axis; the components of the motion along the new  $x'$ ,  $y'$  axes will be

$$\begin{aligned} x' &= \frac{1}{\sqrt{2}}[x_0 \cos 2\pi(v_x t + \beta_x) + y_0 \cos 2\pi(v_y t + \beta_y)], \\ y' &= \frac{1}{\sqrt{2}}[y_0 \cos 2\pi(v_y t + \beta_y) - x_0 \cos 2\pi(v_x t + \beta_x)]. \end{aligned} \quad (10.118)$$

If  $v_x/v_y$  is a rational number, these two expressions will be commensurate, corresponding to closed Lissajous figures of the type shown in Fig. 10.4. But if  $v_x$  and  $v_y$  are incommensurable, the Lissajous figure never exactly retraces its steps and Eqs. (10.118) provide simple examples of multiply periodic series expansions of the form (10.117).

Even when  $q_k$  is a multiply periodic function of all the  $w$ 's, we intuitively feel there must be a special relationship between  $q_k$  and its corresponding  $w_k$  (and



**FIGURE 10.4** Lissajous figures for Eq. (10.118). (a)  $\beta_x = \beta_y = \frac{1}{4}, \frac{v_x}{v_y} = \frac{1}{2}$  (b)  $\beta_x = \frac{1}{4}, \beta_y = 0, \frac{v_x}{v_y} = \frac{1}{3}$ .

therefore  $v_k$ ). After all, the argument culminating in Eq. (10.109) says that when  $q_k$  alone goes through its complete cycle,  $w_k$  increases by unity, while the other  $w$ 's return to their initial values. It was only in 1961 that J. Vinti succeeded in expressing this intuitive feeling in a precise and rigorous statement.\*

Suppose that the time interval  $T$  contains  $m$  complete cycles of  $q_k$  plus a fraction of a cycle. In general, the times required for each successive cycle will be different, since  $q_k$  will not be a periodic function of  $t$ . Then Vinti showed, on the basis of a theorem in number theory, that as  $T$  increases indefinitely,

$$\lim_{t \rightarrow \infty} \frac{m}{T} = v_k. \quad (10.119)$$

The *mean* frequency of the motion of  $q_k$  is therefore always given by  $v_k$ , even when the entire motion is more complicated than a periodic function with frequency  $v_k$ .

Barring commensurability of all the frequencies, a multiply periodic function can always be formed from the generating function  $W$ . The defining equation for  $J_i$ , Eq. (10.102), in effect states that when  $q_i$  goes through a complete cycle; that is, when  $w_i$  changes by unity, the characteristic function increases by  $J_i$ . It follows that the function

$$W' = W - \sum_k w_k J_k \quad (10.120)$$

remains unchanged when *each*  $w_k$  is increased by unity, all the other angle variables remaining constant. Equation (10.120) therefore represents a multiply periodic function that can be expanded in terms of the  $w_i$  (or of the frequencies  $v_i$ ) by a series of the form of Eq. (10.115). Since the transformation equations for the

\*J. Vinti, *J. Res. Nat. Bur. Standards*, **65B**, 131 (1961).

angle variables are

$$w_k = \frac{\partial W}{\partial J_k},$$

it will be recognized that Eq. (10.120) defines a Legendre transformation from the  $q, J$  basis to the  $q, w$  basis. Indeed, comparison with Eq. (9.15) in combination with Eq. (9.12) shows that if  $W(q, J)$  is a generating function of the form  $F_2(q, P)$ , then  $W'(q, w)$  is the corresponding generating function of the type  $F_1(q, Q)$ , transforming in both cases from the  $(q, p)$  variables to the  $(w, J)$  variables. While  $W'$  thus generates the same transformation as  $W$ , it is of course *not* a solution of the Hamilton Jacobi equation.

It has been emphasized that the system configuration is multiply periodic only if the frequencies  $v_i$  are not rational fractions of each other. Otherwise, the configuration repeats after a sufficiently long time and would therefore be simply periodic. The formal condition for the commensurability of two frequencies  $v_i$  and  $v_j$  is that they satisfy the relation  $j_i v_i = j_j v_j$  (no sum) where  $j_i$  and  $j_j$  are nonzero positive integers. For complete commensurability, all pairs of frequencies must satisfy relations of the form

$$j_i v_i = j_k v_k, \quad (\text{no sum}) \quad (10.121)$$

where the  $j_i$  and  $j_k$  are nonzero positive integers.

When we can express any  $v_i$  as a rational fraction of any of the other frequencies, the system is said to be *completely commensurate*. If only  $m + 1$  of the  $n$  frequencies satisfy Eq. (10.121), the system is said to be  $m$ -fold commensurate. For example, consider the set of seven frequencies  $v_1 = 3 \text{ MHz}$ ,  $v_2 = 5 \text{ MHz}$ ,  $v_3 = 7 \text{ MHz}$ ,  $v_4 = 2\sqrt{2} \text{ MHz}$ ,  $v_5 = 3\sqrt{2} \text{ MHz}$ ,  $v_6 = \sqrt{3} \text{ MHz}$ ,  $v_7 = \sqrt{7} \text{ MHz}$ . The first three  $v_1$ ,  $v_2$ , and  $v_3$  are triply commensurate, the next two  $v_4$  and  $v_5$  are doubly commensurate.

There is an interesting connection between commensurability and the coordinates in which the Hamilton–Jacobi equation is separable. It can be shown that the path of the system point for a noncommensurate system completely fills a limited region of both configuration and phase space. This can be seen in the Lissajous figures of incommensurate frequencies.

Suppose the problem is such that the motion in any one of the separation coordinates is simply periodic and has therefore been shown to be independent of the motion of the other coordinates. Hence, the path of the system point as a whole must be limited by the surfaces of constant  $q_i$  and  $p_i$  that mark the bounds of the oscillatory motion of the separation variables. (The argument is easily extended to rotation by limiting all angles to the region 0 to  $2\pi$ .) These surfaces therefore define the volume in space that is densely filled by the system point orbit. It follows that the separation of variables in noncommensurate systems must be unique; the Hamilton–Jacobi equation cannot be separated in two different coordinate systems (aside from trivial variations such as change of scale). The possibility of separating the motion in more than one set of coordinates thus normally provides evidence that the system is commensurate.

The simplest example of being commensurate is degeneracy which occurs when two or more of the frequencies are equal. If two of the force constants in a three-dimensional harmonic oscillator are equal, then the corresponding frequencies are identical and the system is singly degenerate. In an isotropic linear oscillator, the force constants are the same along all directions, all frequencies are equal, and the system is completely degenerate.

Whenever this simple degeneracy is present, the fundamental frequencies are no longer independent, and the periodic motion of the system can be described by less than the full complement of  $n$  frequencies. Indeed, the  $m$  conditions of degeneracy can be used to reduce the number of frequencies to  $n - m + 1$ . The reduction of the frequencies may be most elegantly performed by means of a point transformation of the action-angle variables. The  $m$  degeneracy conditions may be written where  $j_{ki}$  are positive or negative integers

$$\sum_{i=1}^n j_{ki} v_i = 0, \quad k = 1, \dots, m. \quad (10.122)$$

Consider now a point transformation from  $(w, J)$  to  $(w', J')$  defined by the generating function (cf. Eq. (9.26) where the summation convention is used):

$$F_2 = \sum_{k=1}^m \sum_{i=1}^n J'_k j_{ki} w_i + \sum_{k=m+1}^n J'_k w_k. \quad (10.123)$$

The transformed coordinates are

$$\begin{aligned} w'_k &= \sum_{i=1}^n j_{ki}, \quad k = 1, \dots, m, \\ &= w_k, \quad k = m + 1, \dots, n. \end{aligned} \quad (10.124)$$

Correspondingly, the new frequencies are

$$\begin{aligned} v'_k &= \dot{w}'_k = \sum_{i=1}^n j_{ki} v_i = 0 \quad k = 1, \dots, m, \\ &= v_k \quad k = m + 1, \dots, n. \end{aligned} \quad (10.125)$$

Thus in the transformed coordinates,  $m$  of the frequencies are zero, and we are left with a set of  $n - m$  independent frequencies plus the zero frequency. It is obvious that the new  $w'_k$  may also be termed as angle variables in the sense that the system configuration is multiply periodic in the  $w'_k$  coordinates with the fundamental period unity. The corresponding constant action variables are given as the solution of the  $n$  equations of transformation

$$J_i = \sum_{k=1}^m J'_k j_{ki} + \sum_{k=m+1}^n J'_k \delta_{ki}. \quad (10.126)$$

The zero frequencies correspond to constant factors in the Fourier expansion. These are of course also present in the original Fourier series in terms of the  $v$ 's, Eq. (10.110), occurring whenever the indices  $j_i$  are such that degeneracy conditions are satisfied. Since

$$v'_i = \frac{\partial H}{\partial J'_i},$$

the Hamiltonian must be independent of the action variables  $J'_i$  whose corresponding frequencies vanish. In a completely degenerate system, the Hamiltonian can therefore be made to depend upon only one of the action variables.

Note that Hamilton's characteristic function  $W$  also serves as the generating function for the transformation from the  $(q, p)$  set to the  $(w', J')$  set. Since the  $J'$  quantities are  $n$  independent constants, the original constants of integration may be expressed in terms of the  $J'$  set, and  $W$  given as  $W(q, J')$ . In this form, it is a generating function to a new set of canonical variables for which the  $J'$  quantities are the canonical momenta. But by virtue of the point transformation generated by the  $F_2$  of Eq. (10.123), we know that  $w'$  is conjugate to  $J'$ . Hence, it follows that the new coordinates generated by  $W(q, J')$  must be the angle variable  $w'$  set, with equations of transformation given by

$$w'_i = \frac{\partial W}{\partial J'_i}. \quad (10.127)$$

(For a more formal proof of Eq. (10.127) based on the algebraic structure of Eq. (10.123), see Derivation 3.)

The problem of the bound motion of a particle in an inverse-square law central force illustrates many of the phenomena involved in degeneracy. A discussion of this problem also affords an opportunity to show how the action-angle technique is applied to specific systems, and to indicate the connections with Bohr's quantum mechanics and with celestial mechanics. Accordingly, the next section is devoted to a detailed treatment of the Kepler problem in terms of action-angle variables.

## 10.8 ■ THE KEPLER PROBLEM IN ACTION-ANGLE VARIABLES\*

To exhibit all of the properties of the solution, we shall examine the motion in three dimensional space, rather than make use of our a priori knowledge that the orbit lies in a plane. In terms of spherical polar coordinates, the Kepler problem becomes a special case of the general treatment given above in Section 10.5 for central force motion in space. Equations (10.70) through (10.77) can be taken over here immediately, replacing  $V(r)$  wherever it occurs by its specific form

$$V(r) = -\frac{k}{r}. \quad (10.128)$$

\*The summation convention will be resumed from here on.

Since the potential  $V(r)$  depends only upon one of the three coordinates, it follows that the Hamilton-Jacobi equation is completely separable in spherical polar coordinates. We shall confine our discussion to the bound case, that is,  $E < 0$ . Hence, the motion in each of the coordinates will be periodic—libration in  $r$  and  $\theta$ , and rotation in  $\phi$ . The conditions for the application of action-angle variables are thus satisfied, and we can proceed to construct the action variables on the basis of the defining equation (10.102). From Eq. (10.72), it follows that

$$J_\phi = \oint \frac{\partial W}{\partial \phi} d\phi = \oint \alpha_\phi d\phi. \quad (10.129a)$$

Similarly, on the basis of Eq. (10.74),  $J_\theta$  is given by

$$J_\theta = \oint \frac{\partial W}{\partial \theta} d\theta = \oint \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2 \theta}} d\theta. \quad (10.129b)$$

Finally the integral for  $J_r$  from Eq. (10.75), is

$$J_r = \oint \frac{\partial W}{\partial r} dr = \oint \sqrt{2mE + \frac{2mk}{r} - \frac{\alpha_\theta^2}{r^2}} dr. \quad (10.129c)$$

The first integral is trivial;  $\phi$  goes through  $2\pi$  radians in a complete revolution and therefore

$$J_\phi = 2\pi\alpha_\phi = 2\pi p_\phi. \quad (10.130)$$

This result could have been predicted beforehand, for  $\phi$  is a cyclic coordinate in  $H$ , and Eq. (10.130) is merely a special case of Eq. (10.101) for the action variables corresponding to cyclic coordinates. Integration of Eq. (10.129b) can be performed in various ways; a procedure involving only elementary rules of integration will be sketched here. If the polar angle of the total angular momentum vector is denoted by  $i$ , so that

$$\cos i = \frac{\alpha_\phi}{\alpha_\theta}, \quad (10.131)$$

then Eq. (10.129b) can also be written as

$$J_\theta = \alpha_\theta \oint \sqrt{1 - \cos^2 i \csc^2 \theta} d\theta. \quad (10.132)$$

The complete circuital path of integration is for  $\theta$  going from a limit  $-\theta_0$  to  $+\theta_0$  and back again, where  $\sin \theta_0 = \cos i$ , or  $\theta_0 = (\pi/2) - i$ . Hence, the circuital integral can be written as 4 times the integral over from 0 to  $\theta_0$ , or after some manipulation,

$$J_\theta = 4\alpha_\theta \int_0^{\theta_0} \csc \theta \sqrt{\sin^2 i - \cos^2 \theta} d\theta.$$

The substitution

$$\cos \theta = \sin i \sin \psi$$

transforms the integral to

$$J_\theta = 4\alpha_\theta \sin^2 i \int_0^{\pi/2} \frac{\cos^2 \psi d\psi}{1 - \sin^2 i \sin^2 \psi}. \quad (10.133)$$

Finally, with the substitution

$$u = \tan \psi,$$

the integral becomes

$$\begin{aligned} J_\theta &= 4\alpha_\theta \sin^2 i \int_0^\infty \frac{du}{(1+u^2)(1+u^2 \cos^2 i)} \\ &= 4\alpha_\theta \int_0^\infty du \left( \frac{1}{1+u^2} - \frac{\cos^2 i}{1+u^2 \cos^2 i} \right). \end{aligned} \quad (10.134)$$

This last form involves only well-known integrals, and the final result\* is

$$J_\theta = 2\pi \alpha_\theta (1 - \cos i) = 2\pi(\alpha_\theta - \alpha_\phi). \quad (10.135)$$

The last integral (Eq. (10.129c)), for  $J_r$ , can now be written as

$$J_r = \oint \sqrt{2mE + \frac{2mk}{r} - \frac{(J_\theta + J_\phi)^2}{4\pi^2 r^2}} dr. \quad (10.136)$$

After performing the integration, this equation can be solved for the energy  $E \equiv H$  in terms of the three action variables  $J_\phi$ ,  $J_\theta$ ,  $J_r$ . Note that  $J_\phi$  and  $J_\theta$  can occur in  $E$  only in the combination  $J_\theta + J_\phi$ , and hence the corresponding frequencies  $v_\phi$  and  $v_\theta$  must be equal, indicating a degeneracy. This result has not involved the inverse-square law nature of the central force; *any motion produced by a central force is at least singly degenerate*. The degeneracy is of course a consequence of the fact that the motion is confined to a plane normal to the constant angular momentum vector  $\mathbf{L}$ . Motion in this plane implies that  $\theta$  and  $\phi$  are related to each other such that as  $\phi$  goes through a complete  $2\pi$  period,  $\theta$  varies through a complete cycle between the limits  $(\pi/2) \pm i$ . Hence, the frequencies in  $\theta$  and  $\phi$  are necessarily equal.

The integral involved in Eq. (10.136) can be evaluated by elementary means, but the integration is most elegantly and quickly performed using the complex

\*In evaluating the integral of the second term in the final integrand of Eq. (10.134), it has been assumed that  $\cos i$  is positive. This is always possible, since there is no preferred direction for the  $z$  axis in the problem and it may be chosen at will. If  $\cos i$  were negative, the sign of  $\alpha_\phi$  in Eq. (10.135) would be positive. For changes in the subsequent formulas, see Exercise 23.

contour integration method of residues. For the benefit of those familiar with this technique, we shall outline the steps involved in integrating Eq. (10.136). Bound motion can occur only when  $E$  is negative (cf. Section 3.3), and since the integrand is equal to  $p_r = m\dot{r}$ , the limits of the motion are defined by the roots  $r_1$  and  $r_2$  of the expression in the square root sign. If  $r_1$  is the inner bound, as in Fig. 3.6, a complete cycle of  $r$  involves going from  $r_1$  to  $r_2$  and then back again to  $r_1$ . On the outward half of the journey, from  $r_1$  to  $r_2$ ,  $p_r$  is positive and we must take the positive square root. However, on the return trip to  $r_1$ ,  $p_r$  is negative and the square root must likewise be negative. The integration thus involves both branches of a double-valued function, with  $r_1$  and  $r_2$  as the branch points. Consequently, the complex plane can be represented as one of the sheets of a Riemann surface, slit along the real axis from  $r_1$  to  $r_2$  (as indicated in Fig. 10.5).

Since the path of integration encloses the line between the branch points  $r_1$  and  $r_2$ , the method of residues cannot be applied directly. However, we may also consider the path as enclosing all the rest of the complex plane, the direction of integration now being in the reverse (clockwise) direction. The integrand is single-valued in this region, and there is now no bar to the application of the method of residues. Only two singular points are present, namely, the origin and infinity, and the integration path can be distorted into two clockwise circles enclosing these two points. Now, the sign in front of the square root in the integrand must be negative for the region along the real axis below  $r_1$ , as can be seen by examining the behavior of the function in the neighborhood of  $r_1$ . If the integrand is represented as

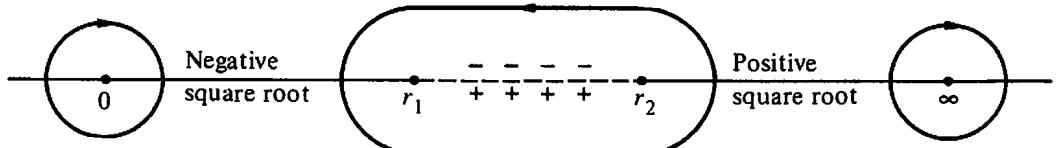
$$-\sqrt{A + \frac{2B}{r} - \frac{C}{r^2}},$$

the residue at the origin is

$$R_0 = -\sqrt{-C}.$$

Above  $r_2$ , the sign of the square root on the real axis is found to be positive, and the residue is obtained by the standard technique of changing the variable of integration to  $z = r^{-1}$ :

$$-\oint \frac{1}{z^2} \sqrt{A + 2Bz - Cz^2} dz. \quad (10.137)$$



**FIGURE 10.5** The complex  $r$  plane in the neighborhood of the real axis; showing the paths of integration occurring in the evaluation of  $J$ .

Expansion about  $z = 0$  now furnishes the residue

$$R_\infty = -\frac{B}{\sqrt{A}}.$$

The total integral is  $-2\pi i$  times the sum of the residues:

$$J_r = 2\pi i \left( \sqrt{-C} + \frac{B}{\sqrt{A}} \right), \quad (10.138)$$

or, upon substituting the coefficients  $A$ ,  $B$ , and  $C$ :

$$J_r = -(J_\theta + J_\phi) + \pi k \sqrt{\frac{2m}{-E}}. \quad (10.139)$$

Equation (10.139) supplies the functional dependence of  $H$  upon the action variables; for solving for  $E$ , we have

$$H \equiv E = -\frac{2\pi^2 mk^2}{(J_r + J_\theta + J_\phi)^2}. \quad (10.140)$$

Note that, as predicted,  $J_\theta$  and  $J_\phi$  occur only in the combination  $J_\theta + J_\phi$ . More than that, all three of the action variables appear only in the form  $J_r + J_\theta + J_\phi$ . Hence, all of the frequencies are equal; *the motion is completely degenerate*. This result could also have been predicted beforehand, for we know that with an inverse-square law of force the orbit is closed for negative energies. With a closed orbit, the motion is simply periodic and therefore, in this case, completely degenerate. If the central force contained an  $r^{-3}$  term, such as is provided by first-order relativistic corrections, then the orbit is no longer closed but is in the form of a precessing ellipse. One of the degeneracies will be removed in this case, but the motion is still singly degenerate, since  $v_\theta = v_\phi$  for all central forces. The one frequency for the motion here is given by

$$v = \frac{\partial H}{\partial J_r} = \frac{\partial H}{\partial J_\theta} = \frac{\partial H}{\partial J_\phi} = \frac{4\pi^2 mk^2}{(J_r + J_\theta + J_\phi)^3}. \quad (10.141)$$

If we evaluate the sum of the  $J$ 's in terms of the energy from Eq. (10.140) the period of the orbit is

$$\tau = \pi k \sqrt{\frac{m}{-2E^3}}. \quad (10.142)$$

This formula for the period agrees with Kepler's third law, Eq. (3.71), if it is remembered that the semimajor axis  $a$  is equal to  $-k/2E$ .

The degenerate frequencies may be eliminated by canonical transformation to a new set of action-angle variables, following the procedure outlined in the previous section. Expressing the degeneracy conditions as

$$v_\phi - v_\theta = 0, \quad v_\theta - v_r = 0,$$

the appropriate generating function is

$$F = (w_\phi - w_\theta)J_1 + (w_\theta - w_r)J_2 + w_r J_3. \quad (10.143)$$

The new angle variables are

$$\begin{aligned} w_1 &= w_\phi - w_\theta \\ w_2 &= w_\theta - w_r, \\ w_3 &= w_r, \end{aligned} \quad (10.144)$$

and, as planned, two of the new frequencies,  $v_1$  and  $v_2$ , are zero. We can obtain the new action variables from the transformation equations

$$\begin{aligned} J_\phi &= J_1, \\ J_\theta &= J_2 - J_1, \\ J_r &= J_3 - J_2, \end{aligned}$$

which yields the relations

$$\begin{aligned} J_1 &= J_\phi, \\ J_2 &= J_\phi + J_\theta, \\ J_3 &= J_\phi + J_\theta + J_r. \end{aligned} \quad (10.145)$$

In terms of these transformed variables the Hamiltonian appears as

$$H = -\frac{2\pi^2 mk^2}{J_3^2}, \quad (10.146)$$

a form involving only that action variable for which the corresponding frequency is different from zero.

If we are willing to use, from the start, our a priori knowledge that the motion for the bound Kepler problem is a particular closed orbit in a plane, then the integrals for  $J_\theta$  and  $J_r$  can be evaluated very quickly and simply. For the  $J_\theta$  integral, we can apply the following procedure. It will be recalled that when the defining equations for the generalized coordinates do not involve time explicitly, then (cf. Eq. (8.20) and the material following (8.20))

$$p_i \dot{q}_i = 2L_2 \dot{q}_i \dot{q}_i = 2T.$$

Knowing that the motion is confined to a plane, we can express the kinetic energy  $T$  either in spherical polar coordinates or in the plane polar coordinates  $(r, \psi)$ . It follows, then, that

$$2T = p_r \dot{r} + p_\theta \dot{\theta} + p_\phi \dot{\phi} = p_r \dot{r} + p_\psi \dot{\psi}, \quad (10.147)$$

where  $p$  ( $\equiv l$ ) is the magnitude of the total angular momentum. Hence, the definition for  $J_\theta$  can also be written as

$$J_\theta \equiv \oint p_\theta d\theta = \oint p d\psi - \oint p_\phi d\phi. \quad (10.148)$$

Because the frequencies for  $\theta$  and  $\phi$  are equal, both  $\phi$  and  $\psi$  vary by  $2\pi$  as  $\theta$  goes through a complete cycle of libration, and the integrals defining  $J_\theta$  reduce to

$$J_\theta = 2\pi(p - p_\phi) = 2\pi(\alpha_\theta - \alpha_\phi),$$

in agreement with Eq. (10.135).

The integral for  $J_r$ , Eq. (10.136), was evaluated in order to obtain  $H \equiv E$  in terms of the three action variables. If we use the fact that the closed elliptical orbit in the bound Kepler problem is such that the frequency for  $r$  is the same as that for  $\theta$  and  $\phi$ , then the functional dependence of  $H$  on  $J$  can also be obtained from Eq. (10.147). In effect then we are evaluating  $J_r$  in a different way. The virial theorem for the bound orbits in the Kepler problem says that (cf. Eq. (3.30))

$$\bar{V} = -2\bar{T},$$

where the bar denotes an average over a single complete period of the motion. It follows that

$$H \equiv E = \bar{T} + \bar{V} = -\bar{T}. \quad (10.149)$$

Integrating Eq. (10.147) with respect to time over a complete period of motion we have

$$\frac{2\bar{T}}{v_3} = J_r + J_\theta + J_\phi = J_3, \quad (10.150)$$

where  $v_3$  is the frequency of the motion, that is, the reciprocal of the period. Combining Eqs. (10.149) and (10.150) leads to the relation

$$-\frac{2}{J_3} = \frac{v_3}{H} = \frac{1}{H} \frac{dH}{dJ_3}, \quad (10.151)$$

where use has been made of Eq. (10.105). Equation (10.151) is in effect a differential equation for the functional behavior of  $H$  on  $J_3$ . Integration of the equation immediately leads to the solution

$$H = \frac{D}{J_3^2}, \quad (10.152)$$

where  $D$  is a constant that cannot involve any of the  $J$ 's, and must therefore depend only upon  $m$  and  $k$ . Hence, we can evaluate  $D$  by considering the elementary case of a circular orbit, of radius  $r_0$ , for which  $J_r = 0$  and  $J_3 = 2\pi p$ . The total energy is here

$$H = -\frac{k}{2r_0} \quad (10.153)$$

(as can most immediately been seen from the virial theorem). Further, the condition for circularity, Eq. (3.40), can be written for the inverse-square force law as

$$\frac{k}{r_0^2} = \frac{p^2}{mr_0^3} = \frac{J_3}{4\pi^2 mr_0^3}. \quad (10.154)$$

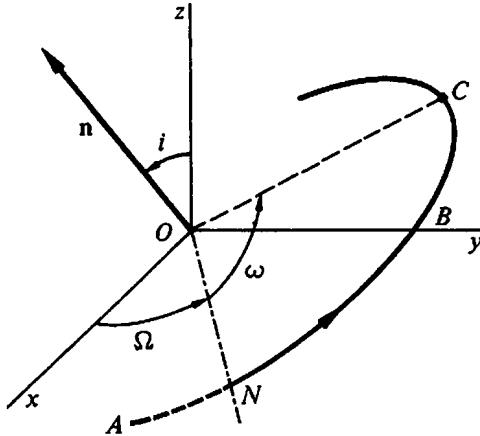
Eliminating  $r_0$  between Eqs. (10.153) and (10.154) leads to

$$H = -\frac{2\pi^2 mk^2}{J_3^2}. \quad (10.155)$$

This result has been derived only for circular orbits. But Eq. (10.152) says it must also be correct for *all* bound orbits of the Kepler problem, and indeed it is identical with Eq. (10.146). Thus, if the existence of a single period for all coordinates is taken as known beforehand, it is possible to obtain  $H(J)$  without direct evaluation of the circuital integrals.

In any problem with three degrees of freedom, there must of course be six constants of motion. It has previously been pointed out that in the Kepler problem five of these are algebraic functions of the coordinates and momenta and describe the nature of the orbit in space, and only the last refers to the position of the particle in the orbit at a given time (cf. Sections 3.7 to 3.9). It is easy to see that five parameters are needed to completely specify, say, the elliptical orbit of the bound Kepler problem in space. Since the motion is in a plane, two constants are needed to describe the orientation of that plane in space. One constant is required to give the *scale* of the ellipse, for example, the semimajor axis  $a$ , and the other the *shape* of the ellipse, say, through the eccentricity  $e$ . Finally, the fifth parameter must specify the *orientation* of the ellipse relative to some arbitrary direction in the orbital plane.

The classical astronomical *elements* of the orbit provide the orbital parameters almost directly in the form given above. Two of the angles appearing in these elements have unfamiliar but time-honored names. Their definitions, and functions as orbital parameters, can best be seen from a diagram, such as is given in Fig. 10.6. Here  $xyz$  defines the chosen set of axes fixed in space, and the unit vector  $\mathbf{n}$  characterizes the normal to the orbital plane. The intersection between the  $xy$  plane and the orbital plane is called the *line of nodes*. There are two points on the line of nodes at which the elliptical orbit intersects the  $xy$  plane; the point at which the particle enters from below into the upper hemisphere (or goes from the “southern” to the “northern” hemispheres) is known as the *ascending node*. In Fig. 10.6, the portion of the orbit in the southern hemisphere is shown, for clarity, as a dashed line. The dot-dashed line  $ON$  is a portion of the line of nodes containing the ascending node. We can measure the direction of  $ON$  in the  $xy$  plane by the angle  $xON$ , which is customarily denoted by  $\Omega$ , and is known as the *longitude*



**FIGURE 10.6** Angular elements of the orbit in the bound Kepler problem.

of the ascending node. Finally, if  $C$  denotes the point of periapsis in the orbit, then the angle  $NOC$  in the orbital plane is denoted by  $\omega$  and is called the *argument of the perihelion*.\* The more familiar angle  $i$ , introduced in Eq. (10.131), is in its astronomical usage known as the *inclination of the orbit*. One usual set of astronomical elements therefore consists of the six constants

$$i, \Omega, a, e, \omega, T,$$

where the last one,  $T$ , is the time of passage through the periapsis point. Of the remaining five, the first two define the orientation of the orbital plane in space, while  $a$ ,  $e$ , and  $\omega$  directly specify the scale, shape, and orientation of the elliptic orbit, respectively.

The action-angle variable treatment of the Kepler problem also leads to five algebraic constants of the motion. Three of them are obvious as the three constant action variables,  $J_1$ ,  $J_2$ , and  $J_3$ . The remaining two are the angle variables  $w_1$  and  $w_2$ , which are constants, because their corresponding frequencies are zero. It must therefore be possible to express the five constants  $J_1$ ,  $J_2$ ,  $J_3$ ,  $w_1$ , and  $w_2$  in terms of the classical orbital elements  $i$ ,  $\Omega$ ,  $a$ ,  $e$ , and  $\omega$ , and vice versa. Some of these interrelations are immediately obvious. From Eqs. (10.145) and (10.135) it follows that

$$J_2 = 2\pi\alpha_\theta \equiv 2\pi l, \quad (10.156)$$

and hence, by Eq. (10.131),

$$\frac{J_1}{J_2} = \cos i. \quad (10.157)$$

As is well known, the semimajor axis  $a$  is a function only of the total energy  $E$  (cf. Eq. (3.61)) and therefore, by Eq. (10.146),  $a$  is given directly in terms of  $J_3$ :

\*This terminology appears to be commonly used even for orbits that are not around the sun. The proper term for orbits about stars is periastra; for Earth-orbiting satellites, this term is perigee.

$$a = -\frac{k}{2E} = \frac{J_3^2}{4\pi^2 mk}. \quad (10.158)$$

In terms of  $J_2$ , Eq. (3.62) for the eccentricities can be written as

$$e = \sqrt{1 - \frac{J_2^2}{4\pi^2 mka}},$$

or

$$e = \sqrt{1 - \left(\frac{J_2}{J_3}\right)^2}. \quad (10.159)$$

It remains only to relate the angle variables  $w_1$  and  $w_2$  to the classic orbital elements. Obviously, they must involve  $\Omega$  and  $\omega$ . In fact, it can be shown that for suitable choice of additive constants of integration they are indeed proportional to  $\Omega$  and  $\omega$ , respectively. This will be demonstrated for  $w_1$ ; the identification of  $w_2$  will be left as an exercise.

The equation of transformation defining  $w_1$  is, by Eq. (10.127),

$$w_1 = \frac{\partial W}{\partial J_1}.$$

It can be seen from the separated form of  $W$ , Eq. (10.71), that  $W$  can be written as the sum of indefinite integrals:

$$W = \int p_\phi d\phi + \int p_\theta d\theta + \int p_r dr. \quad (10.160)$$

As we have seen from the discussion on  $J_r$ , the radial momentum  $p_r$  does not involve  $J_1$ , but only  $J_3$  (through  $E$ ) and the combination  $J_\theta + J_\phi = J_2$ . Only the first two integrals are therefore involved in the derivative with respect to  $J_1$ . By Eq. (10.130),

$$p_\phi = \alpha_\phi = \frac{J_1}{2\pi}, \quad (10.161)$$

and by Eq. (10.74), with the help of Eqs. (10.156) and (10.161),

$$p_\theta = \pm \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2 \theta}} = \pm \frac{1}{2\pi} \sqrt{J_2^2 - \frac{J_1^2}{\sin^2 \theta}}. \quad (10.162)$$

It turns out that in order to relate  $w_1$  to the ascending node, it is necessary to choose the negative sign of the square root.\* The angular variable  $w_1$  is therefore determined by

\*Note that when the particle passes through the ascending node (cf. Fig. 10.6)  $\theta$  is *decreasing* and the corresponding momentum is negative. In calculating  $J_\theta$ , it was not necessary to worry about the choice of sign because in going through a complete cycle both signs are encountered.

$$w_1 = \frac{\phi}{2\pi} + \frac{J_1}{2\pi} \int \frac{d\theta}{\sin^2 \theta \sqrt{J_2^2 - J_1^2 \csc^2 \theta}},$$

or

$$\begin{aligned} 2\pi w_1 &= \phi + \cos i \int \frac{d\theta}{\sin^2 \theta \sqrt{1 - \cos^2 i \csc^2 \theta}} \\ &= \phi + \int \frac{\cot i \csc^2 \theta d\theta}{\sqrt{1 - \cot^2 i \csc^2 \theta}}. \end{aligned}$$

By a change of variable to  $u$ , defined through

$$\sin u = \cot i \csc \theta, \quad (10.163)$$

the integration can be performed trivially, and the expression for  $w_1$  reduces to

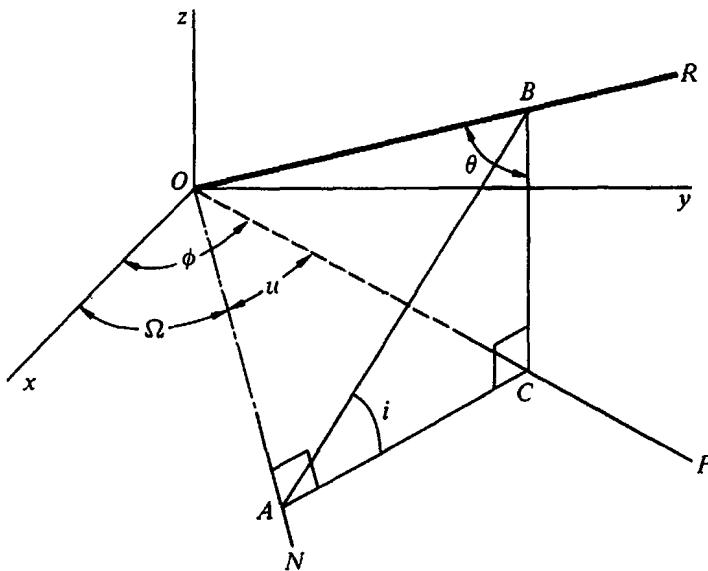
$$2\pi w_1 = \phi - u. \quad (10.164)$$

The angle coordinate  $\phi$  is the azimuthal angle of the projection on the  $xy$  plane measured relative to the  $x$  axis. Clearly, from Eq. (10.163)  $u$  is a function of the polar angle  $\theta$  of the particle. But what is its geometrical significance? We can see what  $u$  is by reference to Napier's rules\* as applied to the spherical triangle defined by the line of nodes, the radius vector, and the projection of the radius vector on the  $xy$  plane. However, it may be more satisfying to indulge in a little trigonometric manipulation and derive the relation *ab initio*. In Fig. 10.7, the line  $ON$  is the line of nodes,  $OR$  is the line of the radius vector at some time, and the dotted line  $OP$  is the projection of the radius vector on the  $xy$  plane. The angle that  $OP$  makes with the  $x$  axis is the azimuth angle  $\phi$ . We contend that  $u$  is the angle  $OP$  makes with the line of nodes. To prove this, imagine a plane normal both to the  $xy$  plane and to the line of nodes, which intersects the radius vector at unit distance  $OB$  from the origin  $O$ . The points of intersection  $A$ ,  $B$ , and  $C$  of this plane, with the three lines from the origin, define with the origin four right triangles. Since  $OB$  has unit length, it follows that  $BC = \cos \theta$  and therefore  $AC = \cos \theta \cot i$ . On the other hand,  $OC = \sin \theta$  and therefore it is also true that  $AC = \sin \theta \sin u$ . Hence,  $\sin u = \cot i \csc \theta$ , which is identical with Eq. (10.163) and proves the stipulated identification of the angle  $u$ . Figure 10.7 shows clearly that the difference between  $\phi$  and  $u$  must be  $\Omega$ , so that

$$2\pi w_1 = \Omega. \quad (10.165)$$

In a similar fashion, we can identify the physical nature of the constant  $w_2$ . Of the integrals making up  $W$ , Eq. (10.160), the two over  $\theta$  and  $r$  contain  $J_2$  and

\*For an explanation of Napier's rules for spherical triangles, see handbooks such as the *Handbook of Mathematical Tables* (Chemical Rubber Publishing Co.) or *Handbook of Applied Mathematics* (Van Nostrand-Reinhold).



**FIGURE 10.7** Diagram illustrating angles appearing in action-angle treatment of the Kepler problem.

are therefore involved in finding  $w_2$ . After differentiation with respect to  $J_2$ , the integral over  $\theta$  can be performed by the same type of trigonometric substitution as employed for  $w_1$ . The corresponding integral over  $r$  can be carried out in a number of ways, most directly by using the orbit equation for  $r$  in terms of the polar coordinate angle in the orbital plane. By suitable choice of the arbitrary lower limit of integration, it can thus be found that  $2\pi w_2$  is the difference between two angles in the orbital plane, one of which is the angle of the radius vector relative to the line of nodes and the other is the same angle but relative to the line of the perihelion. In other words,  $2\pi w_2$  is the argument of the perihelion:

$$2\pi w_2 = \omega. \quad (10.166)$$

Detailed derivation is left to one of the exercises.

The method of action-angle variables is certainly not the quickest way to solve the Kepler problem, and the practical usefulness of the set of variables is not obvious. However, their value has long been demonstrated in celestial mechanics, where they appear under the guise of the *Delaunay variables*.\* As will be seen in Section 12.2, they provide the natural orbital elements that can be used in perturbation theory, to describe the modifications of the nominal Kepler orbits produced by small deviations of the force from the inverse-square law. Many of the basic studies on possible perturbations of satellite orbits were carried out in terms of the action-angle variables.

\*As customarily defined, the Delaunay variables differ from the  $(J_i, w_i)$  set by multiplicative constants.

## DERIVATIONS

1. For a conservative system show that by solving an appropriate partial differential equation we can construct a canonical transformation such that the new Hamiltonian is a function of the new *coordinates* only. (Do not use the exchange transformation,  $F_1$ .) Show how a formal solution to the motion of the system is given in terms of the new coordinates and momenta.
2. In the text, the Hamilton–Jacobi equation for  $S$  was obtained by seeking a contact transformation from the canonical coordinates  $(q, p)$  to the constants  $(\alpha, \beta)$ . Conversely, if  $S(q_i, \alpha_i, t)$  is any complete solution of the Hamilton–Jacobi equation (10.3), show that the set of variables  $(Q_i, p_i)$  defined by Eqs. (10.7) and (10.8) are canonical variables, that is, that they satisfy Hamilton’s equations.
3. In the action-angle formalism, the arguments of Hamilton’s characteristic function are the original coordinates  $q_k$  and the action variables  $J_k$ . In the case of degeneracy, a subsequent canonical transformation is made to new variables  $(w'_i, J'_i)$  from  $(w_k, J_k)$ , in order to replace the degeneracies by zero frequencies. By considering each  $J_k$  a function of the  $J'_i$  quantities as defined by Eq. (10.126), show that it remains true that

$$\frac{\partial W}{\partial J'_i} = w'_i.$$

4. The so-called Poincaré elements of the Kepler orbits can be written as

$$\begin{aligned} w_1 + w_2 + w_3, & \quad J_\phi, \\ \frac{J_r}{\pi} \cos 2\pi(w_2 + w_1), & \quad \frac{J_r}{\pi} \sin 2\pi(w_2 + w_1), \\ \frac{J_\theta}{\pi} \cos 2\pi w_1, & \quad \frac{J_\theta}{\pi} \sin 2\pi w_1. \end{aligned}$$

Show that they form a canonical set of coordinates, with the new coordinates forming the left-hand column, their conjugate momenta being given on the right-hand side.

## EXERCISES

5. Show that the function

$$S = \frac{m\omega}{2}(q^2 + \alpha^2) \cot \omega t - m\omega q\alpha \csc \omega t$$

is a solution of the Hamilton–Jacobi for Hamilton’s principal function for the linear harmonic oscillator with

$$H = \frac{1}{2m}(p^2 + m^2\omega^2q^2).$$

Show that this function generates a correct solution to the motion of the harmonic oscillator.

6. A charged particle is constrained to move in a plane under the influence of a central force potential (nonelectromagnetic)  $V = \frac{1}{2}kr^2$ , and a constant magnetic field  $\mathbf{B}$

perpendicular to the plane, so that

$$\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$$

Set up the Hamilton–Jacobi equation for Hamilton’s characteristic function in plane polar coordinates. Separate the equation and reduce it to quadratures. Discuss the motion if the canonical momentum  $p_\theta$  is zero at time  $t = 0$ .

7. (a) A single particle moves in space under a conservative potential. Set up the Hamilton–Jacobi equation in ellipsoidal coordinates  $u, v, \phi$  defined in terms of the usual cylindrical coordinates  $r, z, \phi$  by the equations

$$r = a \sinh v \sin u, \quad z = a \cosh v \cos u.$$

For what forms of  $V(u, v, \phi)$  is the equation separable?

- (b) Use the results of part (a) to reduce to quadratures the problem of a point particle of mass  $m$  moving in the gravitational field of two unequal mass points fixed on the  $z$  axis a distance  $2a$  apart.

8. Suppose the potential in a problem of one degree of freedom is linearly dependent upon time, such that the Hamiltonian has the form

$$H = \frac{p^2}{2m} - mAtx,$$

where  $A$  is a constant. Solve the dynamical problem by means of Hamilton’s principal function, under the initial conditions  $t = 0, x = 0, p = mv_0$ .

9. Set up the plane Kepler problem in terms of the generalized coordinates

$$u = r + x,$$

$$v = r - x.$$

Obtain the Hamilton–Jacobi equation in terms of these coordinates, and reduce it to quadratures (at least).

10. One end of a uniform rod of length  $2l$  and mass  $m$  rests against a smooth horizontal floor and the other against a smooth vertical surface. Assuming that the rod is constrained to move under gravity with its ends always in contact with the surfaces, use the Hamilton–Jacobi equations to reduce the solution of the problem to quadratures.

11. A particle is constrained to move on a roller coaster, the equation of whose curve is

$$z = A \cos^2 \frac{2\pi x}{\lambda}.$$

There is the usual constant downward force of gravity. Discuss the system trajectories in phase space under all possible initial conditions, describing the phase space orbits in as much detail as you can, paying special attention to turning points and transitions between different types of motion.

12. A particle of mass  $m$  moves in a plane in a square well potential:

$$\begin{aligned} V(r) &= -V_0 & 0 < r < r_0, \\ &= 0 & r > r_0. \end{aligned}$$

- (a) Under what initial conditions can the method of action-angle variables be applied?  
 (b) Assuming these conditions hold, use the method of action-angle variables to find the frequencies of the motion.
13. A particle moves in periodic motion in one dimension under the influence of a potential  $V(x) = F|x|$ , where  $F$  is a constant. Using action-angle variables, find the period of the motion as a function of the particle's energy.
14. A particle of mass  $m$  moves in one dimension under a potential  $V = -k/|x|$ . For energies that are negative, the motion is bounded and oscillatory. By the method of action-angle variables, find an expression for the period of motion as a function of the particle's energy.
15. A particle of mass  $m$  moves in one dimension subject to the potential

$$V = \frac{a}{\sin^2\left(\frac{x}{x_0}\right)}.$$

Obtain an integral expression for Hamilton's characteristic function. Under what conditions can action-angle variables be used? Assuming these are met, find the frequency of oscillation by the action-angle method. (The integral for  $J$  can be evaluated by manipulating the integrand so that the square root appears in the denominator.) Check your result in the limit of oscillations of small amplitude.

16. A particle of mass  $m$  is constrained to move on a curve in the vertical plane defined by the parametric equations

$$\begin{aligned}y &= l(1 - \cos 2\phi), \\x &= l(2\phi + \sin 2\phi).\end{aligned}$$

There is the usual constant gravitational force acting in the vertical  $y$  direction. By the method of action-angle variables, find the frequency of oscillation for all initial conditions such that the maximum of  $\phi$  is less than or equal to  $\pi/4$ .

17. Solve the problem of the motion of a point projectile in a vertical plane, using the Hamilton–Jacobi method. Find both the equation of the trajectory and the dependence of the coordinates on time, assuming the projectile is fired off at time  $t = 0$  from the origin with the velocity  $v_0$ , making an angle  $\alpha$  with the horizontal.
18. For the system described in Exercise 12 of Chapter 6, find a linear point transformation to variables in which the Hamilton–Jacobi equation is separable. By use of the action-angle variables, find the eigenfrequencies of the system.
19. A three-dimensional harmonic oscillator has the force constant  $k_1$  in the  $x$ - and  $y$ -directions and  $k_3$  in the  $z$ -direction. Using cylindrical coordinates (with the axis of the cylinder in the  $z$  direction), describe the motion in terms of the corresponding action-angle variables, showing how the frequencies can be obtained. Transform to the “proper” action-angle variables to eliminate degenerate frequencies.
20. Find the frequencies of a three-dimensional harmonic oscillator with *unequal* force constants using the method of action-angle variables. Obtain the solution for each Cartesian coordinate and conjugate momentum as functions of the action-angle variables.

21. (a) In the harmonic oscillator of Exercise 20, allow all the frequencies to become equal (isotropic oscillator) so that the motion is completely degenerate. Transform to the “proper” action-angle variables, expressing the energy in terms of only one of the action variables.  
(b) Solve the problem of the isotropic oscillator in action-angle variables using spherical polar coordinates. Transform again to proper action-angle variables and compare with the result of part (a). Are the two sets of proper variables the same? What are their physical significances? This problem illustrates the feasibility of separating a degenerate motion in more than one set of coordinates. The nondegenerate oscillator can be separated only in Cartesian coordinates, not in polar coordinates.
22. The motion of a degenerate plane harmonic oscillator can be separated in any Cartesian coordinate system. Obtain the relations between the two sets of action-angle variables corresponding to two Cartesian systems of axes making an angle  $\theta$  with each other. Note that the transformation between the two sets is *not* the orthogonal transformation of the rotation.
23. (a) Evaluate the  $J_\theta$  integral in the Kepler problem by the method of complex contour integration. To get the integral into a useful form, it is suggested that the substitution  $\cos \theta = x \sin i$  might be made.  
(b) Verify the integration procedure used for  $J_\theta$  in the text, carrying out the final integrations in Eq. (10.134).  
(c) Follow the consequences of the inclination being greater than  $90^\circ$ , that is,  $\cos i$  negative. In particular, what are the changes in Eq. (10.135), in the canonical transformations to zero frequencies and therefore in Eqs. (10.145)? Can you write these equations in such a form that they are valid whether  $\cos i$  is positive or negative? Justify your answer.
24. Evaluate the integral for  $J_r$  in the Kepler problem by elementary means. This includes using tables of integrals, but if so, explicit and detailed references should be given to the tables used.
25. Show, but the method outlined in the text (or any other), that  $2\pi w_2$  is  $\omega$ , the argument of the periapsis, in the three-dimensional Kepler problem.
26. Set up the problem of the heavy symmetrical top, with one point fixed, in the Hamilton-Jacobi method, and obtain the formal solution to the motion as given by Eq. (5.63).
27. Describe the phenomenon of small radial oscillations about steady circular motion in a central force potential as a one-dimensional problem in the action-angle formalism. With a suitable Taylor series expansion of the potential, find the period of the small oscillations. Express the motion in terms of  $J$  and its conjugate angle variable.
28. Set up the problem of the relativistic Kepler motion in action-angle variables, using the Hamiltonian in the form given by Eq. (8.54). Show in particular that the total energy (including rest mass) is given by

$$\frac{E}{mc^2} = \frac{1}{\sqrt{1 + \frac{4\pi^2 k^2}{[(J'_3 - J'_2)c + \sqrt{J'^2_2 c^2 - 4\pi^2 k^2}]^2}}}.$$

Note that the degeneracy has been partly lifted, because the orbit is no longer closed, but is still confined to a plane. In the limit as  $c$  approaches infinity, show that this reduces to Eq. (10.146).

# CHAPTER

# 11

## Classical Chaos

We have in the previous chapters devoted most of our attention to integrable problems, that is, problems in which the equations of motion can be integrated to provide solutions in closed form. For example, in Sections 3.7 and 3.8 we found exact solutions for the two-body, inverse-square force law problem by integrations of the equations of motion. For many physical situations exact solutions cannot be found. In the next chapter we shall examine problems with potentials that can be broken into a main integrable part and a weaker additional part that renders the problem nonintegrable, but that can be taken into account by applying classical perturbation theory. A weak interaction term might, for example, couple together two equations of motion so the variables are no longer separable. The present chapter deals with some situations involving perturbations and lack of integrability that cannot be conveniently handled by classical perturbation theory.

If the interaction term is no longer “small” in the sense of classical perturbation theory (cf. Section 12.1), the solutions may become complex and differ considerably from those of the uncoupled equations. In some cases new solutions appear that cannot be generated from the uncoupled equations. These solutions are often well behaved in the sense that a small change in the initial conditions brings about only a small change in the motion. When this is the case, the solutions are referred to as regular or normal. There also exist cases in which the motion evolves in entirely different ways even for nearly identical starting circumstances. Solutions of this type are referred to as *chaotic*. It is important to point out that this chaos still involves deterministic solutions to deterministic equations. They are called chaotic because, although deterministic, they are not predictable because they are highly sensitive to initial conditions. If we consider two bounded solutions in the nonchaotic regime that start nearby within a small region of phase space, the phase space region covered by the solutions at a later time will still be relatively small and compact as expected from Liouville’s theorem (cf. Section 9.9). In the chaotic regime, the sector of phase space covered by these solutions will continually disperse in one or more directions with the passage of time.

*Chaos* is a type of motion that lies between the regular deterministic trajectories arising from solutions of integrable equations and a state of noise or unpredictable stochastic behavior characterized by complete randomness. Chaos exhibits extensive randomness tempered by some regularity. Chaotic trajectories arise from the motion of nonlinear systems, which is nonperiodic, but still somewhat predictable. Specific solutions change exponentially in response to small

changes in the initial conditions. In this chapter we shall examine some of the properties of this chaotic motion, and give examples of it.

This chapter is only an introduction to the subject of chaos; it presents the general principles that underlie chaotic motion. We begin with a discussion of periodic motion in general, and we discuss ways to transform it to circular motions in phase space. Then we add perturbations that disturb the regular motion, and examine the Kolmogorov–Arnold–Moser (KAM) theorem, which provides conditions for the breakdown of regularity. We introduce the Liapunov exponent as a quantitative measure of chaos through dispersion in phase space and use it to summarize some predictions concerning the stability of the solar system. The role played by attractors in nonchaotic motion is explained, as well as the characteristics of the strange attractor involved in chaos. Our next task is to show how to conveniently display the regularities and irregularities of motion with the aid of Poincaré sections. We then examine the motions of independent oscillators and, using the Hénon–Heiles Hamiltonian as an example, we introduce the effect of a perturbation interaction and demonstrate that orbits that are initially regular will, when subject to a continual increase in the magnitude of the perturbing coupling potential, gradually transform to a state of chaos. The logistic equation is treated in detail and used to explain *bifurcations* and *invariants*, including a universal constant associated with chaos. Some brief comments are made on nonintegral dimensionality and fractals before closing.

### 11.1 ■ PERIODIC MOTION

In Chapter 3, we discussed bounded motion with an emphasis on motion in which the orbits are closed; that is, the trajectory repeats itself every period. The simple harmonic oscillator and the Kepler problem are examples of closed periodic motion. In the latter case there are two periodicities, the radial coordinate  $r$  varies from its minimum value  $r_1$  at perihelion to its maximum  $r_2$  at aphelion and then back to perihelion during the time that the angular motion goes from  $\theta = 0$  to  $\theta = 2\pi$ . Hence, the periods for the radial and the angular motions are the same. These periods exemplify two types of motion that are degenerate. We know from Section 3.2 that the rate of change,  $\dot{\theta}$ , depends upon the radial distance  $r$

$$\dot{\theta}(t) = \frac{\ell}{mr^2}, \quad (3.8)$$

and the rate of change of  $r$  is a complicated analytical closed-form expression. The angular speed  $v_\theta = r\dot{\theta}$  depends upon the angle  $\theta$  in the manner sketched in Fig. 3.17. In Chapter 3, we showed how to integrate the equations of motion to obtain the polar coordinate equation for the orbit

$$r = \frac{a(1 - e^2)}{1 + e \cos \theta}, \quad (3.64)$$

where the origin of the angular coordinate,  $\theta = 0$ , is chosen at perihelion. Figures 3.16 and 3.17 present phase space plots in the  $v_r$  versus  $r$  and  $v_\theta$  versus  $\theta$  planes, respectively, for Kepler orbits with the same energy and different eccentricities.

In Section 10.6, we found that a convenient way to represent periodic motion is to carry out a variant of the Hamilton–Jacobi procedure and transform the Hamiltonian to action-angle variables. The new momentum, called the action variable  $J = \oint p dq$  is a constant of the motion, and the new conjugate coordinate  $w$  depends linearly upon the time:  $w = \omega t + \beta$ . We are interested in a Hamiltonian  $\mathcal{H}(q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n; t)$  of a conservative system containing several variables  $p_i, q_i$ , which exhibits bounded motion. If this Hamiltonian  $\mathcal{H}$  is transformed to a new set of canonical variables  $P_i, Q_i$  in which all of the  $Q_i$ 's are cyclic, that is,  $\mathcal{H} = \mathcal{H}(P_1, P_2, \dots, P_n; t)$ , then Hamilton's equations (8.18) can be readily integrated to provide the solution

$$Q_i(t) = w(t) = \omega_i t + \beta_i \quad P_i(t) = P_i(0) - \alpha_i, \quad (11.1)$$

where the  $2n$  constants of integration  $\beta_i$  and  $\alpha_i$  are invariants of the motion. When canonical transformations exist that provide this type of solution, then the Hamiltonian is said to be integrable. This solution is similar to the action-angle variables discussed in Chapter 10. For the motion to remain bounded, that is, confined to a finite region of phase space, the coordinates  $w(t)$ , which are growing linearly with the time, must be arguments of bounded functions, and in many cases, they will be arguments of periodic functions, as is the case with the radial variable  $r$  of Eq. (3.64) quoted above.

In Sections 10.2 and 10.7, we showed that the Hamiltonian of a harmonic oscillator can undergo a canonical transformation to conjugate coordinates and momenta with the time dependencies of Eqs. (11.1). It follows that a Hamiltonian with the coordinates  $Q_i(t)$  and  $P_i(t)$  can be transformed to that of a harmonic oscillator in standard form, with the coordinates  $q'_i, p'_i$ . For the case  $n = 2$ , this gives

$$\mathcal{H} = \frac{p_1'^2}{2m_1} + \frac{1}{2}m_1\omega_1^2 q_1'^2 + \frac{p_2'^2}{2m_2} + \frac{1}{2}m_2\omega_2^2 q_2'^2, \quad (11.2)$$

which corresponds to a system of two uncoupled harmonic oscillators with a Hamiltonian that equals the total energy

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 = E_T, \quad (11.3)$$

where we have, in action variable notation (cf. (10.94))

$$\mathcal{H}_1 = \frac{J_1\omega_1}{2\pi} = E_1 \quad \text{and} \quad \mathcal{H}_2 = \frac{J_2\omega_2}{2\pi} = E_2. \quad (11.4)$$

To visualize the motion, we can express each individual oscillator in normalized coordinates

$$p_i \Rightarrow \frac{p'_i}{(2m_i)^{1/2}} \quad \text{and} \quad q_i \Rightarrow q'_i (\frac{1}{2} m \omega_i^2)^{1/2}. \quad (11.5)$$

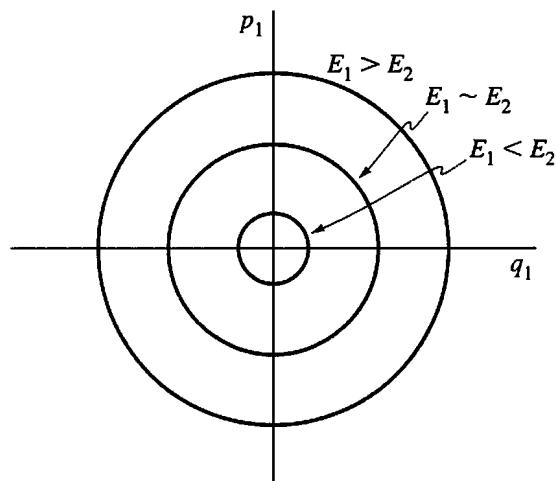
Each part  $\mathcal{H}_i$  of Hamiltonian (11.3) corresponds to the equation of a circle in its  $p_i, q_i$  plane of phase space

$$p_i^2 + q_i^2 = E_i. \quad (11.6)$$

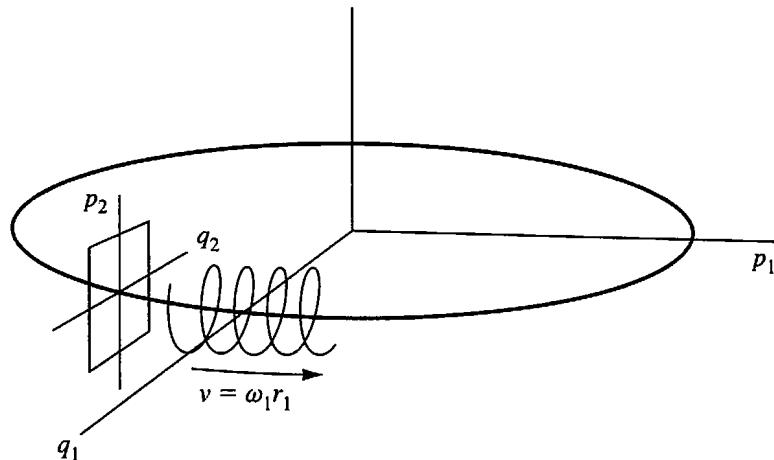
Figure 11.1 illustrates these circles by presenting constant total energy  $E_T = E_1 + E_2$  plots in the  $p_1, q_1$  plane for  $E_1 < E_2$  (small circle),  $E_1 \sim E_2$  (medium-size circle) and  $E_1 > E_2$  (large circle).

This representation of an oscillator by uniform circular motion provides us with an easy way to picture the motion associated with the double oscillator (11.2), where for convenience we select  $\omega_2 \gg \omega_1$ . Consider the movement of the low-frequency oscillator  $\omega_1$  proceeding along a circle of large radius in the  $p_1, q_1$  plane and then plot the trajectory of the high-frequency oscillator  $\omega_2$  along a small circle in a  $p_2, q_2$  plane drawn perpendicular to the circle of  $\omega_1$  and centered on its circumference, as shown in Fig. 11.2 for the case  $\omega_2 \gg \omega_1$ . The joint motion in the total phase space is a spiraling of the system point along the surface of a torus, as illustrated in the figure. If the frequency  $\omega_2$  is a multiple of  $\omega_1$ , meaning that their ratio is an integer

$$\frac{\omega_2}{\omega_1} = n, \quad (11.7)$$



**FIGURE 11.1** Circular orbits in the  $p_1, q_1$  phase space for three values of the energy ratio  $E_1/E_2$  of two uncoupled harmonic oscillators plotted for the same total energy  $E_T = E_1 + E_2$ .



**FIGURE 11.2** Circular motions of a low-frequency ( $\omega_1$ ) harmonic oscillator in the horizontal  $p_1, q_1$  plane and of a high-frequency ( $\omega_2 > \omega_1$ ) harmonic oscillator in the uniformly moving  $p_2, q_2$  vertical plane. The oscillators are uncoupled, and the resultant spiraling motion of the second oscillator generates a torus, as shown.

then the trajectory will close on itself and repeat the same pattern every period  $\tau_1 = 2\pi/\omega_1$ . More generally, if the frequencies are commensurate, meaning that  $n$  in this Eq. (11.7) is a rational number like  $\frac{2}{3}$ , then the orbit will still be closed, but it will trace out more than one path around the  $p_1, q_1$  circle before closing on itself. If, however, the frequencies are incommensurate, meaning that  $n$  in Eq. (11.7) is an irrational number, then the trajectory will never close, but will gradually cover the surface of the torus, without ever passing through exactly the same point twice. Eventually, however, it will pass arbitrarily close to every point on the surface. This is called a dense periodic orbit. Such an orbit is bounded and confined to a surface, but it is not closed.

This approach can be generalized to more than two oscillators. If there are three such oscillators with the frequencies  $\omega_1, \omega_2$ , and  $\omega_3$ , then the motion will be confined to a three-dimensional surface called a 3-torus in the six-dimensional  $p_1, p_2, p_3, q_1, q_2, q_3$  phase space. For  $N$  oscillators, there will be an  $N$ -torus in a  $2N$ -dimensional phase space. It is not easy to visualize the  $N$ -tori for  $N > 2$ .

## 11.2 ■ PERTURBATIONS AND THE KOLMOGOROV–ARNOLD–MOSER THEOREM

In the real world we can often express the dynamics of a system in terms of an integrable Hamiltonian perturbed by a small interaction that makes it nonintegrable. An example is the motion of Earth in a Keplerian orbit around the Sun primarily perturbed by the presence of the planets Mars and Jupiter. This interaction is so weak that there is very little disturbance of Earth's orbit. Weak interactions of this type are most conveniently treated with the aid of canonical perturbation theory, which is explained in detail in Chapter 12. The following outline of the method

discussed in Section 12.2 is sufficient for the consideration of chaos. References are given to the equations in Chapter 12 but reading the chapter is not necessary to follow the arguments, so we have placed this chapter first.

We assume a Hamiltonian  $\mathcal{H}$  involving a dominant interaction arising from an integrable Hamiltonian  $\mathcal{H}_0$  for which the solution is known, plus an additional interaction arising from a small perturbation term  $\Delta\mathcal{H}$

$$\mathcal{H} = \mathcal{H}_0 + \Delta\mathcal{H}. \quad (11.8)$$

It is convenient to use the generating function  $S(q, P, t) = F_2(q, P, t)$  introduced in Section 9.1 to transform the dominant Hamiltonian term  $\mathcal{H}_0$  from the phase space coordinates  $p, q$  to new coordinates  $P, Q$  of a transformed Hamiltonian  $K_0(Q, P)$ , that is identically zero, as was illustrated in the Hamilton–Jacobi approach of Chapter 10. Hamilton’s equations (10.1) for  $K_0 = 0$  provide new coordinates and momenta,  $Q_0$  and  $P_0$ , which are constants of the motion. The same transformation carried out for the total Hamiltonian,  $\mathcal{H} = \mathcal{H}_0 + \Delta\mathcal{H}_0$ , provides a transformed Hamiltonian  $\Delta K_0$ , which can be used to obtain first-order corrections  $P_1, Q_1$  to the time derivatives of the coordinates and momenta via Hamilton’s equations (cf. Equation (12.4))

$$\frac{\partial}{\partial P} \Delta K_0(P, Q) = \dot{Q}_1 \quad \frac{\partial}{\partial Q} \Delta K_0(P, Q) = -\dot{P}_1. \quad (11.9)$$

After differentiation,  $Q$  and  $P$  are replaced in  $\Delta K_0$  by their unperturbed forms, that is, by  $q = Q_0$  and  $p = P_0$ . These expressions (11.9) can be integrated over time to give the first-order determination of  $Q = Q_1$  and  $P = P_1$ . The procedure provides us with a new generating function  $S(Q_1, P_1, t)$ , and hence a new perturbed Hamiltonian  $\Delta K_1$ , which can be iterated to give the next higher-order terms  $Q_2$  and  $P_2$ , and so on. Further cycles of perturbation are obtained by iteration with the aid of the following relations (cf. Eq. (12.6)) with no summation intended:

$$\frac{\partial}{\partial P_1} \Delta K_i(P_i, Q_i) = \dot{Q}_{i+1}, \quad \frac{\partial}{\partial Q_i} \Delta K_i(P_i, Q_i) = -\dot{P}_{i+1}. \quad (11.10)$$

Thus, we have a systematic canonical iteration technique for obtaining better and better approximations to the solution when the perturbation  $\Delta\mathcal{H}$  is present. This method can be continued to higher order, as discussed in Chapter 12.

We have seen that perturbation theory provides us with a solution when  $\Delta\mathcal{H}$  is small relative to  $\mathcal{H}_0$ , but the question arises as to whether the perturbed solution is stable, and whether or not the orbits will remain close to the unperturbed ones over long periods of time. Large perturbations can clearly disturb the regular motion. A theorem known as the *Kolmogorov–Arnold–Moser (KAM) theorem* provides the conditions for the breakdown of regularity. This theorem tells us that

*If the bounded motion of an integrable Hamiltonian  $\mathcal{H}_0$  is disturbed by a small perturbation,  $\Delta\mathcal{H}$ , that makes the total Hamiltonian,  $\mathcal{H} = \mathcal{H}_0 + \Delta\mathcal{H}$ , nonintegrable and if two conditions are satisfied:*

- (a) the perturbation  $\Delta\mathcal{H}$  is small, and
- (b) the frequencies  $\omega_i$  of  $\mathcal{H}_0$  are incommensurate,

then the motion remains confined to an  $N$ -torus, except for a negligible set of initial conditions that result in a meandering trajectory on the energy surface.

Thus, the perturbed orbits will be stable, only slightly altered in shape, and localized in the same region as the unperturbed ones. Another way to say this is to observe that for a perturbation of the Hamiltonian that is sufficiently small, most quasi-periodic orbits will only experience minimal changes. The method of proof for this theorem was originally suggested by Kolmogorov in 1954, and the proofs themselves, approached from different viewpoints, were worked out independently by Arnold and by Moser a decade later. A great deal of mathematical sophistication is needed for the proof, and references can be consulted for details.\* For example, the second condition (b) of the theorem is mathematically more complex than simple incommensurability.

The caveat “except for a negligible set of initial conditions” introduces the possibility of initial conditions for which the theorem does not hold. This is analogous to the case of a differential equation with well-behaved solutions over an entire domain except for one or more singular points where the solutions blow up to infinity. The exceptions are so few that they have very little effect on applications. Chaos can occur when KAM does not hold.

### 11.3 ■ ATTRACTORS

The previous section was concerned with an integrable Hamiltonian  $\mathcal{H}_0$  being disturbed by a small perturbation  $\Delta\mathcal{H}$ . We found that stable orbits of  $\mathcal{H}_0$  persist as slightly modified but still stable orbits of the total Hamiltonian,  $\mathcal{H} = \mathcal{H}_0 + \Delta\mathcal{H}$ . Another case to consider is that of a system in which the initial conditions start the motion on a trajectory that does not lie on a stable path but that evolves toward a particular *fixed point* in phase space or toward a stable orbit in phase space called a *limit cycle*. A fixed point of this type as well as a limit cycle are examples of *attractors*.

In general, an *attractor* is a set of points in phase space to which the solution of an equation evolves long after transients have died out. It might be a point with dimension  $d_A = 0$ , a trajectory or limit cycle orbit (cf. Fig. 11.1) with dimension  $d_A = 1$ , or perhaps a toroidal surface or torus with dimension  $d_A = 2$ . For a regular attractor, the attractor dimension,  $d_A$ , is an integer that is less than the overall dimensions of the phase space. In higher dimensions, the attractors can be  $N$ -dimensional tori, where  $d_A = 2$  for the torus generated by the orbit in Fig. 11.2. There also exist somewhat bizarre types of attractors called *strange attractors*,

\*See, for example, H. Bai-Lin, *Chaos*, Singapore: World Science, 1984; E. A. Jackson, *Perspectives of Nonlinear Dynamics*, Cambridge, England: Cambridge University Press, 1989; L. E. Reichl, *The Transition to Chaos*, Berlin: Springer-Verlag, 1992.

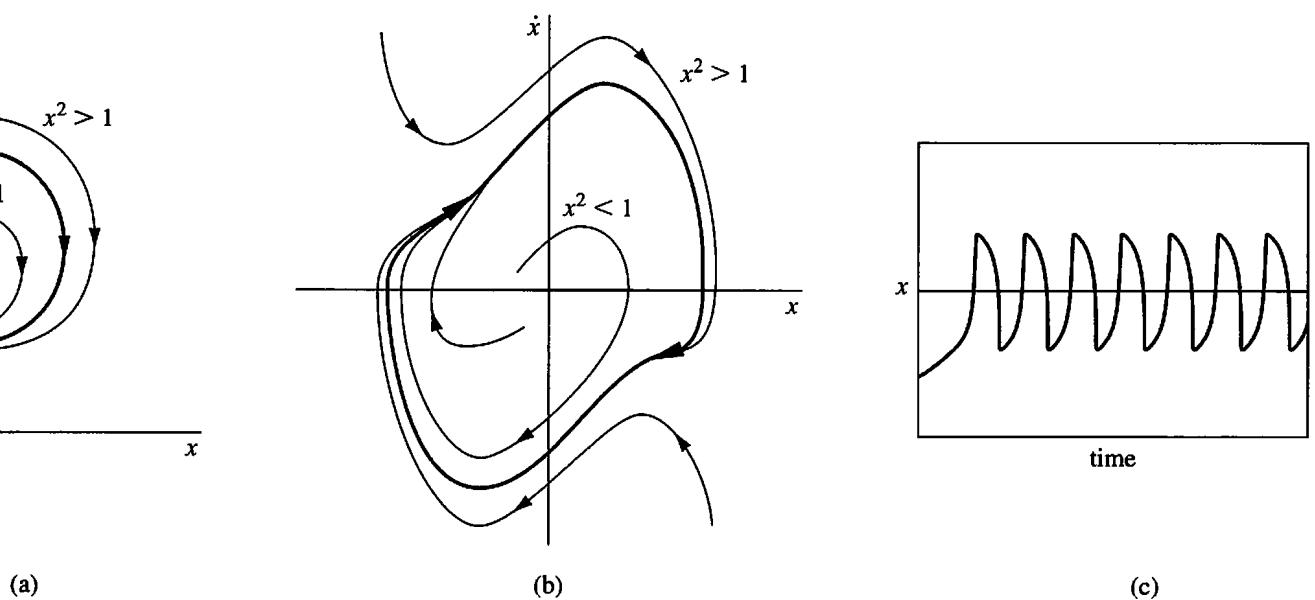
associated with chaos, which tend to be widely dispersed rather than localized in phase space. In addition, they have fractal dimensions—in other words, dimensions that are fractions or irrational numbers rather than whole numbers. These properties, as well as the term *fractal dimension*, are counterintuitive. We shall clarify the meanings of strange attractors and fractal dimensions later in the chapter.

An example of a fixed-point attractor is the equilibrium position of a pendulum at rest. If the pendulum is oscillating while subject to the action of a weak frictional drag force, then successive oscillations will decrease in amplitude until the pendulum finally comes to a stop at its equilibrium position. We say that the motion is drawn to the attractor. If the drag force is a perturbation on the main Hamiltonian, then the motion is underdamped and the pendulum undergoes many oscillations before stopping at the attractor point. If the damping term exceeds the main Hamiltonian term, then the motion is overdamped and the pendulum falls to rest without undergoing any oscillations. Either way, the motion of the pendulum finds its way to the attractor. Being a point, it is clear that the dimensionality of this attractor is zero;  $d_A = 0$ .

An example of a limit cycle type of attractor is provided by the van der Pol equation,

$$m \frac{d^2x}{dt^2} - \epsilon(1 - x^2) \frac{dx}{dt} + m\omega_0^2 x = F \cos \omega_D t, \quad (11.11)$$

which has been employed to describe oscillations in mechanical and electrical systems, as well as cardiac rhythms. If we set  $\epsilon = 0$ , then we have a driven simple harmonic oscillator with a resonant frequency  $\omega_0$  and a driving frequency  $\omega_D$ . If  $\omega_D$  is close to  $\omega_0$ , then the motion repeats itself at the frequency  $\omega_D$  of the applied force. If  $F = 0$ , then the motion will be simple harmonic at the resonant frequency  $\omega_0$ . If the small damping term  $\epsilon(1 - x^2) dx/dt$  is included in the equation, then the motion will be drawn toward the limit cycle, which in this case is a circle of unit radius. If  $x^2 > 1$ , the damping is positive and the motion spirals inward toward the limit cycle, while for  $x^2 < 1$ , the damping is negative and the motion spirals outward toward the limit cycle. Both cases are shown in Fig. 11.3a. The final state of motion has long-term stability since the damping vanishes for  $x = 1$ , and the system point moves along the circular path, which by its nature has dimension  $d_A = 1$ . If  $\epsilon$  is large enough, the damping term becomes comparable in magnitude to the other terms in the equation of motion, and the damping still draws the trajectories toward the limit cycle, but the cycle itself becomes distorted from a circular shape, as shown in Fig. 11.3b. The distortion in shape does not change the dimension of the path, which remains  $d_A = 1$ . In addition, the strong damping causes the previously simple harmonic oscillations  $x = \sin \omega_0 t$  to decrease in frequency and become distorted, as shown in Fig. 11.3c. For very large damping, the shape approximates a square wave.



**FIGURE 11.3** Limit cycles (darkened curves) of the van der Pol equation in the  $\dot{x}, x$  phase space showing (a) circular motion for a small damping coefficient  $\epsilon$ , and (b) distorted curve for large damping. Approaches to the limit cycles via orbits outside and inside them are shown. Part (c) sketches the distorted sine wave obtained for the case of appreciable damping (large  $\epsilon$ ).

## 11.4 ■ CHAOTIC TRAJECTORIES AND LIAPUNOV EXPONENTS

The orbits that we have discussed thus far have been well behaved, and confined to a relatively small region of phase space. Examples are the ellipses of the Kepler problem, the circles of the simple harmonic oscillator, and the limit cycle of the van der Pol equation (11.11). Under certain conditions, trajectories, called chaotic trajectories, will be encountered in which the motion wanders around an extensive and perhaps irregularly shaped region of phase space in a manner that appears to be random, but that in fact is tempered by constraints. This path or region where the meandering takes place is an example of a strange attractor. It is called strange because of its (fractal) geometry and chaotic because of its dynamics.\* The chaotic trajectory roams here and there, back and forth through this strange attractor region seeming to fill the space, but without ever actually passing through the same point twice. In short, *chaotic motion* has affinities with ergodic motion (cf. Section 9.8), with characteristics between regular deterministic trajectories and totally random roaming.

The motion involved in chaos has the properties of mixing, dense quasi-periodic orbits, and sensitivity to initial conditions. The properties are as follows. Mixing means that if we choose two arbitrarily small but nonzero regions,  $I_1$  and  $I_2$ , of the domain of the motion and we follow an orbit that passes through region  $I_1$ , then it will eventually pass through region  $I_2$ . The orbits are quasi-periodic

\*See A. B. Çambel, *Applied Chaos Theory*, New York: Academic Press, 1993, p. 70.

in the sense that they repeatedly and irregularly pass through the whole range of the domain without ever closing on themselves, and without any particular time period associated with successive transits. They are dense because they pass through or arbitrarily close to every point of the domain, a property that conforms with the ergotic hypothesis (cf. Section 9.8). A chaotic orbit that visits and revisits (that is, mixes with) all regions of the available phase space is identified with what is called a *strange attractor*. Its association is not with a localized attractor such as a fixed point or a limit cycle, but rather with a very extended region of phase space, hence the designation strange. The property of ergodicity, which involves covering all accessible regions of a domain, is shared by incommensurate non-chaotic orbits with respect to an ordinary attractor (for example, a torus), and by chaotic orbits with respect to a strange attractor.

Sensitivity to initial conditions means that a small change in the initial conditions can result in a large change in position and velocity many transits or iterations later. For example, a small change can convert a parabolic orbit of the Kepler problem to either a weakly bound elliptic orbit or to a hyperbolic orbit that extends to infinity. In the Hénon–Heiles Hamiltonian, (cf. Section 11.6), a small increase in the energy can induce the onset of chaos with the Liapunov exponent (defined below) giving the time scale for this breakdown of order.

The KAM theorem of the previous section is valid for small perturbations. As the perturbation increases, the effect on the motion of the system becomes more and more pronounced. If the perturbation becomes sufficiently large, the behavior may become chaotic. Then successively calculated orbits move farther and farther away from each other. Even if the first few orbits of a chaotic sequence lie relatively close to the original one, each iteration involves a greater recession than the previous one, so the extent to which they move apart can increase exponentially with the number of iterations. An example is a spaceship in an Earth orbit. A small rocket boost will move it to a nearby orbit whereas a strong boost could throw it out of orbit, heading for outer space. Another common example of how linear and chaotic motions differ when periodicity is not present is turbulence in water. While there is streamline flow, two nearby points in the water stay close together as they move along; after the onset of turbulence the same two points, on average, keep moving farther and farther apart.

A quantitative measure of this exponential divergence is a coefficient,  $\lambda$ , called a *Liapunov exponent*, (sometimes spelled Lyapunov or Ljapunov). In the chaotic region of many systems, if two orbits are separated by the small distance  $s_0$  at the time  $t = 0$ , then at a later time  $t$  their separation is given by

$$s(t) \sim s_0 e^{\lambda t}. \quad (11.12)$$

If  $\lambda > 0$  the motion is chaotic, and the Liapunov exponent  $\lambda$  quantifies the average growth of an infinitesimally small deviation of a regular orbit arising from a perturbation. It sets a time scale  $\tau \sim 1/\lambda$  for the growth of divergences brought about by sufficiently large perturbations. The chaos becomes appreciable for  $t \gg \tau$  when the trajectory winds its way around the extensive, but bounded, phase space

of the strange attractor. Eventually the separation  $s(t)$  becomes comparable to the dimensions of the accessible coordinate space so it can no longer increase further, and from that point on the separations  $s(t)$  vary randomly in time.

If the system evolves by an iterative process rather than by a temporal process then Eq. (11.12) assumes the form

$$s(n) \sim s_0 e^{n\lambda}, \quad (11.13)$$

where  $n$  is the number of iterations, and the exponent  $\lambda$  is now dimensionless. Moreover, this divergence of orbits is not reversible. In a chaotic region it is impossible to reconstruct the distant past history of a system from its present state. This means that current trajectories can no longer be projected back to determine the initial configuration.

If the Liapunov exponent is negative it measures the rate at which a system point approaches a regular attractor. In other words, in the nonchaotic region  $\lambda < 0$  and the distance  $s(t)$  from an attractor at time  $t$  is given by the expression

$$s(t) \sim s_0 e^{-|\lambda|t} \quad (11.14)$$

where  $s_0$  is the initial distance at time  $t = 0$ . For an iterative process we have the analogous expression

$$s(n) \sim s_0 e^{-n|\lambda|} \quad (11.15)$$

for the distance  $s(n)$  after  $n$  iterations. A negative exponent characterizes the rate at which the orbit spirals into the circle on Fig. 11.3a. In the previously considered damped pendulum case the time constant  $\tau$  of the damping process is the reciprocal of the associated negative Liapunov exponent,  $\tau \sim 1/|\lambda|$ .

As an example, consider the elliptic orbit of a planet in the solar system that is perturbed by the gravitational interaction with another planet. The perturbation is nonlinear, and it is also small since the gravitational interactions of the two planets with the much larger Sun are dominant. We might expect that the KAM theorem would predict that any perturbed orbit is stable, but this is not correct for two reasons. First, many natural frequencies in the solar system correspond to resonances involving individual planets and asteroids. Second, many of the objects in the solar system are asteroids, and perturbations resulting from their presence no longer remain small. Both of these effects lead to chaotic results. Some of this chaos simply means that we cannot make exact predictions about the future. Other effects may lead to the eventual ejection of one or more bodies from bound orbits, a possibility that was mentioned in Section 3.12 on the three-body problem.

When we consider natural frequencies, it is not only the orbital periods that are important. The rotation, obliquity (axial tilt), rotational plane, orbital plane, and eccentricity provide some of the other frequencies that may interact in surprising ways. The massive planets of the outer solar system have apparently settled into quasi-periodic orbits of marginal stability. Marginal stability means that their or-

bital motion is stable on a time scale comparable with the age of the solar system. Other orbital parameters occasionally change. The obliquity of Earth's axis is apparently stabilized by the presence of the Moon. Both Venus and Earth interact in a bounded chaotic fashion with little change in their periods. Mercury, Mars, Pluto, and many asteroids may undergo much more chaotic motion.

Calculations, projecting motions for the next 100 Gyr, show that there is a finite probability that Mercury will be ejected or collide with Venus some time during the next 3.5 Gyr. Using the approximation  $\tau \sim 1/|\lambda|$  with  $\tau = 3.5$  Gyr provides a Liapunov exponent  $\lambda \sim 3 \times 10^{-10}$  per year as the time scale for planetary chaos. The eccentricity of the orbit of Mars could increase to 0.2, while its axial tilt can vary by  $60^\circ$ , perhaps sufficient to release water on the surface through the possible melting of its ice caps. Pluto also has chaotic motions, but they seem to be bounded. Thus, chaos has been a mechanism for the reorganization of the planetary bodies since the formation of the solar system.

Motions in both the outer ( $> 2.8$  AU) and inner ( $< 2.5$  AU) asteroid belts are chaotic. The outer belt chaos is dominated by Jupiter and the Jupiter–Saturn–asteroid interactions, while the inner belt chaos involves Mars and Mars–Jupiter–asteroid resonances. These interactions provide a steady impetus for Mars crossing asteroids. Once established along such a path, the Liapunov exponent is much larger, leading to changes in orbit.

We must note that these conclusions are based upon the results of numerical calculations. Every effort has been made to ensure that current limits of numerical accuracy, as well as the inclusion or exclusion of members of the solar family, do not affect the conclusions. Although there is evidence of past chaos in the solar system, we must remember that our future predictions are based upon our model of the solar system, not the system itself. Stability could be better or worse than the model predicts, but the chaos itself is definitely present.

## 11.5 ■ POINCARÉ MAPS

In Section 11.1, we discussed the periodic motion of uncoupled oscillators. When two one-dimensional oscillators become coupled by adding a term such as  $x^2y$  to the Hamiltonian, then the motion becomes rather complex in the four-dimensional  $p_x x p_y y$  phase space, and it is no longer feasible to follow the trajectories. It is more convenient to sample the motion at regular intervals and use the resulting information to deduce some of its general characteristics. A convenient way to sample the motion is to map it on a cross section of phase space.

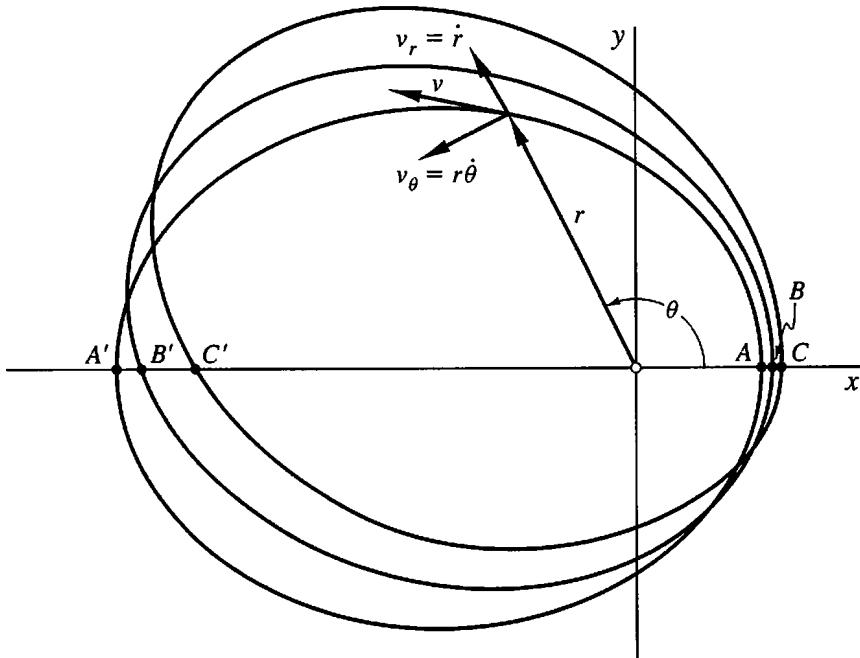
When the total energy,  $E_T$ , of a double oscillator is fixed, the dimensionality of the space is lowered by one, and the motion is confined to a three-dimensional region in this phase space called an *energy hypersurface*. Some authors refer to it as a “three-dimensional energy surface.” To avoid the complications of tracing out orbits wandering around this three-dimensional region, it is more advantageous to study a two-dimensional slice or section through the hypersurface. The slice is called a *Poincaré section*. We calculate the positions of points where or-

bits pass through the section. A convenient choice for this section is either the  $p_x x$  or the  $p_y y$  plane. Since the equations of motion are known via Hamilton's equations (8.18), the positions where successive orbits pass through this two-dimensional section can be calculated. For bounded motion, such sequences of points map out closed curves. The paths on the section defined by these points constitute what is called a *Poincaré map*.

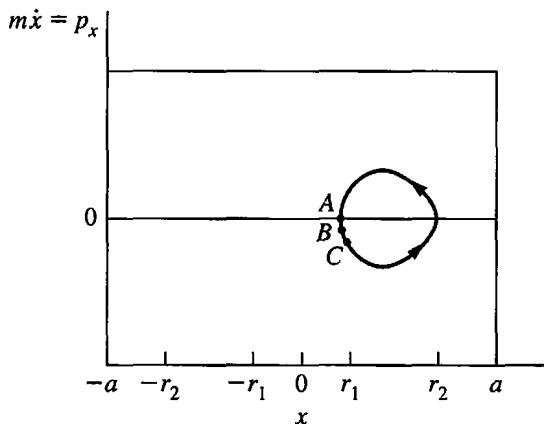
As an example of the determination of a Poincaré map, consider the Kepler problem that was solved in Section 3.7 for the case of negative energies. We now reexamine this problem using Cartesian coordinates  $x$ ,  $p_x = m\dot{x}$ ,  $y$  and  $p_y = m\dot{y}$ , taking into account a perturbation that causes the elliptical orbit to precess in the  $xy$  (that is, in the  $r, \theta$ ) coordinate space plane, as shown in Fig. 11.4. The energy  $E$  is conserved with the value

$$E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 - k(x^2 + y^2)^{-1/2}. \quad (11.16)$$

On this figure we imagine a vertical plane located at the position  $y = 0$ , with the vertical ordinate  $p_x$  axis and the horizontal abscissa  $x$  axis shown in Fig. 11.5. To calculate a Poincaré map on this  $p_x x$  cross section, we start the motion ( $t = 0$ ) at the perihelion point  $A$  of Fig. 11.4 with the initial values  $x = r_1$ ,  $y = 0$ ,  $\dot{x} = 0$ , and the velocity component  $\dot{y}$  a maximum value determined by Eq. (11.16). The polar coordinates for this starting point are  $r = r_1$  and  $\theta = 0$ . The equations of



**FIGURE 11.4** Precessing elliptic orbits of the Keplerian problem sketched in Cartesian coordinate space. The figure shows the vector velocity  $v$  tangent to the orbit at a point  $(r, \theta)$ , together with its radial ( $\dot{r}$ ) and angular ( $r\dot{\theta}$ ) components. Points  $A$ ,  $B$ , and  $C$  along the  $x$  axis near perihelion denote successive penetrations of orbits through the  $x$ ,  $x$  Poincaré section of Fig. 11.5 located along the  $x$  axis where  $y = 0$ .



**FIGURE 11.5** A  $p_x x$  Poincaré section for the Kepler problem with the solid curve on the right tracing out the orbit generated by points  $A, B, C, \dots$  of the precessing ellipse of Fig. 11.4. Points  $A', B', C'$  are not shown but are located at negative values of  $x$ .

motion are used to calculate successive points that trace out the orbit. Every time the orbit passes through the  $p_x x$  section, a point is marked on it indicating the value of  $p_x$ . Since the orbit is fixed for the unperturbed Kepler problem, the orbit will always pass through the same two points on the section, point  $A$  going from back to front and  $A'$  going from front to back, with  $p_x = 0$  for both points, as indicated in Fig. 11.5. Poincaré maps generally only show points going through the section in one direction, which does not include point  $A'$ , so this Poincaré map consists of only one point  $A$ . When the perturbation is taken into account, perhaps arising from the attractive forces of other planets on Earth as it travels around the Sun, then the orbit can precess in time, in the fashion of Fig. 11.4. Successive orbits pass through the  $x$  axis at different orbital distances indicated by points  $A, B, C, \dots$  on Fig. 11.4. These points map onto the  $p_x x$  section at the positions indicated in Fig. 11.5, and trace out the solid curve called the Poincaré map on the right side of the figure. The amount of precession that takes place for each cycle has been greatly exaggerated on these figures.

We have seen that in a four-dimensional phase space a Poincaré section is a two-dimensional slice through a three-dimensional constant-energy hypersurface. More generally, a Poincaré section is a  $2N-2$  dimensional slice through a  $2N-1$  dimensional constant energy hypersurface in a  $2N$  dimensional phase space. Although the concept of a Poincaré section is defined for these higher dimensions, its main usefulness is for the  $N = 2$  case where it provides a two-dimensional representation of the orbits, which is easy to visualize. For  $N > 2$ , it is not nearly as easy to visualize the orbits.

## 11.6 ■ HÉNON-HEILES HAMILTONIAN

Over three decades ago, M. Hénon and C. Heiles were investigating the motion of stars about the galactic center. Two constants of the motion are the vector angular

momentum  $\ell$  and the scalar energy  $E$ . The observed motions of stars near the Sun suggested that one additional constraint might, under certain conditions, restrict the possible motions. Under other energy conditions, however, the motion is not restricted, so only the two standard constants the angular momentum  $\ell$  and the energy  $E$  are available. Rather than solve this problem with the the actual potential of the galaxy, which is relatively unmanageable, Hénon and Heiles restricted the motion to the  $xy$  plane, as in the Kepler problem, and studied a relatively simple analytic potential  $V(x, y)$  that illustrates the general features of the problem.\* This potential, called the *Hénon–Heiles potential*, provides two cubic perturbation terms, which couple together two standard harmonic oscillators, corresponding to the Hamiltonian,

$$\mathcal{H} = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{1}{2}k(x^2 + y^2) + \lambda \left( x^2y - \frac{1}{3}y^3 \right), \quad (11.17)$$

where the coefficient  $\lambda$  is small so the last term serves as a perturbation. These cubic terms prevent the equations of motion from being integrated in closed form. When this Hamiltonian is expressed in polar coordinates  $x = r \cos \theta$ ,  $y = r \sin \theta$  the perturbation potential exhibits threefold symmetry,

$$\mathcal{H} = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{1}{2}kr^2 + \frac{1}{3}\lambda r^3 \sin 3\theta. \quad (11.18)$$

To simplify their computer calculations, Hénon and Heiles set  $p_x = m\dot{x}$  and  $p_y = m\dot{y}$ , expressed the Hamiltonian in normalized form using dimensionless units, and set it equal to a dimensionless energy  $E$ , with  $\lambda = 1$ ,

$$E = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 + \frac{1}{2}x^2 + \frac{1}{2}y^2 + x^2y - \frac{1}{3}y^3. \quad (11.19)$$

The equations of motion, which may be obtained from either Lagrange's equations or Hamilton's equations,

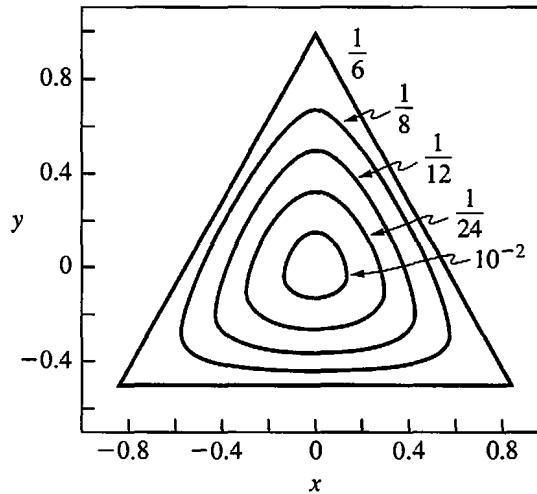
$$\begin{aligned} \ddot{x} &= -x - 2xy \\ \ddot{y} &= -y - x^2 + y^2, \end{aligned} \quad (11.20)$$

are coupled together and nonlinear, so there is no solution in closed form. We can see from the form of the dimensionless potential energy expressed in polar coordinates,

$$V(r, \theta) = \frac{1}{2}r^2 + \frac{1}{3}r^3 \sin 3\theta, \quad (11.21)$$

that for a particular value of  $V$ , the radial coordinate  $r$  attains its maximum value for  $\sin 3\theta = -1$  (that is, for  $\theta = 90^\circ, 210^\circ, 330^\circ$ ), and it attains its minimum

\*M. Hénon, *Numerical Exploration of Hamiltonian Systems*, Course 2 in Chaotic Behavior of Deterministic Systems, at the 1981 *Les Houches École D'Été de Physique Théoretique*, Session 36, G. Iooss, R. H. G. Helleman, and R. Stora (eds.), New York: North Holland, 1983.



**FIGURE 11.6** Hénon–Heiles equipotentials labeled with their dimensionless energies  $E$  plotted on the  $y, x$  plane. Closed curves for energies  $E \leq \frac{1}{6}$  reduce to an equilateral triangle for the limit  $E = \frac{1}{6}$ . Open curves outside the triangle (not shown) exist for higher energies. Adapted from M. Hénon (1983), Fig. 19.

value for  $\sin 3\theta = +1$  (that is, for  $\theta = 30^\circ, 150^\circ, 270^\circ$ ). Figure 11.6 presents equipotential curves (that is, curves of constant  $V$ ) drawn for several values of the energy  $E$ . For the limit  $E \ll \frac{1}{6}$ , not represented in the figure, the cubic perturbation terms  $x^2y - \frac{1}{3}y^3$  are negligible relative to the quadratic harmonic oscillator potential terms,  $\frac{1}{2}(x^2 + y^2)$ , and the curves closely approximate circles centered at  $x = y = 0$ . When the cubic terms are appreciable for  $E < \frac{1}{6}$ , the equipotentials form closed curves as shown, and for  $E = \frac{1}{6}$ , the curve becomes an equilateral triangle with  $r_{\max}/r_{\min} = 2$ . For energies exceeding  $\frac{1}{6}$ , the equipotentials (not shown) lie beyond the equilateral triangle, are open, and diverge to infinity. Thus, the magnitude of the energy determines whether or not the cubic terms constitute a perturbation or serve as main potential terms.

When the energy is fixed at a value  $E < \frac{1}{6}$ , the sum of the terms in the Hamiltonian must be equal to  $E$ , which means that the kinetic and potential terms both satisfy the inequalities

$$\begin{aligned} V(x, y) &\leq E \\ \frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 &\leq E, \end{aligned} \tag{11.22}$$

because the potential is positive definite. The first inequality tells us that any trajectory started inside the closed equipotential curve  $V(x, y) = E$  must remain entirely within that line, the second inequality sets limits to the allowed kinetic energy, and the overall effect is to restrict the motion to a finite region in four-dimensional phase space. To help us visualize what is happening, we examine Poincaré sections in the  $\dot{y}, y$  plane located at  $x = 0$ . The accessible region in such

a section lies within the limits set by letting  $x = 0$  and  $\dot{x} = 0$  in Eq. (11.19),

$$\frac{1}{2}\dot{y}^2 + \frac{1}{2}y^2 - \frac{1}{3}y^3 = E. \quad (11.23)$$

The maximum velocity  $\dot{y}$  occurs at  $y = 0$ , and the extrema of the coordinate  $y$  are found by solving the cubic equation (11.23) with  $\dot{y}$  set equal to zero.

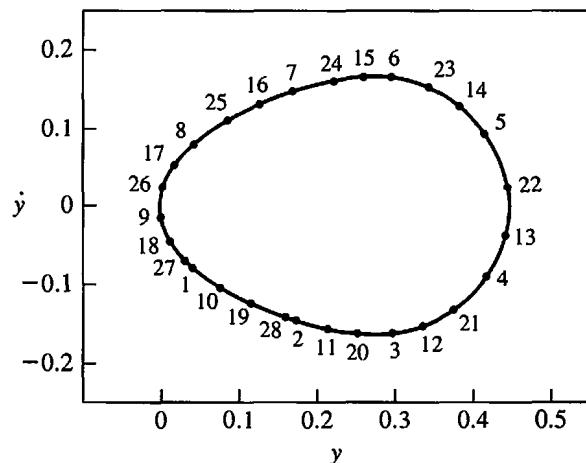
To give an example of the calculation of a Poincaré map on a  $\dot{y}, y$  section located at the position  $x = 0$  in phase space, we follow Hénon and Heiles and select the energy  $E = \frac{1}{12}$  and the values  $\dot{y}_1 = -0.08$ ,  $y = 0.02$  as starting points for the calculation. The initial velocity,  $\dot{x}_1$ , is fixed by Eq. (11.19) with  $x = 0$

$$\dot{x}_1 = \left( 2E - \dot{y}_1^2 - y_1^2 + \frac{2}{3}y_1^3 \right)^{1/2}, \quad (11.24)$$

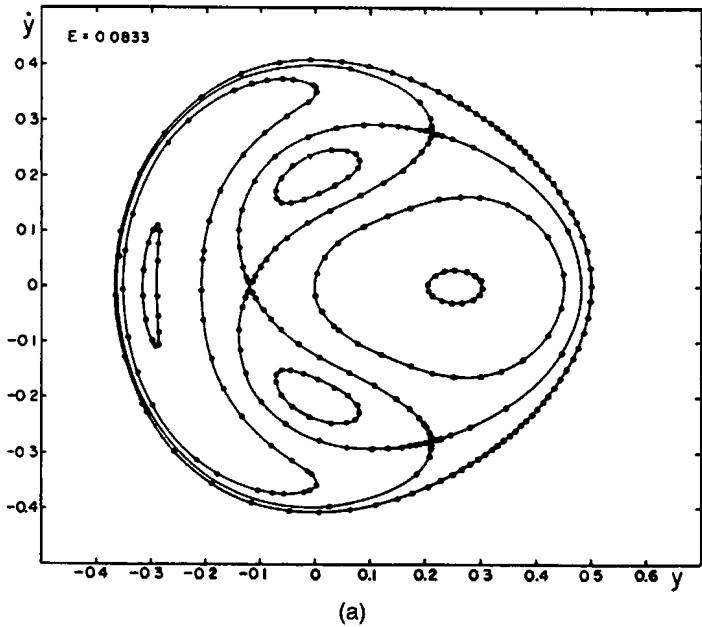
where we set  $x_1 = 0$  since the starting point is on the section. A numerical calculation provides the sequence of points  $(\dot{y}_2, y_2)$ ,  $(\dot{y}_3, y_3)$ ,  $(\dot{y}_4, y_4)$ , ..., which are labeled 2, 3, 4, ... on the Poincaré map of Fig. 11.7. The first eight points lie on a closed curve, as shown, the next nine points retrace this same curve, as do the subsequent points 18, 19, 20, .... A number of trajectories that were calculated by Hénon–Heiles for the same energy and different starting points are displayed in Fig. 11.8(a).

Note that Fig. 11.7 provides an enlargement of the large oval curve on the right side of Fig. 11.8(a). The outermost curve of the latter figure marks the boundary of the accessible region defined by solutions to Eq. (11.23). For  $E = \frac{1}{12}$ , the velocity  $\dot{y}$  reaches its maximum value  $\dot{y} = \pm(2E)^{1/2} = 0.408$  at the position  $y = 0$ , and the coordinate  $y$  attains its extremal values at the velocity  $\dot{y} = 0$  given by the two roots to the cubic equation (11.23),

$$y = \frac{1}{2}, \quad -\frac{1}{2}(\sqrt{3} - 1) \quad (11.25)$$



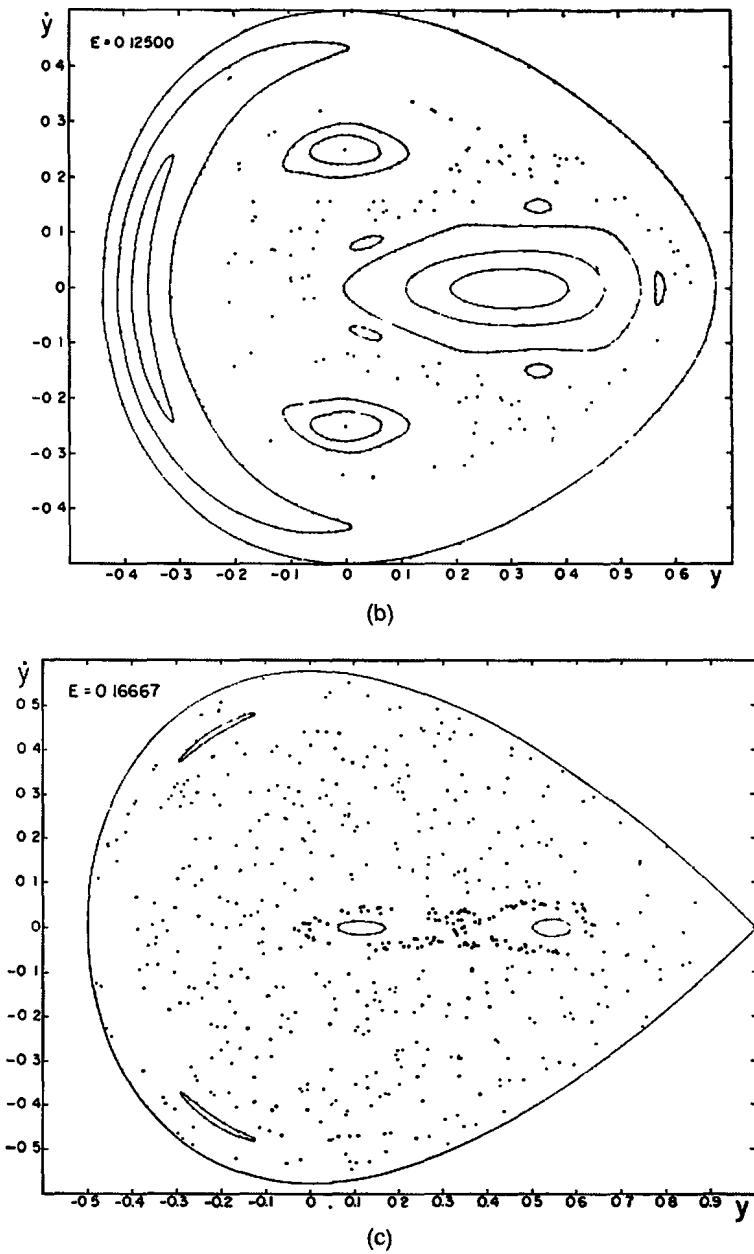
**FIGURE 11.7** Poincaré section in the  $\dot{y}, y$  plane showing the successive points 1, 2, 3, ... of a Hénon–Heiles orbit for the energy  $E = \frac{1}{12}$ . This particular curve also appears on the right side of Fig. 11.8a. From M. Hénon (1983), Fig. 20.



**FIGURE 11.8(a)** Poincaré maps in the  $\dot{y}$ ,  $y$  plane showing several Hénon–Heiles orbits:  
(a)  $E = \frac{1}{12}$  with regular orbits.

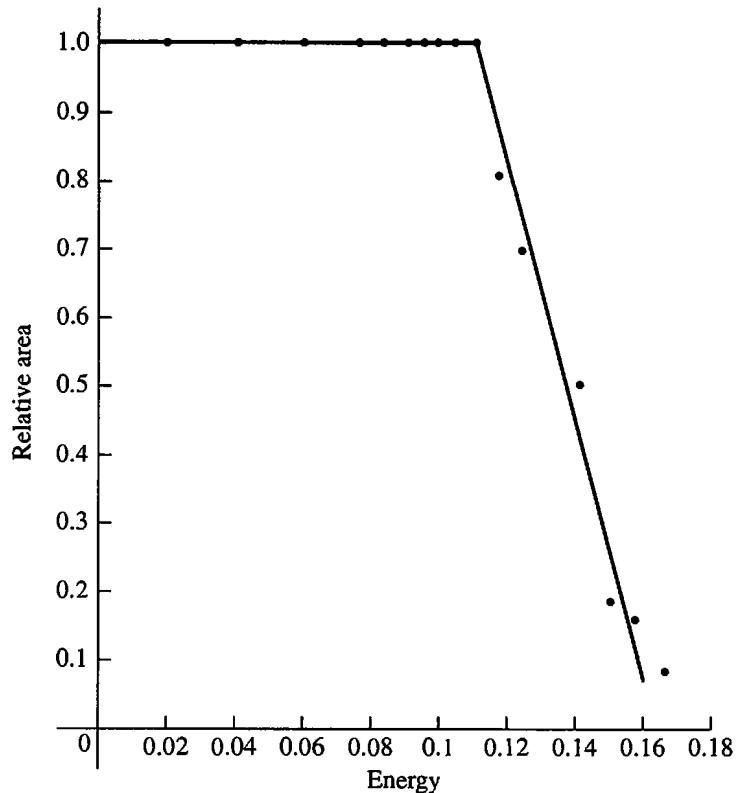
as indicated in Fig. 11.8(a). The third root of the cubic equation  $+\frac{1}{2}(\sqrt{3} + 1)$  violates condition (11.22), so it is not acceptable. The figure shows that there are four regions with oval-shaped orbits, which (if calculated for smaller and smaller circumferences) would shrink to four fixed points called elliptic fixed points. Separating and bounding these regions of elliptic type closed orbits is a single continuous curve that crosses itself three times at what are called hyperbolic points. A horizontal line drawn for  $\dot{y} = 0$  is a line of mirror symmetry with the curves above this line being mirror images of those below it. This symmetry results from the Hamiltonian being invariant under the transformation  $\dot{y} \rightarrow -\dot{y}$ , but not invariant under the transformation  $y \rightarrow -y$  because odd powers of  $y$  in Eq. (11.19) produce asymmetry in the  $y$ -direction.

If the energy is increased to  $E = \frac{1}{8}$  and the calculations are repeated, an unexpected result is obtained. The regions where the oval orbits were found for the lower energy  $E = \frac{1}{12}$  still produce closed trajectories with fixed points at their centers; however, in the regions between these closed trajectories, there is no continuous curve and the points there appear to have no regularity, as shown in Fig. 11.8(b). If we follow the order in which these scattered points appear, we find that, instead of following a regular curve, they jump around in a more or less random fashion from one part of the Poincaré section to another. All of the scattered points on Fig. 11.8(b) arose from the same single chaotic trajectory, and the chaotic region where they appear on the figures constitutes a cross section of a strange attractor. In other words, they all originate from a single orbit meandering through the strange attractor region of phase space and repeatedly penetrating the Poincaré section randomly throughout the chaotic region of this section. Raising the energy still further to the critical value  $E = \frac{1}{6}$  causes the strange attractor



**FIGURE 11.8(b&c)** (b)  $E = \frac{1}{8}$  with regions of regular motion and regions of chaos, and (c)  $E = \frac{1}{6}$  with chaos dominant. The orbit on the right side of (a) is plotted in Fig. 11.7 on an enlarged scale. From M. Hénon (1983), Figs. 21, 22, and 23.

to fill most of the available phase space, and this has the effect of extending the chaotic region to include almost the entire accessible area of Fig. 11.8(c). An index of the extent of the chaos is the fraction of the accessible region where the calculated points lie on regular trajectories. Figure 11.9 shows how the relative area of the regular region declines as the energy increases. The onset of chaos occurs near  $E = \frac{1}{9}$ , beyond which the region of regularity decreases linearly with the energy until complete chaos sets in at about  $E = \frac{1}{6}$ . Calculations for this



**FIGURE 11.9** Fraction of the available Hénon–Heiles Hamiltonian phase space occupied by regular (nonchaotic) orbits plotted as a function of the energy  $E$ . From M. Hénon (1983), Fig. 24.

figure at higher energies are not meaningful because the equipotential lines no longer close on themselves, and the accessible area becomes infinite.

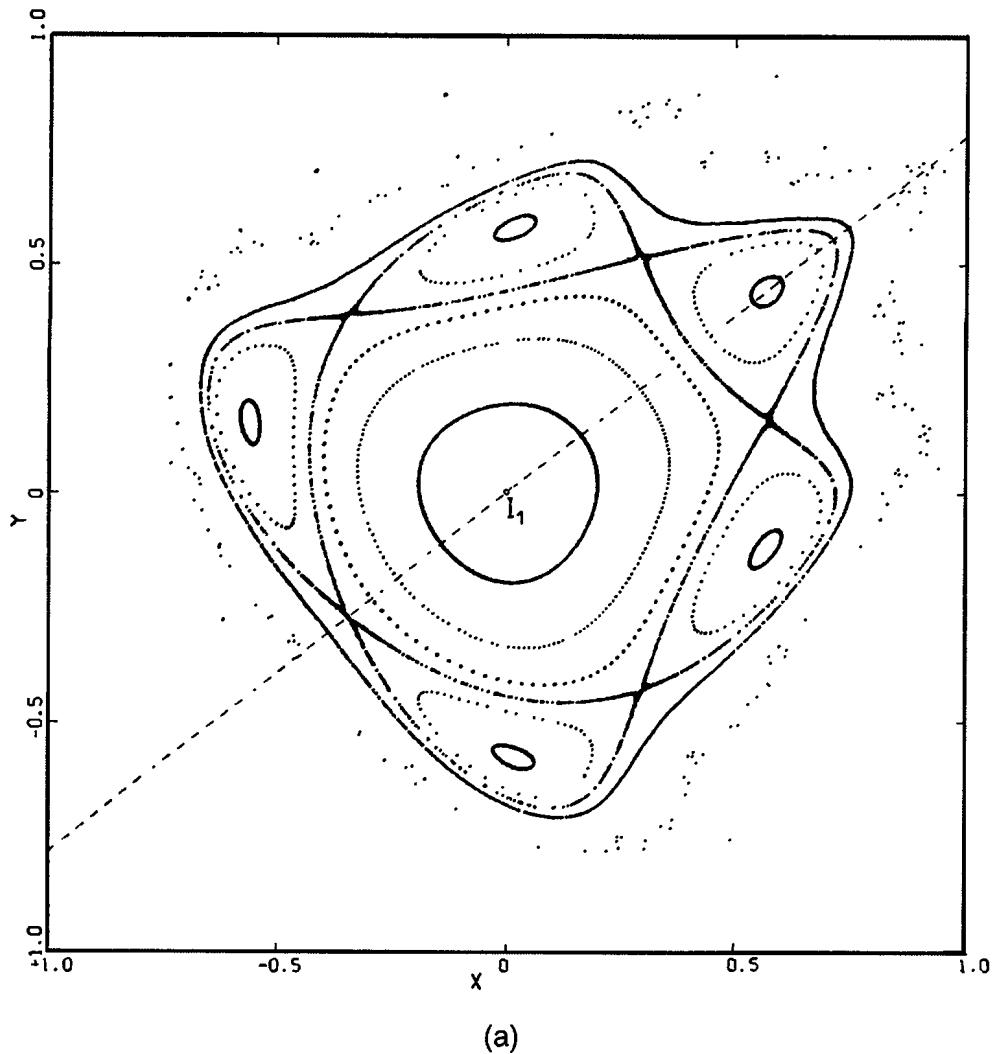
Chaos can also be viewed as a breakdown of integrability. The trajectories of Figs. 11.7 and 11.8(a) for  $E = \frac{1}{12}$  can be obtained by integrating the equations of motion for particular initial conditions; the results obtained by carrying out the integrations are unique and reproducible, and the path followed by the position point is predictable. At the higher energy  $E = \frac{1}{8}$ , the equations are integrable for some initial conditions, but produce points randomly located in the chaotic region for other initial conditions, in accordance with Fig. 11.8(b). For  $E = \frac{1}{6}$ , integrability breaks down over virtually the entire accessible region of phase space depicted in Fig. 11.8(c).

Another interesting feature of chaos is the appearance of what are called *islands*. For very small coupling, such as for energies in the range  $E \sim 10^{-3}$ , the  $\dot{y}$  versus  $y$  section consists of closed orbits slightly perturbed from being circular. The much larger perturbation for the energy  $E = \frac{1}{12}$  produces four sets of elliptic type orbits, and the increase in the energy to  $E = \frac{1}{8}$  results in the appearance of five islands of integrability along the border of the chaotic region on the right side of Fig. 11.8(b). Figure 11.8(c) shows that such islands persist even when almost complete chaos reigns.

In addition to the above features, the chaotic region can exhibit an hierarchy of islands, and these are most easily visualized by plotting constant energy orbits in  $xy$  coordinate space. This can be done for the Hénon–Heiles system, but it will be more instructive for us to plot these coordinate space orbits and display some features of the hierarchy of islands with the aid of another chaotic system called quadratic mapping, which arises from the set of coupled equations

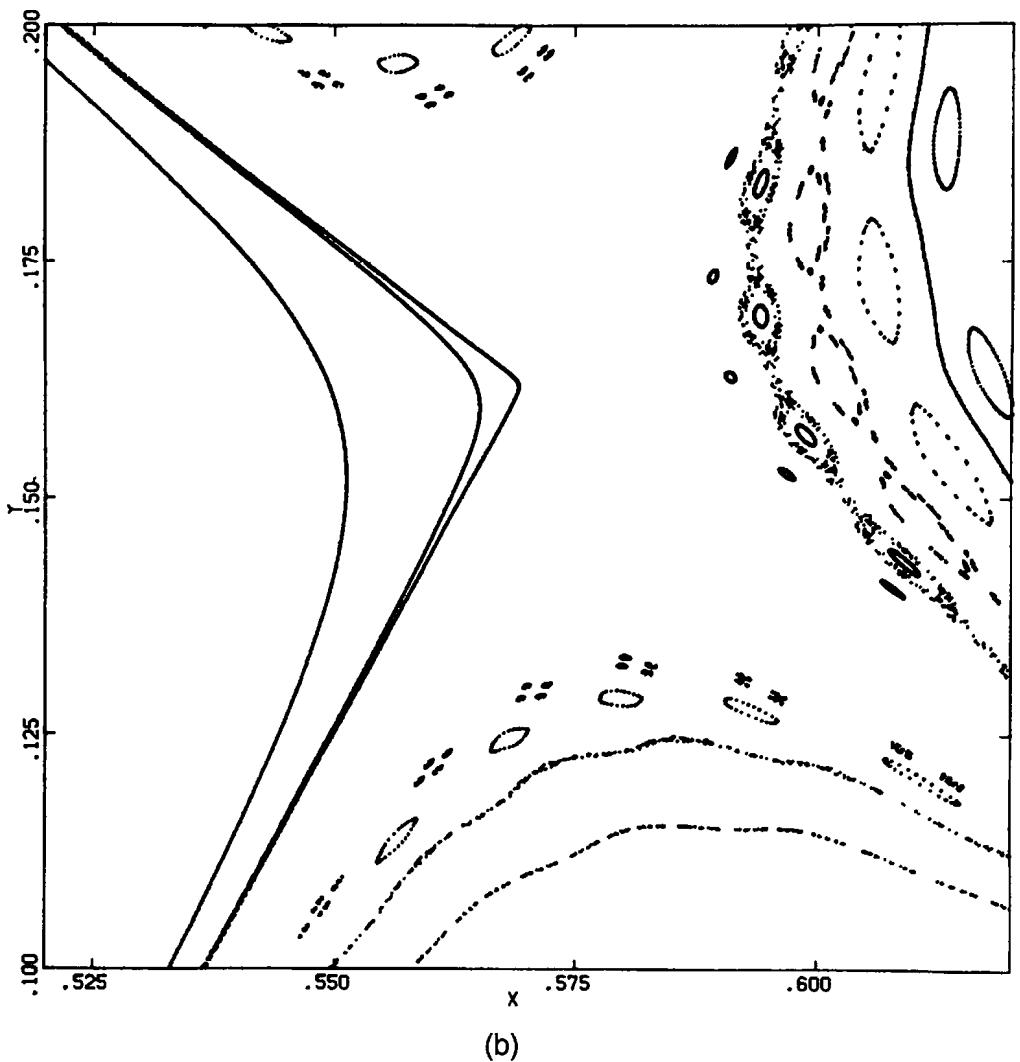
$$\begin{aligned}x_{n+1} &= x_n \cos \alpha - y_n \sin \alpha + x_n^2 \sin \alpha \\y_{n+1} &= x_n \sin \alpha + y_n \cos \alpha - x_n^2 \cos \alpha,\end{aligned}\quad (11.26)$$

where the variables lie in the ranges  $-1 < x < +1$ ,  $-1 < y < +1$ , and  $\alpha$ , which might be called a *control parameter*, determines the extent to which the solutions are regular or chaotic. These equations are solved by an iteration technique similar



**FIGURE 11.10(a)** (a) Trajectories in coordinate space for the quadratic mapping system (11.26) at an energy near the onset of chaos. From M. Hénon (1983), Figs. 33 and 34.

to that described for the logistic equation at the beginning of Section 11.8. Trajectories calculated numerically for the case  $\cos \alpha = 0.4$  are plotted in Fig. 11.10(a). We see from the figure that this system exhibits one main centrally located elliptic-type region, five hyperbolic points where trajectories appear to cross, five outlying elliptic-type regions, and what appears to be a somewhat irregular distribution of dots called islands. The main trajectories can also be referred to as zero-order islands. When the area near one of the hyperbolic points (for example,  $x = 0.57$ ,  $y = 0.15$ ) is enlarged by a factor of 20, we see from Fig. 11.10(b) that the trajectories do not actually cross at a hyperbolic point, but rather there are several series of islands in this region, and some hint of incipient chaos. The well-formed curves on the left are associated with the main central elliptic-type zero-order island region, and the structure at the upper right involves a continuous curve that



**FIGURE 11.10(b)** (b) Enlargement of the rightmost hyperbolic point of (a) showing several orders of “islands.”

encircles the ordered region of Fig. 11.10(a). The long dashed curves at the bottom of Fig. 11.10(b) are part of an outlying elliptic-type region of zero-order islands, and directly above them are first-order islands, each of which has four second-order islands nearby. At the upper border of the figure are first- and second-order islands associated with the zero-order island out of view above the figure. In general, islands tend to be organized in an infinite hierarchy, they are self-similar, and the relatively few islands at one level of enlargement are associated with many islands at the next lower level. Indeed, Fig. 11.10(b) shows islands with a large range of diameters:  $\sim 1.0, 0.3, 0.01, 0.003$ , and  $0.0005$ .

The property of “islands” being replicated at higher and higher levels of magnification is a property characteristic of entities called *fractals*. This *self-similarity* is much more regular in the case of fractals because highly magnified regions can look almost identical to views at much lower magnification. The nonintegral dimensionality associated with a strange attractor that was mentioned earlier in the chapter is also characteristic of fractals. We will have more to say about fractals later in the chapter.

## 11.7 ■ BIFURCATIONS, DRIVEN-DAMPED HARMONIC OSCILLATOR, AND PARAMETRIC RESONANCE

The minimal requirements for a system of first-order equations to exhibit chaos is that they be nonlinear and have at least three variables. While many nonlinear equations in physics are second order, it is possible to reduce a set of second-order nonlinear differential equations to a larger system of first-order nonlinear differential equations. Recall from Section 8.1 that a set of  $N$  second-order Lagrange equations reduces to a set of  $2N$  first-order Hamilton equations. Our present topic deals with the nonlinear analogue of this behavior.

The Hénon–Heiles Hamiltonian satisfies these minimum criteria for chaotic motion. This can be seen by rewriting its two nonlinear second-order equations of motion (11.20) as four first-order equations, two of which are nonlinear

$$\begin{aligned} \frac{dx}{dt} &= v_x & \frac{dv_x}{dt} &= -x - 2xy \\ \frac{dy}{dt} &= v_y & \frac{dv_y}{dt} &= -y + -x^2 + y^2, \end{aligned} \quad (11.27)$$

where there are now four generalized coordinates  $x$ ,  $y$ ,  $v_x$ , and  $v_y$ .

Let us consider, as another example, the driven, damped, harmonic oscillator that has the following equation of motion (cf. Eq. (6.90)):

$$\frac{d^2\theta}{dt^2} + \left(\frac{1}{q}\right) \frac{d\theta}{dt} + \sin\theta = g \cos(\omega_D t), \quad (11.28)$$

where  $\omega_D$  is the driving frequency which is independent of time, and the angle and time coordinates have been renormalized to absorb the excess constants. This

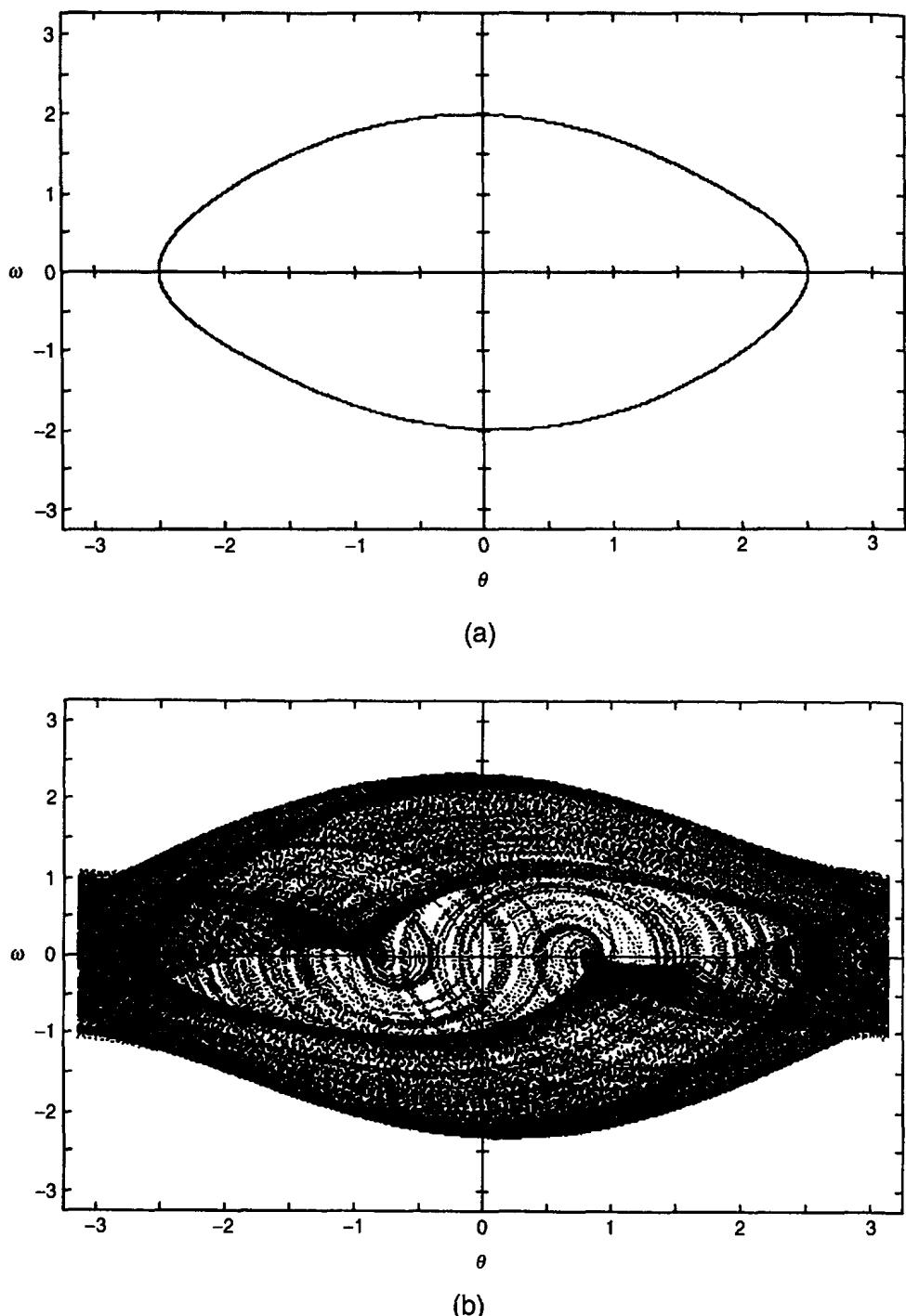
nonlinear second-order differential equation can be converted to a system of three first-order differential equations by writing

$$\begin{aligned}\frac{d\phi}{dt} &= \omega_D \\ \frac{d\theta}{dt} &= \omega \\ \frac{d\omega}{dt} &= -\frac{1}{q}\omega - \sin\theta + g \cos\phi\end{aligned}\tag{11.29}$$

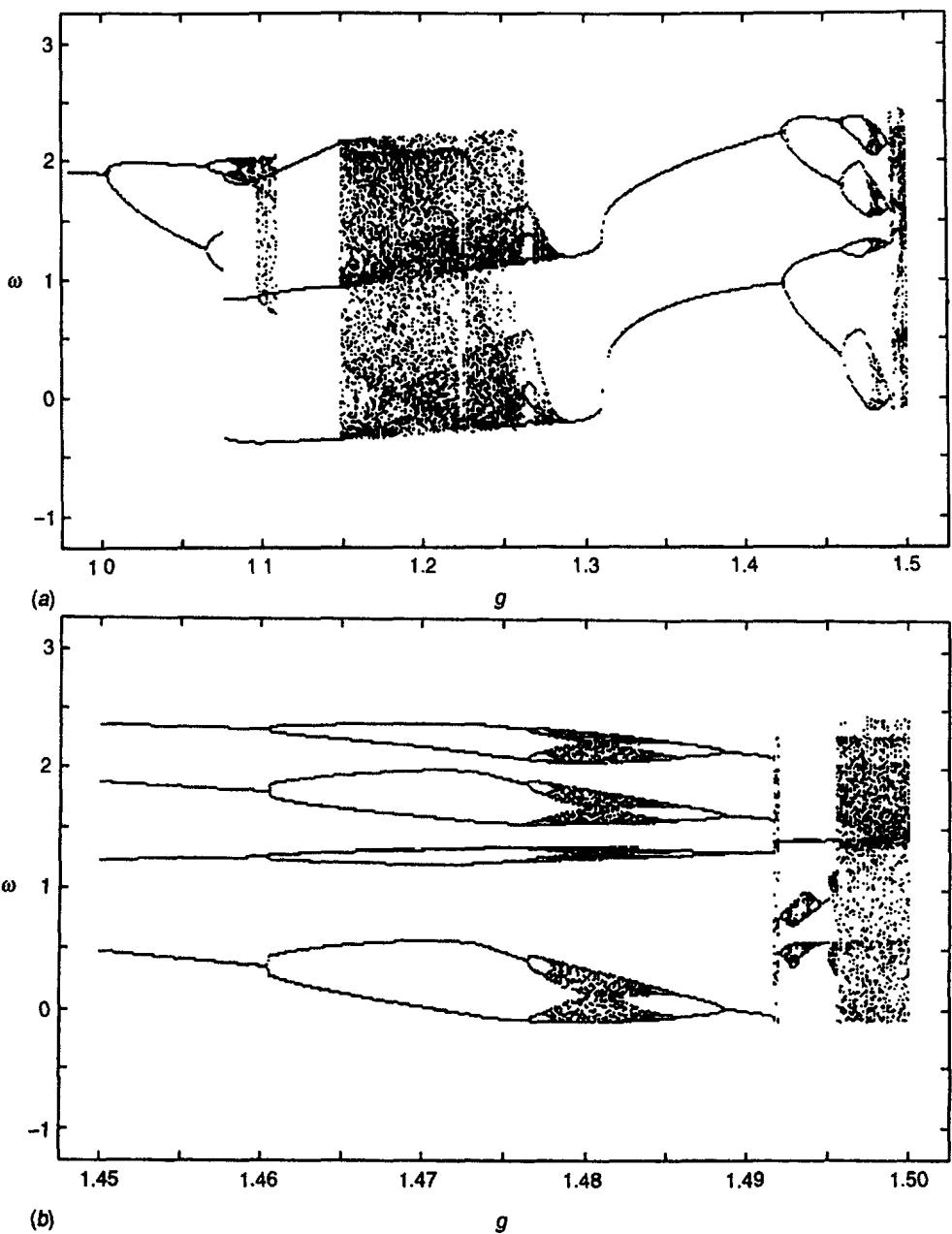
where  $\phi$  is the phase of the driving term. There are now three dependent variables,  $\phi(t)$ ,  $\theta(t)$ ,  $\omega(t)$ , and one independent variable  $t$ . Since the third of these equations is nonlinear, we expect that particular values of the parameters  $q$ ,  $g$ , and  $\omega_D$  might produce chaotic motion. One physical way to justify this expectation is to note that the motion of the pendulum should depend upon the interplay between the “natural” frequency  $\omega$  and the driving frequency  $\omega_D$ .

To obtain quantitative results, we choose  $q = 2$  and let the amplitude  $g$  of the forcing function play the role of what is called a *control parameter*. Such a parameter is an index that delineates regions of normal and chaotic behavior. In Fig. 11.11(a), we show the  $\omega = \dot{\theta}$  versus  $\theta$  Poincaré section for the control parameter  $g = 0.9$ . We see from this figure that the motion is regular, while Fig. 11.11(b) constructed for  $g = 1.15$  displays chaotic motion, that is, randomness in the distribution of points. The periodic nature of the differential equations (11.29) produces regions of stability, and then regions of chaos as the control parameter  $g$  is increased.

If we examine how the frequency of oscillation,  $\omega$ , depends upon this forcing function amplitude  $g$  for a fixed choice of phase,  $\phi$ , we find that the system undergoes a number of *bifurcations* in the measured frequency of the oscillator. At each bifurcation, the number of allowed frequencies doubles. A plot of this is shown in Fig. 11.12. The bifurcations are associated with normal or non-chaotic behavior. The figure also shows shaded regions where the oscillator exhibits chaos. Fig. 11.12(b), which is a factor of ten enlargement of a region of (a), shows that bifurcations and chaos have a complex dependence upon the control parameter. Figures of this type are called *bifurcation diagrams* or *Feigenbaum plots*. A comparison of the two Feigenbaum plots of Fig. 11.12 makes it clear that this system exhibits the property of self-similarity whereby the behavior of  $\omega(t)$  in the neighborhood of one bifurcation resembles that in the neighborhood of other bifurcations, even though the scale or linear dimensions are so much different. It is also evident that the quantity  $g$  seems to “control” the extent to which the system bifurcates and displays chaos. In the Hénon–Heiles system discussed in the previous section, the control parameter is the magnitude of the perturbation  $\Delta\mathcal{H} = x^2y - \frac{1}{3}y^3$ . In the dimensionless units being used there, the effective magnitude of  $\Delta\mathcal{H}$  was set by the choice of energy.



**FIGURE 11.11** Phase space diagram of an orbit of the driven, damped harmonic oscillator (a) in the normal behavior region for  $q = 2$  and control parameter  $g = 0.9$ , and (b) in the chaotic region for  $q = 2$  and  $g = 1.15$ . Reprinted with the permission of Cambridge University Press. From G. L. Baker and J. P. Gollub, *Chaotic Dynamics, An Introduction*, Cambridge, England: Cambridge University Press, 1990, Figs. 3-4a and 3-4c.



**FIGURE 11.12** Feigenbaum plot of the driven, damped harmonic oscillator showing regions of regular and of chaotic behavior. Part (b) is an enlargement of the region on the right side of (a). Reprinted with the permission of Cambridge University Press. From G. L. Baker and J. P. Gollub, *Chaotic Dynamics, An Introduction*, Cambridge, England: Cambridge University Press, 1990, Fig. 4-22.

An example of a parametric harmonic oscillator type system that can become chaotic is the *parametric oscillator*, which satisfies the equation

$$m \frac{d^2x}{dt^2} + G(t, \tau)x = m \frac{d^2x}{dt^2} + (m\omega_0^2 + k(t))x = 0 \quad (11.30)$$

where  $G(t, \tau)$  is the parameter of the oscillator, and the term  $k(t) = k(t + \tau)$  is a perturbation periodic in the time  $\tau$ . Many functions  $k(t)$  produce what is called *parametric resonance*, and we give an example of one. Recall that for a simple rigid rod pendulum of length  $L$ , corresponding to  $k = 0$  in Eq. (11.30), the resonant frequency  $\omega_0 = (g/L)^{1/2}$  and the oscillations can be perturbed by changing the length of the bob. Parametric resonance can be induced in a simple pendulum by shortening the length  $L$  by a small amount  $\Delta L$  when the mass is at its lowest point with the maximum kinetic energy, and increasing the length by the same amount  $\Delta L$  at the top of the motion where the mass is instantaneously at rest, with the kinetic energy zero and the potential energy a maximum. More energy is added at the bottom than is subtracted at the top, so there is a continual increase in energy every cycle.

In general, the evolution in time of the solution of Eq. (11.30) can be highly sensitive to small changes in the initial conditions and the nature of  $k(t)$ . This is a condition for chaos.

## 11.8 ■ THE LOGISTIC EQUATION

Since the driven-damped harmonic oscillator and the parametric resonance oscillator solutions can only be calculated with the aid of sophisticated numerical techniques, we shall consider the detailed analysis of a much simpler mathematical equation called the *logistic equation* or *quadratic iterator*, which lends itself to elementary calculations and exemplifies most of the characteristics of chaos. Its solutions exhibit regularities as well as chaotic behavior. The properties of this equation, using successive iterations, are easy to carry out on a small calculator, and the description of chaos that the calculations provide has much in common with many realistic physical situations. This ubiquitous equation describes behavior in various disciplines such as physics, engineering and economics. For example, in biology it describes population dynamics, or the rise and decline of populations interacting with each other through predator-prey relationships. Other simple functions with a quadratic term also give qualitative and quantitative results similar to those of the quadratic iterator.

The logistic equation is defined by the expression

$$x_{n+1} = ax_n(1 - x_n), \quad (11.31)$$

where  $a$  is the control parameter, with the variable  $x$  is restricted to the domain

$$0 \leq x \leq 1. \quad (11.32)$$

Successive iterations of this equation are expected to bring  $x_{n+1}$  closer and closer to a limiting value,  $x_\infty$ , so that further iterations produce no additional change in  $x_n$ . This limiting value  $x_\infty$  is called a fixed point, and it is obtained by setting  $x_{n+1} = x_n$  in the logistic equation (11.31), which gives

$$x_\infty = \frac{a - 1}{a}. \quad (11.33)$$

Since  $x_n$  is limited to the range given by Eq. (11.32), the control parameter must be positive with the limit  $1 \leq a$ . Equation (11.33) does not set any upper limit on the control parameter, and ordinarily the range  $1 \leq a \leq 4$  is studied.

It is of interest to know the conditions for the fixed point to be stable. For stability, a value of  $x_n$  near the fixed point will iterate to a value  $x_{n+1}$ , which is closer to  $x_\infty$  than  $x_n$  was. To check this, we can select a value of  $x_n$  that is close to the fixed-point value by writing

$$x_n = \frac{a-1}{a} \pm \delta, \quad (11.34a)$$

where  $\delta \ll 1$ . We shall show in Derivation 4 that this gives, to first order in  $\delta$

$$x_{n+1} = \frac{a-1}{a} \pm \delta(2-a). \quad (11.34b)$$

For convergence to  $x_\infty$ , we require the coefficient  $(2-a)$  of  $\delta$  to have an absolute value less than 1, which means that this stable fixed point has the condition

$$1 < a < 3. \quad (11.35)$$

Such a fixed point constitutes an attractor since values of  $x_n$  are attracted to it; that is, they iterate toward it. We see from a left-hand column of Table 11.1 that for the choice  $a = 2$  and the initial value  $x_0 = 0.3$ , less than half a dozen iterations are needed to reach the fixed point  $x_\infty = \frac{1}{2}$  obtained from Eq. (11.33).

It is of interest to find out what happens when we iterate the logistic equation for control parameters beyond the value  $a = 3$ . For  $a = 3.2$  then, we find that after two dozen iterations the value of  $x_n$  alternates between two final values or attractors as follows:

$$\begin{aligned} x_n &= 0.51304 \\ x_{n+1} &= 0.79946, \end{aligned} \quad (11.36)$$

as shown in the center columns of Table 11.1, and for the control parameter  $a = 3.5$ , a double bifurcation corresponds to a fourfold cycle involving the four attractors

$$\begin{aligned} x_n &= 0.501 \\ x_{n+1} &= 0.875 \\ x_{n+2} &= 0.383 \\ x_{n+3} &= 0.827. \end{aligned} \quad (11.37)$$

There is an eightfold cycle for  $a = 3.55$ , a sixteenfold cycle for  $a = 3.566, \dots$ . Figures 11.13(a) and Fig. 11.14 illustrate the bifurcations. These Feigenbaum diagrams, which plot  $x_\infty$  against  $a$ , show how the number of values of  $x_\infty$  succes-

**TABLE 11.1** Examples of iterations of the logistic equation (11.31) before bifurcation with control parameter  $a = 2.0$  (left side), after one bifurcation with control parameter  $a = 3.2$  (center), and in the chaotic region ( $a > a_\infty$ ) with control parameter  $a = 4.0$  (right side). In the normal regions, values of  $|x_n - x_\infty|$  are given, and in the chaotic region, values of  $\Delta x_n = |x_n - x'_n|$  are given for two iterations  $x_n$  and  $x'_n$ , which start close together.

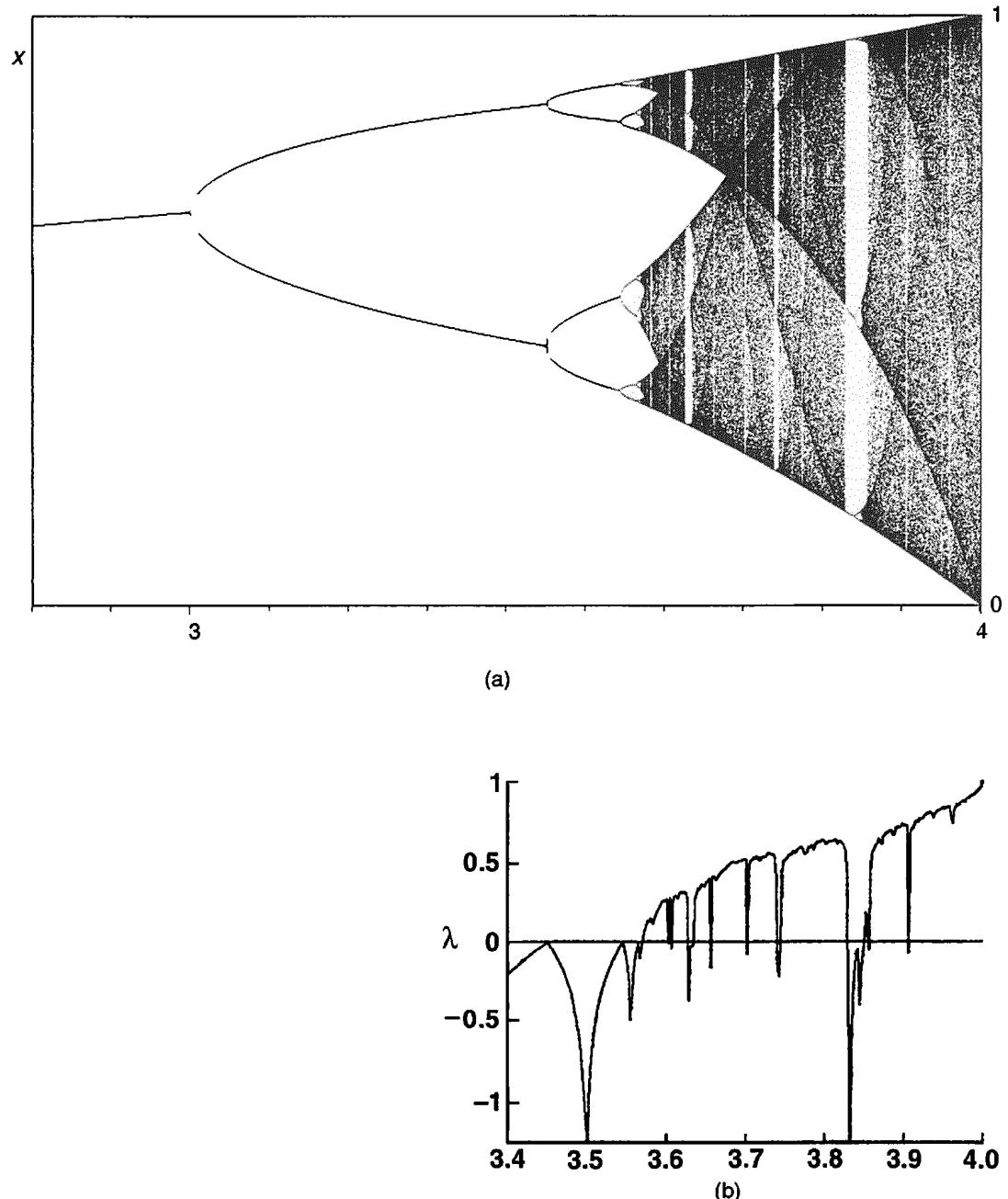
Normal, $a = 2.0$			Normal, $a = 3.2$			Chaotic, $a = 4.0$		
$n$	$x_n$	$ x_n - x_\infty  \times 10^4$	$x_n$	$ x_n - x_\infty  \times 10^4$	$x_n$	$x'_n$	$\Delta x_n \times 10^4$	
0	0.3000	2000	0.3000	2130	0.3000	0.3001	1	
1	0.4200	800	0.6720	1590	0.8400	0.8402	2	
2	0.4872	128	0.7053	942	0.5376	0.5372	4	
3	0.4997	3	0.6651	1521	0.9943	0.9948	5	
4	0.5000	0	0.7128	867	0.0225	0.0220	5	
5	0.5000	0	0.6551	1421	0.0879	0.0859	20	
6	0.5000	0	0.7230	765	0.3208	0.3143	65	
7	0.5000	0	0.6408	1278	0.8716	0.8621	95	
8	0.5000	0	0.7365	630	0.4476	0.4755	279	
9	0.5000	0	0.6210	1080	0.9890	0.9976	86	
10	0.5000	0	0.7531	264	0.0434	0.0096	338	
11		0.5950	820		0.1661	0.0381	1280	
12		0.7711	284		0.5542	0.1465	4077	
13		0.5647	517		0.0734	0.4714	4268	
14		0.7866	129		0.2720	0.9967	7247	
15		0.5372	242		0.7922	0.0199	7723	
16		0.7960	35		0.6586	0.2877	3709	
17		0.5204	74		0.8999	0.8197	802	
18		0.7987	8		0.3619	0.5911	2292	
19		0.5146	16		0.9237	0.9668	431	
20		0.7993	2		0.2819	0.1282	1537	
21		0.5133	3		0.8097	0.4472	3625	

sively doubles: 1, 2, 4, 8, ..., for increasing control parameter  $a$  until the value

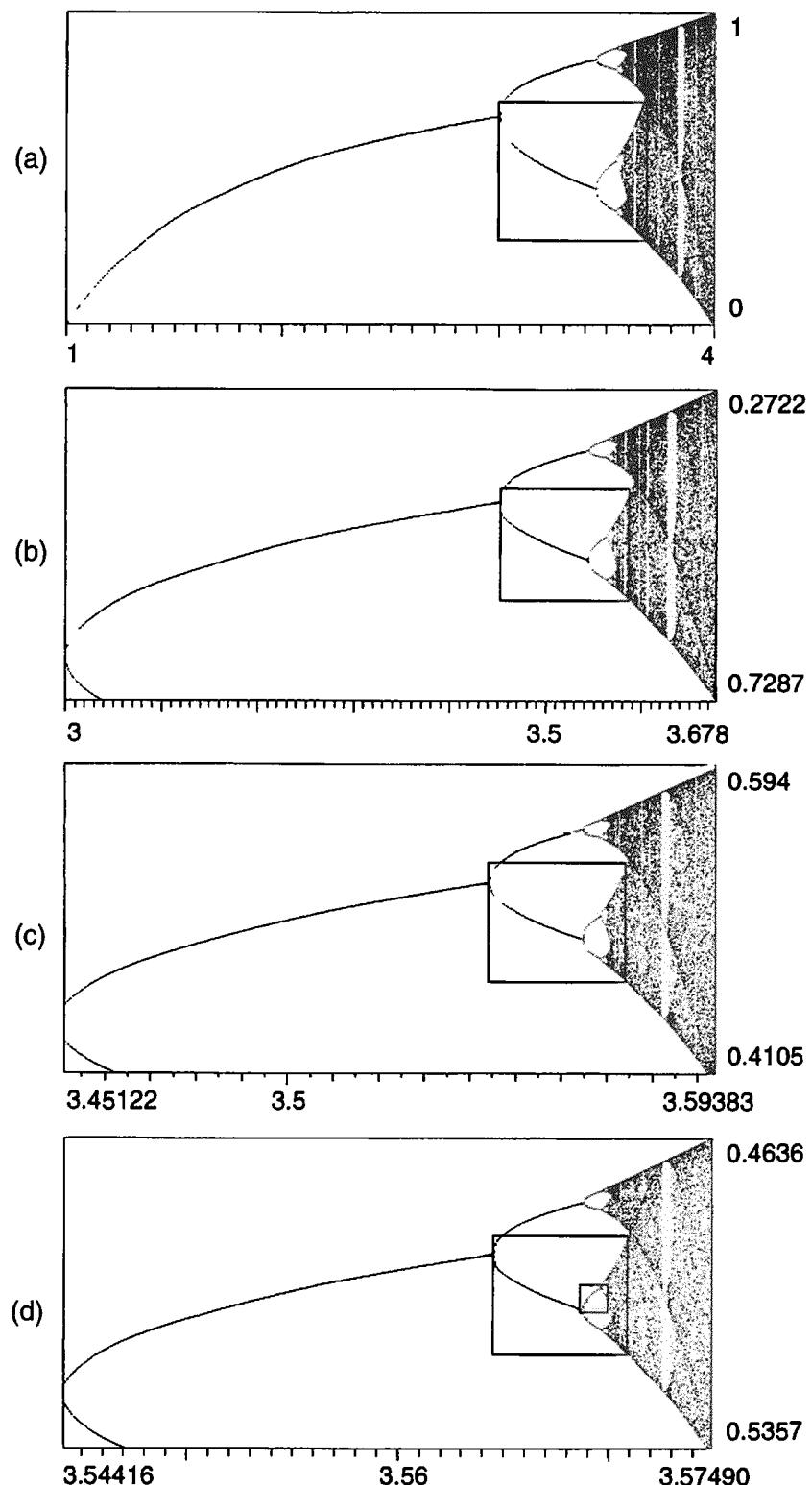
$$a_\infty = 3.5699456 \dots \quad (11.38)$$

called the Feigenbaum point is reached, beyond which the behavior becomes chaotic. For the choice of control parameter  $a = 4.0$  in the chaotic region beyond  $a_\infty$ , successive  $x_n$ -terms generate a sequence of what seems like random numbers. If we start with two very close initial values, such as  $x_0 = 0.3000$  and  $x'_0 = 0.3001$ , we see from the right-hand column of Table 11.1 that after 10 or 11 iterations  $x_n$  and  $x'_n$  become widely separated from each other, and their difference  $\Delta x_n = |x_n - x'_n|$  becomes comparable to their values. Additional iterations produce seemingly random values of  $x_n$  and  $x'_n$ .

A Feigenbaum diagram has some other interesting properties. When the region near each bifurcation is enlarged, we find successive bifurcations that are



**FIGURE 11.13** Correlation between Feigenbaum plot (a) and the Liapunov exponent  $\lambda$  (b) of the logistic equation in the control parameter range from  $a = 3.4$  to  $a = 4.0$ . The figures are aligned with corresponding values of  $a$  to show how sharp minima in  $\lambda$  correlate with bands of normal behavior embedded in the chaos. The Liapunov exponent is negative in the range  $a < a_\infty$  of normal behavior, and positive in the chaotic region  $a > a_\infty$ , except where regions of normal behavior appear in the chaos beyond  $a = a_\infty$ . From Peitgen et al. (1992), Fig. 11–1 (upper figure) and R. Shaw, Z. Naturforsch., 36a, 80 (1981) (lower figure).



**FIGURE 11.14** Feigenbaum diagram of the logistic equation over a wide range (1–4) of control parameter  $\alpha$  (a). Diagrams (b), (c), and (d) show successively greater enlargements of regions near bifurcations. Note the reversals in order of the ordinate scales on the right side of successive figures. From Peitgen et al. (1992), Fig. 11.3.

self-similar to each other, but on successively smaller scales. This is illustrated graphically in the sequence of enlargements, Figs. 11.14(a–d). The ratio of the horizontal spacing between successive bifurcations converges to a limit called the Feigenbaum number  $\delta$

$$\delta = \lim_{n \rightarrow \infty} \frac{a_n - a_{n-1}}{a_{n+1} - a_n} = 4.6692016\dots, \quad (11.39)$$

and the ratio of successive vertical spacings also converges to a limit  $\alpha$ :

$$\alpha = \lim_{n \rightarrow \infty} \frac{x_n - x_{n-1}}{x_{n+1} - x_n} = 2.50290787\dots \quad (11.40)$$

The Feigenbaum number  $\delta$  is a universal constant found with many chaotic systems, but the numbers  $\alpha$  and  $a_\infty$  depend upon the specific model, which in the present case is the logistic equation. Another interesting property of a Feigenbaum diagram is the presence of regions of normality embedded in the chaos. This is evident in Fig. 11.13(a), and is more prominent in the expanded diagrams of Fig. 11.15, which display bifurcations for three levels of enlargements. Each enlargement displays more bifurcations and new regions of normality within the chaos. The fractal property of self-similarity is evident.

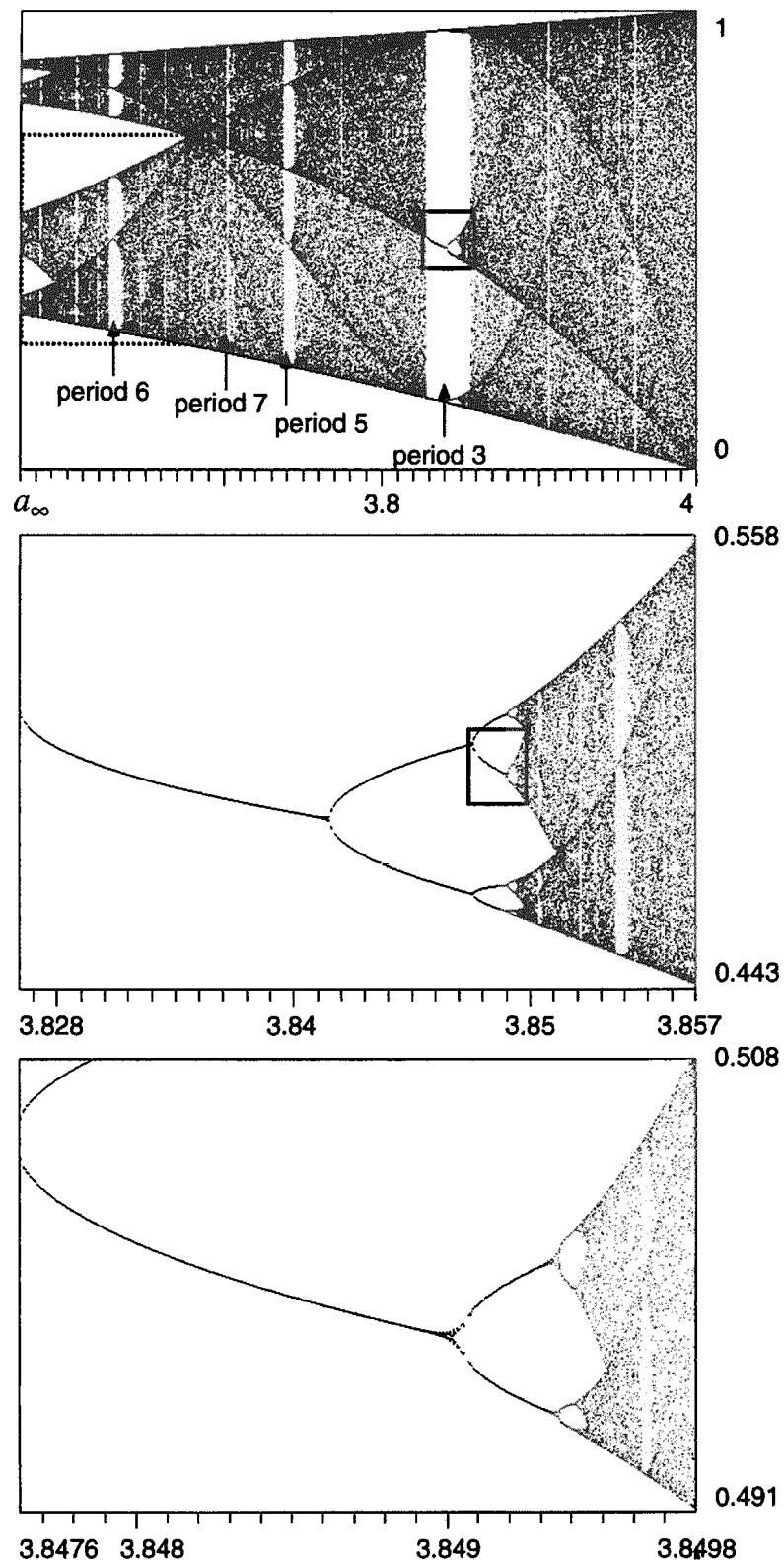
In Section 11.4, we discussed how the rate of approach to a normal-state fixed point or to randomization in the chaotic region is determined by the value of the associated Liapunov exponent  $\lambda$ . This exponent  $\lambda$  from Eq. (11.13) is dimensionless, and we write for the normal and chaotic regions, respectively, as

$$|x_n - x_\infty| = e^{n\lambda} = e^{-n|\lambda|} \quad (\text{normal region}) \quad (11.41a)$$

$$|x_n - x'_n| = e^{n\lambda} = e^{n|\lambda|} \quad (\text{chaotic region}). \quad (11.41b)$$

Note that the exponent  $n\lambda$  is written as  $-n|\lambda|$  for the normal region because  $\lambda$  is negative there. In the normal region  $x_n \Rightarrow x_\infty$  for large  $n$ , so the difference  $|x_n - x_\infty|$  goes to zero. In the chaotic region, the difference  $|x_n - x'_n|$  grows exponentially until it becomes comparable to the overall range of values, namely  $0 < x < 1$ , which means exponential growth in separation until perhaps  $|x_n - x'_n| > 0.2$ . Further iterations keep this separation  $x_n - x'_n$  in the approximate range  $0.2 < x < 1$ . These behaviors are clear from the data in the right-hand columns of Table 11.1. Figure 11.13 shows how the Liapunov exponent depends upon the control parameter. We see from the figure that  $\lambda$  is negative in the normal range  $a < a_\infty$ , and rises to zero at bifurcation points, as can be seen by comparing Figs. 11.13(a) and (b). It is positive in the chaotic region where  $a > a_\infty$ , except where regions of normal behavior that appear white in Fig. 11.13(a) are embedded in the chaos. Near control parameter  $a = 3.83$ , we see three successive minima of  $\lambda$  in the region of negative values that correspond to the period doublings visible in Fig. 11.13(a), and that appear considerably enlarged in Fig. 11.15.

We must remember when studying systems such as the logistic equation that values of  $x_n$  obtained from the iterative process of Eq. (11.31) do not correspond



**FIGURE 11.15** Feigenbaum diagram of the logistic equation showing regions of normal behavior embedded in regions of chaos. Three successive enlargement figures are shown, as indicated by their abscissa and ordinate scales. From Peitgen et al. (1992), Fig. 11.41.

to a particle moving in space. Successive iterations merely illustrate some of the properties of chaos. We must maintain a clear distinction between the chaos that results from simplified models such as that of Hénon–Heiles, and the actual motion of real stars in the galaxy. These simplified models display many of the features that are found in numerical solutions that more closely approximate the real world, but they cannot make reasonable quantitative predictions about the onset of chaos in real physical situations.

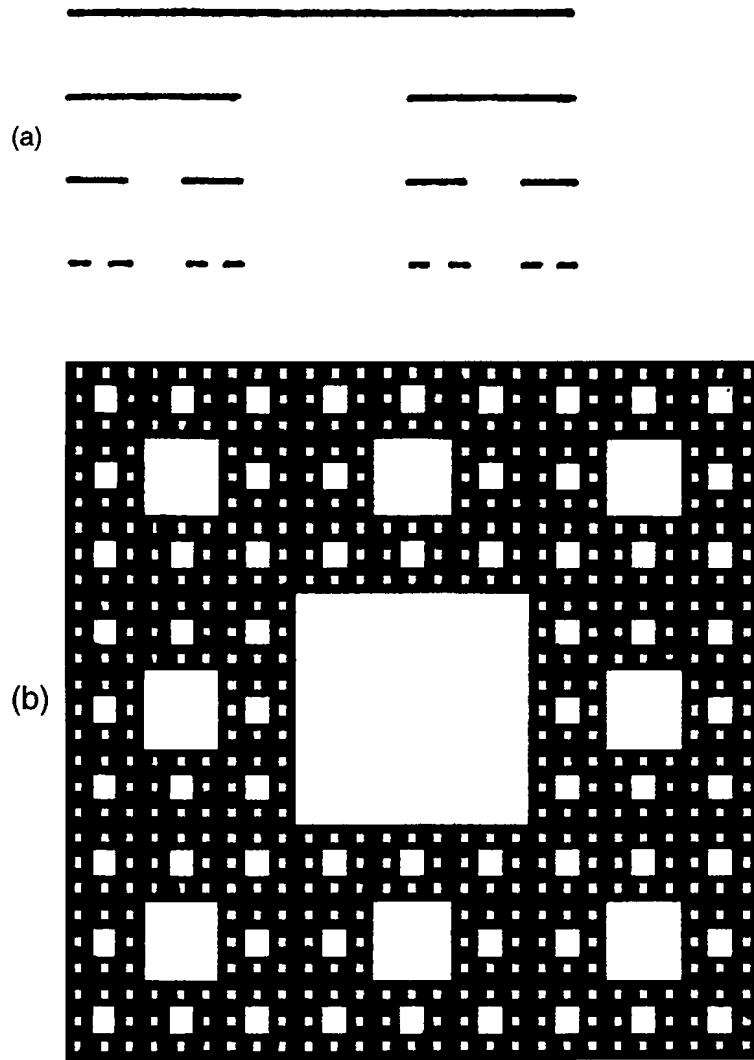
By way of summary, we have seen that where the logistic equation behaves in a normal manner, the solutions occur at values of  $x$  called attractors, stable fixed points that constitute one-dimensional analogues of limit cycles. Beyond this, successive bifurcations are found. In the chaotic region the equation generates numbers in a random manner so that if we start with a value of  $x$  in one small interval, the iteration will eventually produce a number in another previously designated small interval, corresponding to the property of mixing. We also saw that in the chaotic region two points that are initially very close generate successive sequences that do not remain near each other, corresponding to the property of sensitivity to initial conditions. There are also regions of order with attractors, period doublings, and negative Liapunov exponents imbedded in the chaos.

### 11.9 ■ FRACTALS AND DIMENSIONALITY

The phenomenon of “islands” being replicated at higher and higher levels of magnification, as described in Section 11.6, is characteristic of many chaotic systems, and also of entities called fractals. A fractal is an object or set with nonintegral dimensions that exhibits the property of self-similarity. For example, consider a line segment, remove its middle third to produce two line segments, remove the middle third of these latter line segments to produce a total of four, and so on, as indicated in Fig. 11.16(a). If this process of removing the middle third of successively smaller line segments is continued indefinitely, we end up with a series of dots with characteristic spacings called a *Cantor set*. The Cantor set at various stages in its generation is self-similar in the sense that magnifications of the set at later stages of generation have the same appearance as the set itself at earlier stages of formation. The dimensionality of the Cantor set is a little more subtle to deduce because the recursion process of its generation continually increases the number and reduces the size of the residual “dots.”

Before discussing the dimensionality of the Cantor set it will be helpful to say a few words about dimensionality  $d$  in ordinary Cartesian or Euclidian space. In one dimension consider a line segment of length  $a_0$  divided into a large number of equal subdivisions each of length  $a \ll a_0$ . In two dimensions we have a square of side  $a_0$  subdivided into many equal subdivisions each of side  $a \ll a_0$ . In three dimensions the same type tiny squares are made of a cube of side  $a_0$ . In each case the total number of subdivisions, which we denote by  $N(a)$ , is given by

$$N(a) = (a_0/a)^d$$



**FIGURE 11.16** Recursive procedure that generates (a) the Cantor set, and (b) the Sierpinski carpet shown after four steps of iteration. From R. J. Creswick, H. A. Farach, and C. P. Poole, Jr., *Introduction to Renormalization Groups in Physics*, New York: Wiley (1992), Figs. 1.1.1 and 1.2.3.

where the dimensionality  $d = 1, 2, 3$  for these three cases. Solving this expression for the dimensionality of the space we obtain

$$d = \frac{\log N(a)}{\log(a_0/a)} \quad (11.42)$$

This formula for the dimension  $d$  is intuitively obvious for systematic subdivisions of ordinary Euclidian space in any number of dimensions. We will also find it applicable for what we might call the pathological subdivisions of space that are characteristic of fractals. In this application the dimensionality  $d$  determined by the application of Eq. (11.42) is called the Hausdorff or fractal dimension  $d_F$ .

Returning to the Cantor set, it involves subdividing a line originally of length  $a_0$ , which is one-dimensional; that is, its Euclidean dimensionality  $d_E = 1$ . Eventually we feel by intuition that after an infinity of splittings the lines diminish to points that have a dimensionality of zero, and we say that the topological dimensionality of the Cantor set  $d_T = 0$ . Further consideration, however, leads us to think that the limit is never really reached, and that any large but finite number of splittings still leaves an enormous number of infinitesimal one-dimensional line segments present. This suggests that we need another way to assign dimensionality. This can be done by noting that at the  $n$ th level of subdivision the line segments are of length  $a = a_0/3^n$ , and the number of them  $N(a)$  is  $2^n$ . Thus, we have

$$\begin{aligned} a &= 3^{-n}a_0 \\ N(a) &= 2^n. \end{aligned} \quad (11.43)$$

The *fractal dimension* or *Hausdorff dimension*  $d_F$  is defined by the expression

$$d_F = \frac{\log N(a)}{\log(a_0/a)}. \quad (11.44)$$

This definition is chosen to be consistent with the results of Eq. (11.42). Inserting Eqs. (11.43) into Eq. (11.44) to get for the Cantor set

$$d_F = \frac{\log 2}{\log 3} = 0.6309. \quad (11.45)$$

In the following discussion, we shall use  $d_E$  for the initial Euclidean dimension,  $d_T$ , for the final limiting Euclidean (called topological) dimension, and  $d_F$  for the counterintuitive non-integer dimension characteristic of fractals and strange attractors. The fractal dimension  $d_F$  is always between the two limiting values  $d_T$  and  $d_E$ ,

$$d_T < d_F < d_E, \quad (11.46)$$

and we see that this relation is satisfied for the Cantor set

$$0 < 0.6309 < 1. \quad (11.47)$$

It will be instructive to determine the fractal dimensions of an initially two-dimensional ( $d_E = 2$ ) self-similar figure called the Sierpinski carpet of linear dimension  $a_0$  and area  $A_0 = a_0^2$  illustrated in Fig. 11.16(b). To start, a square is divided into nine squares of length  $a = a_0/3$  and area  $A = a^2 = (a_0/3)^2$ , and the middle square removed. Then each of the remaining eight squares is divided into nine smaller squares, and the middle one of each is removed. The figure shows the fourth step in this iteration process. At the  $n$ th level of subdivision, the squares are of length  $a = a_03^{-n}$  and the number of them  $N(a)$  is  $8^n$ . Thus, we have

$$a = a_0 3^{-n}$$

$$N(a) = 8^n. \quad (11.48)$$

The appropriate limit is a set of edges of squares delineating intersecting jagged filamentary lines, which become progressively thinner and thinner with successive iterations, appearing to approach  $d_T = 1$ . The fractal dimension  $d_F$  is again given by Eq. (11.44),

$$d_F = \frac{\ln 8}{\ln 3} = 1.8928, \quad (11.49)$$

and Eq. (11.46) is satisfied by the Sierpinski carpet, as expected,

$$1 < 1.8928 < 2. \quad (11.50)$$

In the general case of a fractal object in a  $d_E$ -dimensional Euclidean space, we define the fractal dimensionality  $d_F$ , also referred to as the capacity dimension, by a covering of the region occupied by the object by  $d_E$ -dimensional spheres in accordance with the expression

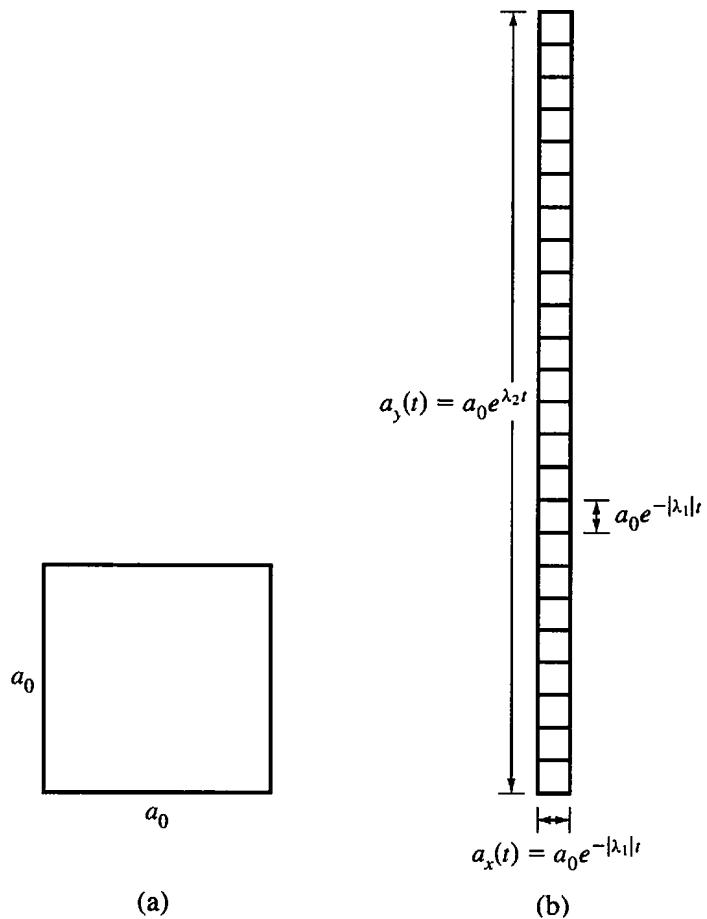
$$d_F = \lim_{r \rightarrow 0} \frac{\log N(r)}{\log (r_0/r)}, \quad (11.51)$$

where  $r$  is the radius of the  $d_E$ -dimensional spheres. This definition is clearly independent of the value of  $r_0$ . If  $d_E = 2$ , the sphere is a 2-sphere or circle of radius  $r$ , and if  $d_E = 1$ , the “sphere” is a 1-sphere or line segment of length  $2r$ . In the case of the Cantor set the object being covered by line segments or one dimensional spheres of radius  $r = a/2$  is the multitude of residual line segments after many subdivisions. In the Sierpinski carpet case the covering is by circles of radius  $r = a/\sqrt{2}$ , where a circle of radius  $r = a_0/\sqrt{2}$  covers the initial square before any subdivisions. Fractal dimensions have been evaluated for many chaotic systems.\* For example, the logistic equation was quoted as having a strange attractor dimension of 0.538, which is between the topological dimension  $d_T = 0$  corresponding to the individual points  $x_n$  and the Euclidean dimension  $d_E = 1$  corresponding to the range of  $x$  given by Eq. (11.32). The driven-damped pendulum with the equation of motion (11.25) exists in two-dimensional  $(x, y)$  Euclidean space, and has one-dimensional orbits of the type shown in Fig. 11.11(a). Its fractal dimensionality determined from Liapunov exponents ranges from 1.2 to 1.4 for various damping factors, which is between the values of  $d_T = 1$  and  $d_E = 2$  that we just mentioned.

We saw in the previous section that chaotic systems exhibit a type of self-similarity, but less regular than in the case of systematically constructed fractals such as those in Fig. 11.16. This does, however, suggest that chaotic systems could have a fractal-type nature, and that nonintegral dimensionality might be a

\*See A. B. Çambel, *Applied Chaos Theory*, New York: Academic Press, 1993, p. 70; G. L Baker and J. P. Gollub, *Chaotic Dynamics, An Introduction*, Cambridge, England: Cambridge University Press, 1990.

characteristic of chaos. This suggestion is correct. The fractal dimensionality  $d_F$  of a *strange attractor* can be calculated from the Liapunov exponents associated with its expansion in phase space. To illustrate this, we consider the particular case of a strange attractor in two-dimensional configuration space, which evolves in time by continuously expanding in one direction and continuously contracting in its orthogonal direction in such a manner that its area  $A(t)$  continuously decreases in magnitude with the passage of time. This permits it to continuously elongate and meander throughout the available regions of phase space. We start with a square zone in the  $x$ ,  $y$  plane of a chaotic region with the initial dimension  $a_0$  in the  $x$ - and  $y$ -directions and the corresponding initial area,  $A_0 = a_0^2$ , as shown in Fig. 11.17a. This means that  $d_E = 2$ . If the area evolves in time by contracting in the  $x$ -direction with the negative Liapunov exponent  $\lambda_1$  and expanding in the  $y$ -direction with the positive Liapunov exponent  $\lambda_2$  it gets continuously thinner and evolves toward a line of topological dimension  $d_T = 1$ . In terms of these Liapunov exponents, the  $x$ - and  $y$ -dimensions of the area have the respective time



**FIGURE 11.17** Role of the Liapunov exponents  $\lambda_1 < 0$  and  $\lambda_2 > 0$ , subject to the condition  $|\lambda_1| > \lambda_2$ , in the evolution of an initially square area (a) in phase space that expands along one coordinate direction and contracts along the other with the passage of time (b).

dependencies from Eq. (11.12),

$$a_x(t) = a_0 e^{-|\lambda_1|t} \quad a_y(t) = a_0 e^{\lambda_2 t}, \quad (11.52)$$

and the area  $A(t)$  evolves in time as

$$A(t) = A_0 e^{(\lambda_2 - |\lambda_1|)t}, \quad (11.53)$$

where  $A_0 = a_0^2$ . Since  $\lambda_1$  is negative and  $\lambda_2$  is positive, it is necessary to have  $|\lambda_1| > \lambda_2$  so that the area (11.50) will continually decrease with time. The feature of a continuous decrease in the fractal area  $A(t)$  of Eq. (11.50) is analogous to the continuous decrease in overall length of the line segments in the Cantor set, and of the continuous decrease in the net remaining area in the Sierpinski carpet case, as the iterations progress to the limit  $n \Rightarrow \infty$ .

If we consider the evolved elongated area  $A(t)$  as containing a number  $N(t)$  of small squares of individual area  $\Delta A(t) = a_x^2$ , as indicated in Fig. 11.17b then we have

$$\Delta A(t) = a_0^2 e^{-2|\lambda_1|t}, \quad (11.54)$$

where  $\lambda_1$  is negative, and

$$N(t) = \frac{A(t)}{\Delta A(t)} = \frac{a_0^2 e^{(\lambda_2 - |\lambda_1|)t}}{a_0^2 e^{-2|\lambda_1|t}} = e^{(\lambda_2 + |\lambda_1|)t}. \quad (11.55)$$

By analogy with Eq. (11.44), the strange attractor dimension  $d_F$ , is given by

$$d_F = \frac{\log N(t)}{\log(a_0/a_x(t))} = 1 + \frac{\lambda_2}{|\lambda_1|}, \quad (11.56)$$

which has a nonintegral or fractal value. For the present case,  $\lambda_2 < |\lambda_1|$ , so Eq. (11.46) is satisfied with  $1 < d_F < 2$ . Thus, a strange attractor is related to a fractal in the sense that its dimension is “strange”; that is, it is not an integer.

There is a fundamental difference between the time evolution and the space-filling effect of regular trajectories and chaotic trajectories. We saw in Section 11.1 how the orbits of incommensurate oscillators can “fill” the space of a torus by ranging over the entire domain. However, technically speaking, these regular orbits do not occupy any of the area of the toroidal surface because they are one-dimensional curves without any width, meaning that the actual area taken up by them is zero. Chaotic orbits also range over their entire domain of phase space, but they do so by occupying area in this space. What is strange is that the more the chaotic orbits “fill” phase space, the smaller the area that they actually occupy (cf Eq. (11.53)). This makes it appropriate to refer to the domain over which the chaotic orbits roam as a strange attractor. The onset of chaos may be looked upon as the increase in the dimension of a regular orbit from its topological value  $d_T = 1$  to its fractal value  $1 < d_F < 2$  as it begins to occupy space in an area of Euclidean dimension  $d_E = 2$ . The fractal dimension may be looked

upon as an index of how much space is occupied by the fractal orbit. We should of course continue to bear in mind the fact that the main difference between the space-“filling” aspects of regular and chaotic orbits is that in the regular incommensurate case the space is “filled” in a systematic manner by the predetermined spiraling motion around the torus, while in the chaotic case the orbit “fills” space in a random, meandering, manner.

This nonintuitive manner in which chaotic orbits in a sense spread out more and more, and in another sense become more attenuated, as they develop in time is very analogous to the behavior of fractals. We saw above how the Cantor set and the Sierpinski carpet illustrated in Fig. 11.16 both become more disperse and more attenuated as they go through successive iterations, always remaining finite throughout the process. There is an analogue of the Sierpinski carpet in three-dimensional Euclidian space called a Sierpinski sponge, which evolves in an analogous manner through an iterative process, dispersing through space while losing volume in accordance with a fractal dimension. Chaotic trajectories are indeed closely related to fractals.

In our treatment of the quantitative aspects of chaos, we have placed more emphasis on the fractal property of nonintegral dimensionality than we have on its property of self-similarity. In the applications of fractals outside the domain of classical mechanics, the emphasis is often more on the self-similarity aspect. Many books display beautiful pictures of precisely drawn figures that illustrate self-similarity down to infinite levels of subdivision, such as the Sierpinski carpet sketched in Fig. 11.16. There are also examples from nature, such as the dendritic growth of the branches of a tree, in which the self-similarity is more approximate and irregular.

## DERIVATIONS

1. Show that the system  $y_{n+1} = 1 - \gamma y_n^2$  with  $-1 < y < 1$  and  $0 < \gamma \leq 2$  can be transformed to the logistic equation (11.31) by the substitution  $y = cx + d$ . Find  $\gamma$ ,  $c$ , and  $d$  in terms of the control parameter  $a$  of the logistic equation.
2. Show that the Hénon–Heiles Hamiltonian (11.17) can be written in polar coordinates as

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{1}{2}kr^2 + \frac{1}{3}\lambda r^3 \sin 3\theta.$$

This form explicitly exhibits the threefold symmetry.

3. Show that for the energy  $E = \frac{1}{6}$ , the bounding equipotential ( $V(r, \theta) = \frac{1}{6}$ ) for the dimensionless Hénon–Heiles potential

$$V(r, \theta) = \frac{1}{2}r^2 + \frac{1}{3}r^3 \sin 3\theta,$$

forms an equilateral triangle in the  $x, y$  plane (cf. Fig. 11.6).

4. Show that Eq. (11.34b) follows from inserting Eq. (11.34a) into Eq. (11.31). In addition, show that the stability range given in the text (cf. Eq. (11.35)) also follows.

## EXERCISES

Most of the following exercises are best completed using a personal computer able to run programs such as *Maple*<sup>TM</sup>, *Mathematica*<sup>TM</sup>, or *Maxima*<sup>TM</sup>. In these exercises the notation  $dz/dt = \dot{z}$  is used.

5. Find the first three bifurcations for the system  $y_{n+1} = 1 - by_n^2$ , where  $-1 < y < 1$  and  $0 < b \leq 2$ .
6. In an attempt to predict weather patterns, Edward N. Lorenz developed a model in 1969 with the following three coupled equations (Lorenz model) in  $x(t)$ ,  $y(t)$ , and  $z(t)$ :

$$\frac{dx}{dt} = \sigma(y - x), \quad \frac{dy}{dt} = rx - y - xz, \quad \frac{dz}{dt} = xy - bz,$$

where  $\sigma$ ,  $r$ , and  $b$  are positive constants and  $x$ ,  $y$ , and  $z$  are real. Lorenz chose, for physical reasons,  $\sigma = 10$  and  $b = \frac{8}{3}$ , and the parameter  $r$  is increased from 0. Let  $x(0) = 2$ ,  $y(0) = 5$ , and  $z(0) = 5$ . Investigate the behavior for

- (a)  $r = 0, 10$ , and  $20$ ,  $0 \leq t < 20$
- (b)  $r = 28$ ,  $0 \leq t < 20$ , where chaotic behavior sets in for  $t \approx 7$

In both cases, investigate the trajectories by using either three-dimensional plots of the coordinates  $x(t)$ ,  $y(t)$ ,  $z(t)$  for different time steps or, if your numeric programs do not generate such plots, plot  $x(t)$  versus  $t$ .

7. A system of equations simpler than the Lorenz equations of Exercise 6 was proposed by O. E. Rössler in 1976, with only one nonlinear system coupling term. This system had no physical intent except to show chaos.

$$\frac{dx}{dt} = -(y + z), \quad \frac{dy}{dt} = x + ay, \quad \frac{dz}{dt} = b + z(x - c),$$

with  $a$ ,  $b$ , and  $c$  positive constants and  $x(t)$ ,  $y(t)$ , and  $z(t)$  real.

- (a) Take  $a = b = 0.2$ , and initial conditions  $x(0) = -1$ ,  $y(0) = z(0) = 0$ . Investigate the effects of changing  $c$  around  $c = 5.7$ , holding  $a$  and  $b$  fixed.
- (b) Take  $a = b = 0.2$ ,  $c = 5.7$  and investigate the effects of changing the initial conditions starting with  $x(0) = -1$ ,  $y(0) = z(0) = 0$ .

8. The general forced damped oscillator equation studied by F. Duffing in 1918 (*Duffing oscillator*) can be written as

$$\frac{d^2x}{dt^2} + 2\gamma \frac{dx}{dt} + \alpha x + \beta x^3 = F \cos \omega t$$

- (a) Take  $\alpha = 1$ ,  $\beta = 0.2$ ,  $\gamma = 0$ ,  $F = 4.0$ ,  $[dx/dt]_{t=0} = 0$  and choose a set of values of  $x_{t=0}$  and  $\omega$  to show that the amplitude (absolute magnitude of the maximum  $x$ ) of the steady-state oscillation shows hysteresis. This is best done by plotting the behavior for increasing  $\omega$  until there is a jump in the amplitude and then continuing the plot for slowly decreasing  $\omega$  from a value slightly larger than where the jump occurred.
- (b) Having solved part (a), pick a value of  $\omega$  in the range of the jump and slightly vary  $F$  to determine how the amplitude varies for a fixed  $\omega$  as  $F$  is changed.

9. Study the van der Pol equation (11.11),

$$m \frac{d^2x}{dt^2} - \varepsilon(1 - x^2) \frac{dx}{dt} + m\omega_0^2 x = F \cos \omega_D t$$

- (a) For the initial conditions near  $x = 0.5$  and  $dx/dt = 0$  for the values of  $\varepsilon = 0, 0.1, 0.2$ , and  $0.3$ . Plot  $x(t)$  as a function of time to determine empirically the rate at which the orbit approaches the attractor at  $x = 1$ .
- (b) Repeat for the initial conditions  $x = 1.5, dx/dt = 0$ .

10. Construct the Poincaré section  $xp$  for the particular Duffing oscillator

$$\frac{d^2x}{dt^2} + 0.7 \frac{dx}{dt} + x^3 = 0.75 \cos t,$$

where  $p = \dot{x} = \frac{dx}{dt}$ , with initial conditions  $x(0) = \frac{dx}{dt}(0) = 0$ .

11. Construct the Poincaré section, as in Exercise 10, for the *inverted Duffing oscillator*,

$$\frac{d^2x}{dt^2} + 0.5 \frac{dx}{dt} - x + x^3 = F \cos t,$$

for values of  $F$  in the range 0.24 to 0.35. This oscillator is said to be inverted because the coefficient of the linear term is negative.

12. The diffusion equation is  $\partial u / \partial t = \eta \nabla^2 u$  where  $u(x, t)$  is the density and  $\eta$  is the diffusion constant. The model of diffusion by Witten and Sadler can be approximated for numerical integration in two dimensions by considering a two-dimensional square lattice and defining the size of a cluster as the minimum radius that includes all of its particles. Mathematically perform the following:

- (a) Place a particle at the center of a  $25 \times 25$  lattice of spacing  $a$ .
- (b) Place a particle at a random position away from the center but not adjacent to the center and allow this particle to randomly move one location at a time until it either leaves the lattice or becomes adjacent to the original particle. For the latter eventuality, draw a circle centered on the center of the cluster that just includes these two particles. Call this radius  $R_{\min}$ . After completing this step  $R_{\min} = a/2$ .
- (c) Repeat this process by adding additional particles at random, increasing  $R_{\min}$  if necessary to include all adjacent particles.
- (d) After a reasonable number of particles,  $N$ , are aggregated, calculate the fractal dimension,  $D$ , by the rule

$$D = \frac{\ln N}{\ln R_{\min}}.$$

13. Construct a Poincaré section for the Hénon–Heiles potential. It is suggested that you make the plot in the  $\dot{y}y$  plane so that you can compare your results with Figs. 11.7 and 11.8. Choose an energy,  $E$ , and initial conditions,  $x = 0$  and  $\dot{x} = 0$ , and initial conditions on  $y$  and  $\dot{y}$  to satisfy the energy condition and find the boundary curve. Relax the condition on  $\dot{x}$  and choose conditions on  $\dot{x}, y$ , and  $\dot{y}$  that satisfy the energy condition for  $x = 0$ . Integrate the equations of motion to find the crossings.

- (a) Choose  $E = \frac{1}{12}$ ,  $y_0 = 0.01$ ,  $\dot{y}_0 = 0.02$ , and  $x_0 = 0$ . Use the energy equation to determine  $\dot{x}_0$ . Integrate the equations to find the values of  $t$  where  $x(t) \approx 0$  saving

the values  $t$ ,  $x(t) \approx 0$ ,  $\dot{x}(t)$ ,  $y(t)$ ,  $\dot{y}(t)$ . Find the first 27 crossings and compare with Fig. 11.7.

(b) Repeat this process for  $E = \frac{1}{8}$  and plot the chaotic behavior.

14. Construct the entries in Table 11.1 for  $a = 3.55$  and  $a = 3.60$ .
15. Refer to Figs. 11.13 and 11.15 for the logistic equation. Find the values of the three cycle attractors embedded in the region of chaos. Use the control parameter  $a = 3.83$ . Also find the values of the next higher cycle obtained for a larger control parameter in this same embedded region of normality.
16. Show that in the control parameter range between the first and second bifurcations of the logistic equation the two final values of the attractors,  $x_n$  and  $x_{n+1}$  such as those given by Eq. (11.36) satisfy the cubic equation

$$a^3 x^2(2 - x) - a^2(a + 1)x + (a^2 - 1) = 0.$$

# CHAPTER

# 12

# Canonical Perturbation Theory

## 12.1 ■ INTRODUCTION

Almost all of the problems in classical mechanics discussed in Chapters 1–10, whether in the text or in the exercises, have had exact solutions. Nevertheless, it should be clear from Chapter 11 on chaos that the great majority of problems in classical mechanics cannot be solved exactly. We have found solutions for the two-body Kepler problem, but with the exception of a few special cases the classical motion of three-point bodies acted upon only by their mutual gravitational forces has proved intractable (see Section 3.12). Even for two bodies the solutions are implicit; no closed explicit formula can be found for the coordinates as a function of time (cf. Section 3.8). There is thus considerable incentive for developing approximate methods of solution.

It often happens, fortunately, that in a physical problem that cannot be solved directly the Hamiltonian differs only slightly from the Hamiltonian for a problem that can be solved rigorously. The more complicated problem is then said to be a *perturbation* of the soluble problem, and the difference between the two Hamiltonians is called the *perturbation Hamiltonian*. Perturbation theory consists of techniques for obtaining approximate solutions based on the smallness of the perturbation Hamiltonian and on the assumed smallness of the changes in the solutions. We know from the discussion in Chapter 11 that even when the change in the Hamiltonian is small, the eventual effect of the perturbation on the motion can be large. This suggests that any perturbation solution must be carefully analyzed to be sure that it is physically correct.

The development of perturbation theory goes back to the earliest days of celestial mechanics. Newton realized, for example, that most of the oscillations in the Moon's motion were the result of small changes in the attraction to the Sun as the Moon revolves about Earth. His initial attempts at a lunar theory including these effects corresponded roughly to a form of perturbation theory. Many of the subsequent developments in the formal structure of classical mechanics, such as Hamilton's canonical theory, stemmed in large measure from the desire to perfect perturbation techniques in celestial mechanics. The need for predicting highly accurate orbits for space vehicles and the enormously increased capacity for numerical computations have spurred further improvements in perturbation theory.

Classical perturbation theory can be divided into two approaches: time-dependent and time-independent perturbations. The terminology is chosen with an eye to perturbation theory as developed for quantum mechanics, and indeed there are many points of analogy between the classical perturbation techniques and their quantum counterparts. Generally speaking, classical perturbation theory is considerably more complicated than the corresponding quantum mechanical version. We shall treat time-dependent perturbation first as being the easier form to understand. While perturbation theory can be developed for all versions of classical mechanics, it is simplest to use the Hamilton–Jacobi formulation.

## 12.2 ■ TIME-DEPENDENT PERTURBATION THEORY

Let  $H_0(q, p, t)$  represent the Hamiltonian for the soluble, unperturbed problem. We imagine the solution has been obtained through Hamilton's principal function  $S(q, \alpha, t)$ , which generates a canonical transformation in which the new Hamiltonian,  $K_0$ , for the unperturbed problem is identically zero. The transformed canonical variables,  $(\alpha, \beta)$ , are then all constant in the unperturbed situation. Now let us consider the perturbed problem for which we write the Hamiltonian as (cf. Eq. (11.8))

$$H(q, p, t) = H_0(q, p, t) + \Delta H(q, p, t). \quad (12.1)$$

As has been emphasized before, the canonical property of a given coordinate transformation is independent of the particular form of the Hamiltonian. Therefore, the transformation

$$(p, q) \rightarrow (\alpha, \beta)$$

generated by  $S(q, \alpha, t)$  remains a canonical transformation for the perturbed problem. Only now the new Hamiltonian will not vanish and the transformed variables may not be constant. For the perturbed problem, the transformed Hamiltonian will be

$$K(\alpha, \beta, t) = H_0 + \Delta H + \frac{\partial S}{\partial t} = \Delta H(\alpha, \beta, t). \quad (12.2)$$

Hence, the equations of motion satisfied by the transformed variables are now

$$\dot{\alpha}_i = -\frac{\partial \Delta H(\alpha, \beta, t)}{\partial \beta_i}, \quad \dot{\beta}_i = \frac{\partial \Delta H(\alpha, \beta, t)}{\partial \alpha_i}. \quad (12.3)$$

Equations (12.3) are rigorous; no approximation has yet been made. If the set of  $2n$  equations can be solved for  $\alpha_i$  and  $\beta_i$  as functions of time, then the equations of transformation between  $(p, q)$  and  $(\alpha, \beta)$  give  $q_j$  and  $p_j$  as functions of time, that is, solve the problem. However, the exact solution of Eqs. (12.3) is usually no less difficult to obtain than for the original equations of motion. The use of

Eqs. (12.3) as an alternative approach to the rigorous solution is therefore not particularly fruitful.

In the perturbation technique, however, advantage is taken of the fact that  $\Delta H$  is small. The quantities  $(\alpha, \beta)$ , while no longer constant, therefore do not change rapidly, at least compared to the explicit dependence of  $\Delta H$  on time. A first-order approximation to the time variation of  $(\alpha, \beta)$  is obtained by replacing  $\alpha$  and  $\beta$  on the *right-hand side* of Eqs. (12.3) by their *constant* unperturbed values:

$$\dot{\alpha}_{1i} = -\frac{\partial \Delta H(\alpha, \beta, t)}{\partial \beta_i} \Big|_0, \quad \dot{\beta}_{1i} = \frac{\partial \Delta H(\alpha, \beta, t)}{\partial \alpha_i} \Big|_0. \quad (12.4)$$

Here  $\alpha_{1i}$  and  $\beta_{1i}$  stand for the first-order perturbation solutions for  $\alpha_i$  and  $\beta_i$ , respectively, and the vertical lines with subscript 0 indicate that after differentiation  $\alpha$  and  $\beta$  are to be replaced by their unperturbed forms; that is, the constants  $(\alpha_0, \beta_0)$ . Equations (12.4) can be placed in matrix form by designating  $\gamma$  as the column matrix of the  $\beta$  and  $\alpha$  canonical variables, so that

$$\dot{\gamma}_i = \mathbf{J} \frac{\partial \Delta H(\gamma, t)}{\partial \gamma} \Big|_0, \quad (12.5)$$

where  $\mathbf{J}$  is the matrix given by Eq. (8.38a). Equations (12.4) can now be integrated directly to yield the  $\alpha_1$  and  $\beta_1$  as functions of time. Through the transformation equations, we then obtain  $(q, p)$  as functions of time to first order in the perturbation. Clearly, the second-order perturbation is obtained by using the first-order dependence of  $\alpha$  and  $\beta$  on time in the right-hand sides of Eqs. (12.4), and so on. In general, the  $n$ th-order perturbation solution is obtained by integrating the equations (in matrix form) for  $\gamma_n$  given by

$$\dot{\gamma}_n = \mathbf{J} \frac{\partial \Delta H(\gamma, t)}{\partial \gamma} \Big|_{n-1}. \quad (12.6)$$

As a trivial example of these procedures, let us consider as the unperturbed system the force-free motion in one dimension of a particle of mass  $m$ . The unperturbed Hamiltonian is

$$H_0 = \frac{p^2}{2m}.$$

The momentum  $p$  is clearly conserved; call its constant value  $\alpha$ . For this system the Hamilton–Jacobi equation is

$$\frac{1}{2m} \left( \frac{\partial S}{\partial x} \right)^2 + \frac{\partial S}{\partial t} = 0. \quad (12.7)$$

Because the system is conservative and  $x$  is cyclic, we know immediately that the solution for Hamilton's principal function is

$$S = \alpha x - \frac{\alpha^2 t}{2m}. \quad (12.8)$$

The transformed momentum is  $\alpha$ ; the transformed constant coordinate is

$$Q \equiv \beta = \frac{\partial S}{\partial \alpha} = x - \frac{\alpha t}{m}$$

or

$$x = \frac{\alpha t}{m} + \beta, \quad (12.9)$$

the expected solution for the force-free motion. While Eq. (12.9) is obvious a priori, this formal derivation via the Hamilton–Jacobi equation at least shows that  $\alpha$  and  $\beta$ , so defined, form a canonical set.

Now suppose the perturbation Hamiltonian is

$$\Delta H = \frac{m\omega^2 x^2}{2}, \quad (12.10)$$

where  $\omega$  is some constant. The total Hamiltonian is

$$H = H_0 + \Delta H = \frac{1}{2m}(p^2 + m^2\omega^2 x^2). \quad (12.11)$$

We are thus considering the harmonic oscillator potential as a perturbation on force-free motion! In terms of the  $\alpha, \beta$  variables, the perturbation Hamiltonian, by Eq. (12.9), is

$$\Delta H = \frac{m\omega^2}{2} \left( \frac{\alpha t}{m} + \beta \right)^2. \quad (12.12)$$

In the perturbed system, the equations of motion for  $\alpha, \beta$  are (cf. Eqs. (12.3))

$$\dot{\alpha} = -m\omega^2 \left( \frac{\alpha t}{m} + \beta \right), \quad (12.13a)$$

$$\dot{\beta} = \omega^2 t \left( \frac{\alpha t}{m} + \beta \right). \quad (12.13b)$$

Note that

$$\dot{\beta} + \frac{t}{m}\dot{\alpha} = 0. \quad (12.14)$$

A rigorous solution of Eqs. (12.13) can be obtained by taking the time derivative of Eq. (12.13a):

$$\ddot{\alpha} = -\omega^2\alpha - m\omega^2 \left( \dot{\beta} - \frac{\dot{\alpha}t}{m} \right) = -\omega^2\alpha. \quad (12.15)$$

Thus,  $\alpha$  in the perturbed system rigorously has a simple harmonic variation with time. From Eqs. (12.13a) and (12.9), it follows  $x = -\dot{\alpha}/(m\omega^2)$ , and hence the solution for  $x$  is also simple harmonic motion. Considered as rigorous equations of motion, Eqs. (12.13) therefore lead properly to the correct and well-known solution.

But now let us treat  $m\omega^2$  ( $\equiv k$ , the force constant) as a small parameter and seek perturbation solutions. The first-order perturbation is obtained by replacing  $\alpha$  and  $\beta$  on the right by their unperturbed values  $\alpha_0$  and  $\beta_0$ . For simplicity, we shall take  $x = 0$  initially, so that  $\beta_0 = 0$ ; the initial value of  $p$  is then  $\alpha_0$ . The first-order equations of motion are then

$$\dot{\alpha}_1 = -\omega^2 \alpha_0 t, \quad \dot{\beta}_1 = \alpha_0 \frac{\omega^2 t^2}{m}, \quad (12.16)$$

with immediate solutions

$$\alpha_1 = \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2}, \quad \beta_1 = \frac{\alpha_0 \omega^2 t^3}{3m}. \quad (12.17)$$

Solutions for  $x$  and  $p$  to first order are then

$$x = \frac{\alpha_1 t}{m} + \beta_1 = \frac{\alpha_0}{m\omega} \left( \omega t - \frac{\omega^3 t^3}{6} \right), \quad (12.18a)$$

and

$$p = \alpha_1 = \alpha_0 \left( 1 - \frac{\omega^2 t^2}{2} \right). \quad (12.18b)$$

Substituting Eqs. (12.17) for  $\alpha$  and  $\beta$  on the right-hand side of Eqs. (12.13), the second-order equations of motion become

$$\begin{aligned} \dot{\alpha}_2 &= -\alpha_0 \omega^2 \left( t - \frac{\omega^2 t^3}{6} \right), \\ \dot{\beta}_2 &= \frac{\alpha_0 \omega^2}{m} \left( t^2 - \frac{\omega^2 t^4}{6} \right), \end{aligned} \quad (12.19)$$

with solutions

$$\begin{aligned} \alpha_2 &= \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2} + \frac{\omega^2 \alpha_0 t^4}{24}, \\ \beta_2 &= \frac{\alpha_0 \omega^2}{m} \left( \frac{t^3}{3} - \frac{\omega^2 t^5}{30} \right). \end{aligned} \quad (12.20)$$

The corresponding second-order solutions for  $x$  and  $p$  are

$$x = \frac{\alpha_0}{m\omega} \left( \omega t - \frac{\omega^3 t^3}{3!} + \frac{\omega^5 t^5}{5!} \right),$$

$$p = \alpha_0 \left( 1 - \frac{\omega^2 t^2}{2!} + \frac{\omega^4 t^4}{4!} \right). \quad (12.21)$$

By now we have enough to see where the  $n$ th-order solution is going. The quantities in the parentheses in Eqs. (12.21) are the first three terms in the expansion of the sine and cosine, respectively. In the limit of infinite order of perturbation, clearly

$$x \rightarrow \frac{\alpha_0}{m\omega} \sin \omega t, \quad p \rightarrow \alpha_0 \cos \omega t,$$

which are the standard solutions consistent with the initial conditions.

The constant transformed variables  $(\alpha, \beta)$  incorporate information on the parameters of the unperturbed orbit. Thus, if the Kepler problem in three dimensions describes the unperturbed system, then a suitable set of  $(\alpha, \beta)$  are the Delaunay variables, that is, the constant action variables  $J_1$  and the constant terms in the corresponding angle variables  $w_i$ . We have seen in Section 10.8 that the Delaunay variables are simply related to the orbital parameters—semimajor axis, eccentricity, inclination, and so on. The effect of the perturbation is to cause these parameters to vary with time. If the perturbation is small, the variation of the parameters within one period of the unperturbed motion will also be small. Time-dependent perturbation theory thus implies a picture in which the perturbed system moves during small intervals of time in an orbit of the same functional form as the unperturbed system, an orbit whose parameters however will be changing in time. The unperturbed orbit along which the system is momentarily traveling is sometimes described as the “osculating orbit.” In position and tangent direction, it matches instantaneously the true trajectory.

As determined by a perturbation treatment, the parameters of the osculating orbit may vary with time in two ways. There may be a periodic variation, in which a parameter comes back to an initial value in a time interval that to first order is usually the period of the unperturbed motion. Or there may remain a net increment in the value of the parameter at the end of each successive orbital period—and the perturbed parameters are said to exhibit *secular* change. Periodic effects of perturbation do not change the average parameters of the orbit; on the whole, the trajectory remains looking much like the unperturbed orbit. A secular change, no matter how small per orbital period, means that eventually, after many periods, the instantaneous perturbed parameters may be quite different from their unperturbed values. Therefore, the major interest in a perturbation calculation will often be in the secular terms only, and the periodic effects may be eliminated early in the game by averaging the perturbation over the unperturbed period. Effectively, this is what was done in Section 5.8 when the perturbing

gravitational potential of the oblate Earth was averaged over the satellite period (cf. Eq. (5.90)).\*

Often we would like to determine the time dependence of the orbital “constants”—for example, eccentricity, or inclination—directly, rather than through the intermediary of the canonical set  $(\alpha, \beta)$ . This can be done easily through the Poisson bracket formalism. Let  $c_i$  be any set of  $2n$  independent functions of the  $(\alpha, \beta)$  constants of the unperturbed system:

$$c_i = c_i(\alpha, \beta). \quad (12.22)$$

One or more of the  $c_i$  may be the desired orbital parameters. Then in the perturbed system the time dependence of the  $c_i$  quantities is determined by the equations of motion

$$\dot{c}_i = [c_i, K] = [c_i, \Delta H]. \quad (12.23)$$

But  $\Delta H(\alpha, \beta, t)$  may equally well, by the inverse of Eqs. (12.22), be considered a function of the  $c$ ’s and  $t$ , so that (cf. Eq (9.68))

$$\begin{aligned} [c_i, \Delta H] &\equiv \frac{\partial c_i}{\partial \boldsymbol{\eta}} \mathbf{J} \frac{\partial \Delta H}{\partial \boldsymbol{\eta}} = \frac{\partial c_i}{\partial \boldsymbol{\eta}} \mathbf{J} \frac{\partial \Delta H}{\partial c_j} \frac{\partial c_j}{\partial \boldsymbol{\eta}} \\ &= [c_i, c_j] \frac{\partial \Delta H}{\partial c_j}. \end{aligned}$$

Hence,

$$\dot{c}_i = [c_i, c_j] \frac{\partial \Delta H}{\partial c_j}. \quad (12.24)$$

As with Eqs. (12.3), Eqs. (12.24) are rigorous equations of motion for the  $c_i$ ’s. They become first-order perturbation equations when the right-hand sides, including the Poisson brackets, are evaluated for the unperturbed motion. In general the  $n$ th-order perturbation is obtained when the right-hand sides are evaluated in terms of the  $(n - 1)$ st order of perturbation. Equations (12.24) thus correspond, in generalized form, to Eqs. (12.6).

\*The circumstances are often more complicated than as described in this paragraph. For example, the periodic variation of orbital parameters can exhibit more than one period. This would obviously occur when the perturbing potential has its own intrinsic periodicity, for example, the varying perturbation of the Sun’s gravity on Earth–Moon orbit as Earth revolves around the Sun. Multiply periodic behavior can also appear through interactions between perturbations. Thus, the periodic perturbation of satellite parameters can show both short and long periods, and it is necessary to average over both kinds of periods to find the secular perturbation effects. Sometimes the dividing line between periodic and secular perturbations becomes a bit vague. What may appear as a secular perturbation in first order will at times on closer examination turn out to be a periodic perturbation with a very long period, as we discovered in Section 11.1 with the harmonic oscillator perturbation calculation. Depending on the purpose of the calculation, it may still be advisable to treat it as a secular perturbation term. Nonetheless, the distinction between periodic and secular terms remains useful and normally straightforward, especially in first-order perturbation theory.

A version of Eqs. (12.24) expressed in Lagrange brackets (cf. Eq. (9.79)) is often found in the literature of celestial mechanics. Multiply the equation for  $c_i$ , by the Lagrange bracket  $\{c_k, c_i\}$  and sum over  $i$ :

$$\{c_k, c_i\}\dot{c}_i = \{c_k, c_i\}\{c_i, c_j\} \frac{\partial \Delta H}{\partial c_j}.$$

By the theorem expressed in Eq. (9.83), this reduces to

$$-\frac{\partial \Delta H}{\partial c_j} = \{c_j, c_i\}\dot{c}_i. \quad (12.25)$$

Historically, the perturbation equations of celestial mechanics are expressed in terms of the *disturbing function*  $R$ , defined as  $-\Delta H$ , so that Eqs. (12.25) appear as

$$\frac{\partial R}{\partial c_j} = \{c_j, c_i\}\dot{c}_i. \quad (12.25')$$

Equations (12.24) or (12.25) are frequently denoted as the *Lagrange perturbation equations*.

### 12.3 ■ ILLUSTRATIONS OF TIME-DEPENDENT PERTURBATION THEORY

**A. Period of the plane pendulum with finite amplitude.** In the limit of small oscillations a plane pendulum behaves like a harmonic oscillator and is isochronous; that is, the frequency is independent of the amplitude. As the amplitude increases, however, the correct potential energy deviates from the harmonic oscillator form, and the frequency shows a small dependence on the amplitude. The small difference between the potential energy and the harmonic oscillator limit can be considered as the perturbation Hamiltonian, and the shift in frequency derived from the time variation of the perturbed phase angle.

The Hamiltonian for a plane pendulum, consisting of a mass point  $m$  at the end of a weightless rod of length  $l$ , is

$$H = \frac{p^2}{2ml^2} + mgl(1 - \cos \theta), \quad (12.26)$$

where, for simplicity, the momentum conjugate to  $\theta$  is denoted by  $p$ . Expanding the  $\cos \theta$  term in a Taylor series, the Hamiltonian can be written as

$$H = \frac{p^2}{2ml^2} + \frac{mgl\theta^2}{2} \left(1 - \frac{\theta^2}{12} + \frac{\theta^4}{360} - \dots\right). \quad (12.27)$$

The small amplitude limit consists of dropping all but the first term in the parentheses. We can get an idea of the magnitude of the correction terms by introducing

artificially a parameter

$$\theta_1^2 = \frac{2E}{mgl} \quad (12.28)$$

and the related parameter

$$\lambda = \frac{\theta_1^2}{6} = \frac{E}{3mgl}$$

The series in the parentheses (cf. (12.27)) then looks like

$$1 - \frac{\lambda}{2} \left( \frac{\theta}{\theta_1} \right)^2 + \frac{\lambda^2}{10} \left( \frac{\theta}{\theta_1} \right)^2 - \dots$$

Now, the ratio  $\theta/\theta_1$  rises to the order of unity at the maximum amplitude. Indeed,  $\theta_1$  is the maximum amplitude of oscillation when  $E$ , and therefore the amplitude, is small. Hence, the rate of convergence of the expansion is determined by the magnitude of  $\lambda$ .

If only one correction term is retained, first-order perturbation introduces terms of the order  $\lambda$  in the motion. Second-order perturbation with the same perturbation Hamiltonian introduces  $\lambda^2$  terms. Thus, to obtain modifications of the motion consistently correct to  $\lambda^2$ , we would have to compute second-order perturbation on the  $\lambda$  term in the Hamiltonian, and first-order perturbation on the  $\lambda^2$  term in the Hamiltonian. We shall here content ourselves with a consistent treatment to order  $\lambda$ ; that is, retain only the first correction term in the Hamiltonian and carry out a first-order perturbation solution.

The unperturbed Hamiltonian derived from Eq. (12.27) can be put in the form of a harmonic oscillator by writing it as (cf. Eq. (10.18))

$$H = \frac{1}{2I}(p^2 + I^2\omega^2\theta^2), \quad (12.29)$$

where  $I = ml^2$ , the moment of inertia of the pendulum, and

$$\omega^2 = \frac{mgl}{I} = \frac{g}{l}. \quad (12.30)$$

A suitable set of canonical variables corresponding to a vanishing  $K$  for the unperturbed system are the action variable  $J$  and the phase angle  $\beta$  in the angle variable:

$$w = vt + \beta, \quad v = \frac{\omega}{2\pi}. \quad (12.31)$$

The effect of the perturbation is to cause both  $J$  and  $\beta$  to vary with time. The equations of transformation relating  $p$  and  $\theta$  to  $J$  and  $\beta$ , respectively, have already been given in Eqs. (10.96) and (10.97), which here take the form

$$\begin{aligned}\theta &= \sqrt{\frac{J}{\pi I\omega}} \sin 2\pi(\nu t + \beta), \\ p &= \sqrt{\frac{IJ\omega}{\pi}} \cos 2\pi(\nu t + \beta).\end{aligned}\quad (12.32)$$

In the unperturbed system  $J$  and  $\beta$  are constant and Eqs. (12.32) constitute the complete solutions for the motion. But the equations remain valid for the perturbed case, only  $J$  and  $\beta$  have time dependencies to be determined.

The unperturbed Hamiltonian is  $H_0 = J\nu$ , but the perturbation Hamiltonian takes the form

$$\Delta H = -\frac{mgl}{24} \theta^4 = -\frac{J^2}{24\pi^2 ml^2} \sin^4 2\pi(\nu t + \beta). \quad (12.33)$$

The first-order time dependence of  $\beta$  and  $J$  are to be obtained from

$$\dot{\beta} = \frac{\partial \Delta H}{\partial J}, \quad j = -\frac{\partial \Delta H}{\partial \beta}, \quad (12.34)$$

where on the right-hand side of each equation the unperturbed solutions for  $J$  and  $\beta$  are to be used; that is,  $J$  and  $\beta$  are considered constant. Thus,

$$\dot{\beta} = -\frac{J}{12\pi^2 ml^2} \sin^4 2\pi(\nu t + \beta). \quad (12.35)$$

Equation (12.35) says that to first order,  $\dot{\beta}$  varies over the cycle of the unperturbed oscillation. But there is a net value for  $\dot{\beta}$  when averaged over a complete cycle, for the average of  $\sin^4$  is  $\frac{3}{8}$ . Hence,  $\dot{\beta}$  exhibits a secular perturbation at a constant rate given by

$$\bar{\dot{\beta}} = -\frac{J}{32\pi^2 ml^2}. \quad (12.36)$$

Viewed over times long compared to the unperturbed period,  $\beta$  has a time dependence

$$\beta = \bar{\dot{\beta}}t + \beta_0. \quad (12.37)$$

Such a variation, when inserted in Eq. (12.32), says that, on average, the first-order solution is still simple-harmonic with a frequency

$$\nu' = \nu + \bar{\dot{\beta}}.$$

Now, in the unperturbed motion

$$J = \frac{2\pi E}{\omega} = 2\pi\omega \frac{El}{g},$$

so that  $\dot{\beta}$ , Eq. (12.36), becomes

$$\dot{\beta} = -\frac{\omega E}{16\pi mgl} = -\frac{v\theta_1^2}{16}. \quad (12.38)$$

The first-order fractional change in the frequency at a finite amplitude  $\theta_1$  is therefore

$$\frac{\Delta\nu}{\nu} = \frac{\dot{\beta}}{\nu} = -\frac{\theta_1^2}{16}, \quad (12.39)$$

a well-known result that can also be obtained by approximating the elliptic-function representation of the motion.

From Eqs. (12.33) and (12.34), it is seen that to first order the time variation of  $J$  is

$$\dot{J} = -\frac{J^2}{3\pi ml^2} \sin^3 2\pi(vt + \beta) \cos 2\pi(vt + \beta).$$

The average of  $\sin^3 \phi \cos \phi$  over even a half period of  $\phi$  is zero; hence,  $J$  shows no secular perturbation. We would expect this result physically, as  $J$  is a measure of the amplitude of the oscillations (cf. Eqs. (12.32)), and the perturbation would not be such as to cause the amplitude to grow or decay with time.

**B. A central force perturbation of the bound Kepler problem.** In Exercise 21, Chapter 3, it was shown rigorously that if a potential with a  $1/r^2$  form is added to the Coulomb potential, the orbit in the bound problem is an ellipse in a *rotating* coordinate system. In effect, the ellipse rotates, and the periapsis appears to precess. Here we will find the precession rate by first-order perturbation theory, considering a somewhat more general form for the perturbing potential.

Suppose the total potential is

$$V = -\frac{k}{r} - \frac{h}{r^n}, \quad (12.40)$$

where  $n$  is an integer greater than or equal to +2. The constant  $h$  will be assumed to be such that the second term is a small perturbation on the first for the range of  $r$  considered. The perturbation Hamiltonian is thus

$$\Delta H = -\frac{h}{r^n}, \quad n \geq 2. \quad (12.41)$$

In the unperturbed problem the angular position of the periapsis in the plane of the orbit is given by the constant  $\omega = 2\pi w_2$  (cf. Eq. (10.166)). With the perturbation,  $\omega$  has a time dependence determined by

$$\dot{\omega} = 2\pi \frac{\partial \Delta H}{\partial J_2} = \frac{\partial \Delta H}{\partial l}, \quad (12.42)$$

using the relation  $J_2 = 2\pi l$  (Eq. (10.156)). First-order perturbation results are obtained by evaluating  $\Delta H$ , and the derivative, in terms of the unperturbed motion. Further, the instantaneous change in  $\omega$  is rarely of interest. In most situations where the perturbation formalism is of value,  $\dot{\omega}$  is so small the change in  $\omega$  is difficult or impossible to perceive within a single orbital period, and it is sufficient to measure only the secular change in  $\omega$  after many orbits. Therefore, what is wanted is  $\bar{\omega}$  averaged over a time interval  $\tau$ , the period of the unperturbed orbit:

$$\bar{\omega} \equiv \frac{1}{\tau} \int_0^\tau \frac{\partial \Delta H}{\partial l} dt.$$

The derivative can be taken outside the integral sign, since  $\tau$  is a function of  $J_3$  only (Eq. (10.142) combined with Eq. (10.146)), whereas the derivative is with respect to  $l = J_2/2\pi$ . Hence,

$$\bar{\omega} = \frac{\partial}{\partial l} \left( \frac{1}{\tau} \int_0^\tau \Delta H dt \right) = \frac{\partial \bar{\Delta H}}{\partial l}. \quad (12.43)$$

But the time average of the perturbation Hamiltonian is here

$$\bar{\Delta H} = -h \overline{\left( \frac{1}{r^n} \right)} = -\frac{h}{\tau} \int_0^\tau \frac{dt}{r^n}. \quad (12.44)$$

By using the conservation of angular momentum in the form  $l dt = mr^2 d\psi$ , the integral can be converted into one over  $\psi$ :

$$\bar{\Delta H} = -\frac{mh}{l\tau} \int_0^{2\pi} \frac{d\psi}{r^{n-2}} \quad (12.45)$$

$$= -\frac{mh}{l\tau} \left( \frac{mk}{l^2} \right)^{n-2} \int_0^{2\pi} [1 + e \cos(\psi - \psi')]^{n-2} d\psi, \quad (12.45')$$

where  $r$  has been expressed in terms of  $\psi$  through the orbit equation, Eq. (3.56) (with  $\psi$  used in place of  $\theta$ ). In general, only terms involving even powers of the eccentricity  $e$  will give nonvanishing contributions to the integral. The derivative with respect to  $l$  also involves  $e$  and its powers, since, by Eq. (10.159),  $e$  is a function only of  $J_2$  and  $J_3$ .

Two special cases are of particular interest. One occurs when  $n = 2$ , mentioned briefly at the start of this illustration. The average perturbation Hamiltonian is then simply

$$\bar{\Delta H} = -\frac{2\pi mh}{l\tau},$$

and the secular precession rate is

$$\bar{\omega} = \frac{2\pi mh}{l^2\tau}, \quad (12.46)$$

which agrees with Exercise 21 of Chapter 3.

The other case of interest is for  $n = 3$  (a  $1/r^3$  perturbation potential), for which Eq. (12.45') reduces to

$$\overline{\Delta H} = -\frac{2\pi m^2 h k}{l^3 \tau}$$

and

$$\bar{\omega} = \frac{6\pi m^2 h k}{l^4 \tau}. \quad (12.47)$$

What makes this choice of  $n$  of particular significance is that general relativity theory predicts a correction to Newtonian motion that can be construed as an  $r^{-3}$  potential. The so-called Schwarzschild spherically symmetric solution of the Einstein field equations corresponds for weak fields to an additional Hamiltonian term in the Kepler problem of the form of Eq. (12.41), with  $n = 3$  and

$$h = \frac{kl^2}{m^2 c^2}, \quad (12.48)$$

so that Eq. (12.47) becomes

$$\bar{\omega} = \frac{6\pi k^2}{\tau l^2 c^2}. \quad (12.49)$$

To apply Eq. (12.49) to the secular precession rate for the precession of a body revolving around the Sun,  $k$  is set equal to  $GMm$  and Eq. (3.63), valid for the unperturbed ellipse, is used

$$l^2 = mka(1 - e^2). \quad (12.50)$$

Equation (12.49) can then be put in the form

$$\bar{\omega} = \frac{6\pi}{\tau(1 - e^2)} \left( \frac{R}{a} \right), \quad (12.51)$$

where  $R$  is the so-called gravitational radius of the Sun is

$$R = \frac{GM}{c^2} = 1.4766 \text{ km.} \quad (12.52)$$

For the planet Mercury,  $\tau = 0.2409$  sidereal years,  $e = 0.2056$ , and  $a = 5.790 \times 10^7$  km; Eq. (12.51) then predicts a precession of the perihelion of Mercury arising from general relativity at an average rate of

$$\bar{\omega} = 42.98''/\text{century}.$$

The observed secular precession of the perihelion of Mercury is over 100 times larger than this value, namely  $5599.74 \pm 0.41''/\text{century}$ . Most of this is due to the precession of the equinoxes, of the remainder, about  $531.54''/\text{century}$  arises from perturbations of the orbit of Mercury by other planets. Only after these two sets of effects are subtracted from the observed precession does the small general relativity effect of approximately  $43''/\text{century}$  become visible. The currently accepted observational value is stated to be  $43.1'' \pm 0.5''/\text{century}$ ; the deviation from the theoretical prediction is not considered significant.

One point remains to be made. In the application to relativistic effects, the constant  $h$ , Eq. (12.48), is a function of the value of  $l$ . It might be asked therefore that in finding  $\dot{\omega}$ , why doesn't the derivative with respect to  $l$  act also on  $h$ ? The key here is that  $h$  is not functionally dependent on  $l$  as a canonical momentum, Equation (12.48) says only how the value of the constant  $h$  is determined in terms of the value of the orbit parameter  $l$ . In other words, the perturbation potential is a function of the dynamical variables only through  $r$ ; it is not to be construed as velocity dependent.

**C. Precession of the equinoxes and of satellite orbits.** The family of problems to be considered here was discussed previously in Section 5.8, which bears the same title. We wish to describe the relative motion of two bodies interacting through their gravitational attraction, one a spherically symmetric or point body, the other being slightly oblate with a resultant gravitational quadrupole moment. The effect of the slight oblate shape of Earth is physically that the torques exerted by the Sun and Moon on the equatorial bulge cause Earth's rotation axis to precess very slowly. Reciprocally, the effect on an object orbiting around Earth, such as the Moon or an artificial satellite, is to cause the plane of the orbit to precess about the figure axis of Earth. The small magnitude of the gravitational quadrupole term, manifested by the very slow rate of precession, suggests that a perturbation treatment should be an extremely good approximation. We shall actually examine here only the case of the perturbation of a satellite's orbit; the reciprocal phenomenon of the precession of the equinoxes proceeds very similarly (though with different notation) from the same perturbation Hamiltonian, and will be left for the exercises.

Since the emphasis here will be on a point satellite moving about a much more massive Earth, the notation of Section (5.8) will be reversed here and  $m$  used to denote the mass of the satellite while  $M$  stands for Earth's mass. The total potential acting on the satellite, by Eq. (5.88), is then

$$V = -\frac{k}{r} + \frac{k}{M} \frac{(I_3 - I_1)}{r^3} P_2(\gamma), \quad (12.53)$$

where  $k = GMm$ ,  $P_2(\gamma)$  is the second-order Legendre polynomial, and  $\gamma$  is the cosine of the angle  $\theta$  between the radius vector to the satellite and Earth's figure axis. For the perturbation Hamiltonian, we therefore have

$$\Delta H = k \frac{I_3 - I_1}{2Mr^3} (3\cos^2\theta - 1). \quad (12.54)$$

The polar angle  $\theta$  can be expressed in terms of the inclination angle of the orbit,  $i$ , and the angle of the radius vector in the orbital plane relative to the periapsis,  $\psi$ , (the so-called true anomaly) by the relation\*

$$\cos \theta = \sin i \sin(\psi + \omega), \quad (12.55)$$

where  $\omega$  is the argument of the periapsis. A small amount of manipulation enables us to rewrite the angular dependence of  $\Delta H$  as

$$3 \cos^2 \theta - 1 = (\frac{1}{2} - \frac{3}{2} \cos^2 i) - \frac{3}{2} \sin^2 i \cos 2(\psi + \omega). \quad (12.56)$$

Now, because of the small size of the perturbation, the chief interest is in the cumulative effects of the secular portion. Thus, the precession of the orbital plane shows up as a secular change in  $\Omega$ , the angle of the line of nodes (or longitude of the ascending node). By the same argument used in the previous illustration we can obtain the secular effects by averaging  $\Delta H$  prior to taking derivatives:

$$\begin{aligned} \overline{\Delta H} &\equiv \frac{1}{\tau} \int_0^\tau \Delta H dt = \frac{m}{l\tau} \int_0^{2\pi} r^2 \Delta H d\psi \\ &= \frac{m^2 k^2 (I_3 - I_1)}{2Ml^3 \tau} \int_0^{2\pi} (1 + e \cos \psi) (3 \cos^2 \theta - 1) d\psi. \end{aligned} \quad (12.57)$$

The term in  $\cos 2(\psi + \omega)$  in Eq. (12.56) gives zero contribution to the integral because it is orthogonal, in the interval of integration, to both  $l$  and  $\cos \psi$ . Hence the averaged perturbation Hamiltonian is

$$\overline{\Delta H} = \frac{\pi m^2 k^2 (I_3 - I_1)}{2Ml^3 \tau} (1 - 3 \cos^2 i). \quad (12.58)$$

In view of Eqs. (10.157) and (10.165) linking  $\Omega$  and  $i$  with the action-angle variables, the first-order perturbation value for  $\dot{\bar{\Omega}}$  is to be found from

$$\dot{\bar{\Omega}} = 2\pi \bar{\dot{w}_1} = 2\pi \frac{\partial \overline{\Delta H}}{\partial J_1} = \frac{1}{l} \frac{\partial \overline{\Delta H}}{\partial \cos i}$$

or

$$\dot{\bar{\Omega}} = -\frac{3\pi m^2 k^2 (I_3 - I_1) \cos i}{Ml^4 \tau}.$$

Finally, using Eq. (12.50), the average fractional change in  $\Omega$  per unperturbed revolution is

\*Equation (12.55) can be obtained in many ways, for example, by matrix rotation of the plane of the orbit into the  $xy$  plane. It is given, most simply perhaps, by some old-fashioned trigonometric reasoning based on Fig. 10.7. As  $OB = 1$ ,  $BC = \cos \theta$ , but  $AB = \sin(\psi + \omega)$  and therefore  $BC$  is also  $\sin i \sin(\psi + \omega)$ .

$$\frac{\bar{\Omega}\tau}{2\pi} = -\frac{3}{2} \frac{I_3 - I_1}{Ma^2} \frac{\cos i}{(1 - e^2)^2}, \quad (12.59)$$

which is the appropriate generalization of Eq. (5.96) to an elliptic satellite orbit.

Once the average perturbation Hamiltonian is known, the effect of the perturbation on other average parameters of the orbit can be found. Thus, the secular precession of the periapsis in the plane of the orbit is immediately given by

$$\bar{\omega} = 2\pi \bar{\dot{\omega}}_2 = 2\pi \frac{\partial \bar{\Delta H}}{\partial J_2} = \frac{\partial \bar{\Delta H}}{\partial l}.$$

The canonical variable  $J_2$  occurs in  $\Delta H$  as given by Eq. (12.58) in two forms: in the  $l^3$  term in the denominator and in the term containing  $\cos i = J_1/J_2$ . Upon carrying out the derivative, it is found that

$$\frac{\bar{\omega}\tau}{2\pi} = \frac{3}{4} \frac{I_3 - I_1}{Ma^2(1 - e^2)} (5 \cos^2 i - 1). \quad (12.60)$$

The maximum value of  $\bar{\omega}$  is thus about the same as that of  $\bar{\Omega}$ , but the dependence upon  $i$  is quite different. At critical inclinations of  $63^\circ 26'$  and  $116^\circ 34'$ , the precession of the periapsis vanishes (at least to first order) and changes sign above and below these points. It is clear that, to first order, there is no secular change in either  $a$  or  $e$ , since  $\bar{\Delta H}$  does not contain the constant parts of any of the angle variables. The shape and size of the osculating ellipse, when averaged over the orbital period, thus does not change with time.

It may be noted from the last two illustrations that the general relativity correction and the gravitational quadrupole field both give rise to a precession of the periapsis of an orbiting body. The former is believed to be the more dominant factor contributing to the observed precession of the perihelion of Mercury, since the measured quadrupole component of the Sun's mass is too small.

## 12.4 ■ TIME-INDEPENDENT PERTURBATION THEORY

Consider conservative periodic separable systems of arbitrary number of degrees of freedom with a perturbation parameter  $\epsilon$ . For the unperturbed problem, we assume a set of action-angle variables  $(J_{0i}, w_{0i})$  such that the unperturbed Hamiltonian,  $H_0$ , is a function only of the action variables  $J_{0i}$ , and correspondingly, the  $w_{0i}$  are then linear functions of time. In the notation of Eq. (10.110'), the relation between, say,  $q_k$  and the  $w_{0i}$  can be written compactly as

$$q_k = \sum_j A_j^{(k)}(\mathbf{J}_0) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0}, \quad (12.61)$$

where  $\mathbf{j}$ ,  $\mathbf{w}_0$ , and  $\mathbf{J}_0$  are  $n$ -dimensional vectors of the integer indices, angle variables, and action variables, respectively.

In the perturbed system,  $(\mathbf{w}_0, \mathbf{J}_0)$  remain a valid canonical set of variables. When expressed in terms of the set  $(\mathbf{w}_0, \mathbf{J}_0)$ , the perturbed Hamiltonian can be expanded in powers of a small perturbation parameter  $\epsilon$ :

$$H(\mathbf{w}_0, \mathbf{J}_0, \epsilon) = H_0(\mathbf{J}_0) + \epsilon H_1(\mathbf{w}_0, \mathbf{J}_0) + \epsilon^2 H_2(\mathbf{w}_0, \mathbf{J}_0) + \dots \quad (12.62)$$

We seek a canonical transformation from  $(\mathbf{w}_0, \mathbf{J}_0)$  to a new set  $(\mathbf{w}, \mathbf{J})$ , such that the  $\mathbf{J}$  are all constants and the  $\mathbf{w}$  therefore linear functions of time. In this set,  $H$  is a function only of  $\mathbf{J}$  (and  $\epsilon$ ) and, in its functional form with respect to  $\mathbf{J}$ , will be written as

$$\alpha(\mathbf{J}, \epsilon) = \alpha_0(\mathbf{J}) + \epsilon \alpha_1(\mathbf{J}) + \epsilon^2 \alpha_2(\mathbf{J}) + \dots \quad (12.63)$$

To obtain the perturbed frequencies through a given order in  $\epsilon$ , it suffices to find the appropriate functions  $\alpha_0, \alpha_1, \dots$ , for then the vector representing the frequencies is

$$\boldsymbol{\nu} = \boldsymbol{\nu}_0 + \epsilon \frac{\partial \alpha_1}{\partial \mathbf{J}} + \epsilon^2 \frac{\partial \alpha_2}{\partial \mathbf{J}} + \dots \quad (12.64)$$

The generator of the canonical transformation from  $(\mathbf{w}_0, \mathbf{J}_0)$  to  $(\mathbf{w}, \mathbf{J})$  is  $Y(\mathbf{w}_0, \mathbf{J}, \epsilon)$ , with a corresponding expansion in  $\epsilon$ :

$$Y(\mathbf{w}_0, \mathbf{J}, \epsilon) = \mathbf{w}_0 \cdot \mathbf{J} + \epsilon Y_1(\mathbf{w}_0, \mathbf{J}) + \epsilon^2 Y_2(\mathbf{w}_0, \mathbf{J}) + \dots \quad (12.65)$$

We seek to find  $Y$  as the solution of the appropriate Hamilton–Jacobi equation:

$$H\left(\mathbf{w}_0, \frac{\partial Y}{\partial \mathbf{w}_0}, \epsilon\right) = \alpha(\mathbf{J}, \epsilon). \quad (12.66)$$

As before, the terms in  $\alpha$  to a given order in  $\epsilon$  are found by expanding both sides in powers of  $\epsilon$  and collecting coefficients of the same order on both sides. We shall illustrate the process for a second-order calculation, where the Hamilton–Jacobi equation reduces to

$$H_0\left(\frac{\partial Y}{\partial \mathbf{w}_0}\right) + \epsilon H_1\left(\mathbf{w}_0, \frac{\partial Y}{\partial \mathbf{w}_0}\right) + \epsilon^2 H_2\left(\mathbf{w}_0, \frac{\partial Y}{\partial \mathbf{w}_0}\right) = \alpha_0(\mathbf{J}) + \epsilon \alpha_1(\mathbf{J}) + \epsilon^2 \alpha_2(\mathbf{J}). \quad (12.67)$$

Each of the terms on the left are functions of  $\epsilon$  through the derivative of  $Y$ :

$$\mathbf{J}_0 = \frac{\partial Y}{\partial \mathbf{w}_0} = \mathbf{J} + \epsilon \frac{\partial Y_1}{\partial \mathbf{w}_0} + \epsilon^2 \frac{\partial Y_2}{\partial \mathbf{w}_0}. \quad (12.68)$$

We again expand the terms  $H_i$  in a Taylor series around  $\mathbf{J}_0 = \mathbf{J}$ , retaining terms of order  $\epsilon^2$  in  $H_0$  and of order  $\epsilon$  in  $H_i$ , with  $\mathbf{J}_0$  replaced directly by  $\mathbf{J}$  in  $H_2$ . The expansions for  $H_0$  and  $H_1$ , in matrix notation, are then

$$\begin{aligned} H_0 \left( \frac{\partial Y}{\partial \mathbf{w}_0} \right) &= H_0(\mathbf{J}) + \left( \epsilon \frac{\partial Y_1}{\partial \mathbf{w}_0} + \epsilon^2 \frac{\partial Y_2}{\partial \mathbf{w}_0} \right) \frac{\partial H_0}{\partial \mathbf{J}} \\ &\quad + \frac{1}{2} \left( \epsilon \frac{\partial Y_1}{\partial \mathbf{w}_0} \right) \frac{\partial^2 H_0}{\partial \mathbf{J} \partial \mathbf{J}} \left( \epsilon \frac{\partial Y_1}{\partial \mathbf{w}_0} \right) \end{aligned} \quad (12.69)$$

$$H_1 \left( \mathbf{w}_0, \frac{\partial Y}{\partial \mathbf{w}_0} \right) = H_1(\mathbf{w}_0, \mathbf{J}) + \epsilon \frac{\partial Y_1}{\partial \mathbf{w}_0} \frac{\partial H_1}{\partial \mathbf{J}}. \quad (12.70)$$

Collecting powers of  $\epsilon$  in Eq. (12.67) then leads to the following expressions for the first three terms in  $\alpha$ :

$$\alpha_0 = H_0(\mathbf{J}), \quad (12.71a)$$

$$\alpha_1 = \mathbf{v}_0 \frac{\partial Y_1}{\partial \mathbf{w}_0} + H_1(\mathbf{w}_0, \mathbf{J}), \quad (12.71b)$$

$$\alpha_2 = \mathbf{v}_0 \frac{\partial Y_2}{\partial \mathbf{w}_0} + \Phi_2(\mathbf{w}_0, \mathbf{J}), \quad (12.71c)$$

where

$$\Phi_2(\mathbf{w}_0, \mathbf{J}) = H_2(\mathbf{w}_0, \mathbf{J}) + \frac{\partial Y_1}{\partial \mathbf{w}_0} \frac{\partial H_1}{\partial \mathbf{J}} + \frac{1}{2} \frac{\partial Y_1}{\partial \mathbf{w}_0} \frac{\partial^2 H_0}{\partial \mathbf{J} \partial \mathbf{J}} \frac{\partial Y_1}{\partial \mathbf{w}_0}. \quad (12.72)$$

Again, the equation of transformation linking  $\mathbf{w}$  and  $\mathbf{w}_0$  is given by

$$\mathbf{w} = \frac{\partial Y}{\partial \mathbf{J}} = \mathbf{w}_0 + \epsilon \frac{\partial Y_1}{\partial \mathbf{J}} + \epsilon^2 \frac{\partial Y_2}{\partial \mathbf{J}} + \dots \quad (12.73)$$

In order for the  $(q, p)$  set to be periodic in both  $\mathbf{w}_0$  and  $\mathbf{w}$  with period 1, all of the  $Y_k$  terms must be periodic functions of  $\mathbf{w}_0$ , that is, of the form

$$Y_k(\mathbf{w}_0, \mathbf{J}) = \sum_{\mathbf{j}} B_{\mathbf{j}}^{(k)}(\mathbf{J}) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0}. \quad (12.74)$$

Hence, all derivatives of  $Y_k$  with respect to  $\mathbf{w}_0$  have no constant term, and the first terms on the right of Eqs. (12.71b,c) do not contribute to the  $J$  dependence. Equations (12.71) can therefore also be written as

$$\alpha_0(\mathbf{J}) = H_0(\mathbf{J}) \quad (12.75a)$$

$$\alpha_1(\mathbf{J}) = \overline{H_1(\mathbf{w}_0, \mathbf{J})}, \quad (12.75b)$$

$$\alpha_2(\mathbf{J}) = \overline{\Phi_2(\mathbf{w}_0, \mathbf{J})}, \quad (12.75c)$$

where the bar denotes an average over the periods of all  $\mathbf{w}_0$ . We can conveniently express all of Eqs. (12.75) in a common format by

$$\alpha_i(\mathbf{J}) = \overline{\Phi_i(\mathbf{w}_0, \mathbf{J})}, \quad (12.75')$$

where  $\Phi_0 = H_0$  and  $\Phi_1 = H_1$ . In addition, Eqs. (12.71) have counterparts periodic in  $\mathbf{w}_0$  with zero mean:

$$\nu_0 \frac{\partial Y_i}{\partial \mathbf{w}_0} = \overline{\Phi_i} - \Phi_i. \quad (12.76)$$

Note that in second-order perturbation the terms in  $Y_1$  do not necessarily vanish in the mean. It is true that the derivatives of  $Y_1$  themselves have zero mean, but they are multiplied by other functions that will be periodic in  $\mathbf{w}_0$ , and there is no guarantee that the average of the product vanishes. Hence, to find the second-order correction to the frequencies, we need to know the first-order canonical transformation. (Analogously in quantum mechanics, a second-order eigenvalue involves first-order corrections of the wave function.) In principle, the coefficients  $B_J^{(1)}$  defining  $Y_1$  through Eq. (12.74) can be found directly from Eq. (12.76) for  $i = 1$ . Subtraction of the average means that  $H_1 - \overline{H_1}$  can be expanded in a Fourier series analogous to Eqs. (12.61) or (12.74) but without any constant term:

$$H_1 - \overline{H_1} = \sum_{\mathbf{j} \neq 0} C_{\mathbf{j}}(\mathbf{J}) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0}. \quad (12.77)$$

Using the derivative of  $Y_1$  in Eq. (12.76) with respect to one of the  $\mathbf{w}_0$ , say  $w_{0k}$ , will bring down a factor  $2\pi i j_k$ . Hence, the matrix product on the left-hand side of Eq. (12.76) can be written

$$\nu_0 \frac{\partial Y_1}{\partial \mathbf{w}_0} = \sum_{\mathbf{j} \neq 0} B_{\mathbf{j}}^{(1)}(\mathbf{J}) 2\pi i (\mathbf{j} \cdot \nu_0) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0}. \quad (12.78)$$

From Eqs. (12.76) and (12.77), the coefficients in the series for  $Y_1$  can be obtained as

$$B_{\mathbf{j}}^{(1)}(\mathbf{J}) = \frac{C_{\mathbf{j}}(\mathbf{J})}{2\pi i (\mathbf{j} \cdot \nu_0)}, \quad \mathbf{j} \neq 0. \quad (12.79)$$

It is true the constant terms in  $Y_1$  are not determined in this way, but it is only the derivatives of  $Y_1$  that enter into the expressions for  $\alpha_i$  and these do not involve the constant terms (cf. Eqs (12.71)).

While we have carried out the procedure in detail only for second-order perturbation, it is easy to see that the general form of the higher-order calculations must be similar; only the details of the algebra will be more complex. For the  $i$ th order perturbation, we will again be able to write  $\alpha_i$  in the form

$$\alpha_i(\mathbf{J}) = \nu_0 \frac{\partial Y_i}{\partial \mathbf{w}_0} + \Phi_i(\mathbf{w}_0, \mathbf{J}). \quad (12.71d)$$

The first term on the right will come from the first-derivative term in the Taylor expansion of  $H(\mathbf{J}_0)$  about  $\mathbf{J}_0 = \mathbf{J}$ , where all terms in the difference  $\mathbf{J}_0 - \mathbf{J}$  are kept through order  $\epsilon^i$ . Only in this term will  $Y_i$  appear; hence,  $\Phi_i$  can contain only

the generators  $Y_k$  for order less than  $i$ . By virtue of the arguments already used for first- and second-order perturbations, the first term on the right in the previous equation (12.71d) has zero mean when averaged over complete cycles in  $\mathbf{w}_0$ , and hence, Eqs. (12.75) and (12.76) are valid in all orders. Of course, for  $i > 2$ ,  $\Phi_i$  becomes increasingly more complicated than Eq. (12.72), but it always contains only such functions as have already been found in lower order calculations. Thus, step by step, we could in principle work up to any order perturbation.

There are practical problems in such a series of calculations of course, but the most serious and obvious conceptual difficulty occurs if the unperturbed system is degenerate. As we see from Eq. (10.122), the existence of a degeneracy means there will be at least one vector of indices  $\mathbf{j}$  such that  $\mathbf{j} \cdot \mathbf{v}_0 = 0$ . The corresponding coefficient  $B_j^{(1)}$  in the Fourier series for  $Y_1$  will therefore, by Eq. (12.79), blow up. Indeed, something similar takes place even when the unperturbed system is *not* degenerate. Even if the frequencies are not exactly equal, as we go to higher and higher values of the integer indices in  $\mathbf{j}$ , eventually there will be found a vector  $\mathbf{j}$  for which  $\mathbf{j} \cdot \mathbf{v}_0$  is very small even if not zero, and the corresponding coefficients  $B$  become very large (the so-called problem of “small divisors”).\* This crudely qualitative observation is the basis of the elegant proof by Poincaré at the end of the last century that the Fourier series for  $Y_1$ , and therefore for the motion, are only semiconvergent. Nonetheless, the series can be truncated at some reasonable values of the indices and still give extremely precise results, at least for times that are not too long.

We shall discuss later what can be done in the presence of degeneracy, but at this point it may be well to illustrate a second-order calculation with a specific example of a system with one degree of freedom.

Consider a one-dimensional *anharmonic oscillator*, that is, one with a  $q^3$  term in the potential energy. The Hamiltonian can be written as

$$H = \frac{1}{2m} \left[ p^2 + m^2 \omega_0^2 q^2 \left( 1 + \epsilon \frac{q}{q_0} \right) \right], \quad (12.80)$$

where  $\omega_0$  is the unperturbed angular frequency:

$$\omega_0 = 2\pi v_0 = 2\pi \sqrt{\frac{k}{m}},$$

$q_0$  is a reference amplitude that can be left unspecified for the moment, and  $\epsilon$  is a small dimensionless parameter. Taken as an expansion in powers of  $\epsilon$ ,  $H$  consists of the terms

$$H_0 = \frac{1}{2m} (p^2 + m^2 \omega_0^2 q^2), \quad (12.81a)$$

\*Similar phenomena, it will be recalled, are found in quantum mechanics, where degeneracy means that there are several states with the same energy  $E$ . Denominators of the form  $E_l - E_j$  will then vanish, or become small even if there is no exact degeneracy.

$$H_1 = \frac{m\omega_0^2 q^3}{2q_0}, \quad (12.81b)$$

and

$$(12.81c)$$

$$H_i = 0, \quad i \geq 2. \quad (12.81d)$$

Using the unperturbed action-angle variables  $(J_0, w_0)$  as canonical variables the nonvanishing parts of  $H$  can, by Eqs. (10.96) and (10.97), be written as

$$H_0 = J_0 v_0 \quad (12.82a)$$

and

$$H_1 = \frac{m\omega_0^2}{2q_0} \left( \frac{J_0}{\pi m\omega_0} \right)^{3/2} \sin^3 2\pi w_0. \quad (12.82b)$$

The recipes of Eqs. (12.75a,b) then give as the lowest two terms in  $\alpha(J)$

$$\alpha_0(J) = H_0 v_0; \quad \alpha_1(J) = 0.$$

To obtain the second-order term  $\alpha_2(J)$ , we note that since  $H_0$  is linear in  $J$ , and  $H_2$  vanishes, then  $\Phi_2$  (cf. Eq. (12.72)) reduces to

$$\Phi_2 = \frac{\partial Y_1}{\partial w_0} \frac{\partial H_1}{\partial J}.$$

But the vanishing of  $\overline{H_1}$  means that Eq. (12.76) for  $i = 1$  has the simple form

$$\frac{\partial Y_1}{\partial w_0} = -\frac{H_1}{v_0}.$$

Combining these two results leads to

$$\Phi_2 = -\frac{1}{2v_0} \frac{\partial H_1^2}{\partial J}. \quad (12.83)$$

Now from Eq. (12.82b),

$$H_1^2(w_0, J) = \frac{v_0 J^3}{2\pi^2 m q_0^2} \sin^6 2\pi w_0,$$

leading to

$$\Phi_2(w_0, J) = -\frac{3J^2}{4\pi^2 m q_0^2} \sin^6 2\pi w_0. \quad (12.84)$$

Since the average of  $\sin^6$  over one period is  $\frac{15}{48}$ ,  $\alpha_2(J)$  is simply

$$\alpha_2(J) = -\frac{15J^2}{64\pi^2mq_0^2}, \quad (12.85)$$

and to second order in  $\epsilon$  the perturbed frequency is

$$\nu = \frac{\partial \alpha}{\partial J} = \nu_0 - \epsilon^2 \frac{15J}{32\pi^2mq_0^2}. \quad (12.86)$$

It is convenient to use for  $q_0$  the maximum amplitude the oscillator would have for the given energy in its unperturbed form, so that to lowest order

$$\frac{m\omega_0^2 q_0^2}{2} = E,$$

or, since  $E = J\omega_0/(2\pi)$ ,

$$mq_0^2 = \frac{J}{\pi\omega_0}. \quad (12.87)$$

In terms of this reference amplitude, Eq. (12.86) is equivalent to saying that the second-order fractional shift in the frequency is simply

$$\frac{\Delta\nu}{\nu_0} = -\frac{15}{16}\epsilon^2. \quad (12.88)$$

Mention has already been made of the difficulties that appear in perturbation theory arising out of the existence of degeneracy, for example, the vanishing (or near vanishing) of  $\mathbf{j} \cdot \boldsymbol{\nu}_0$  in the denominators of Eq. (12.79). Treatment of degeneracies in classical perturbation theory is much more complicated than in quantum mechanics. The mathematics that has been brought to bear on the problem is both subtle and complicated, and a full exposition would be out of place here. Only some brief and introductory remarks can be made at this point.

We speak of exact (or “proper”) degeneracy, as in Section 10.7, when the unperturbed frequencies  $\nu_0$  are such that there are one or more sets of integers  $\mathbf{j}$  for which  $\mathbf{j} \cdot \boldsymbol{\nu}_0 = 0$ . As has been pointed out in Section 10.7, we can then transform to a new set of variables  $(J_0, \omega_0)$  for which the degeneracies appear as zero frequencies and the remaining nonzero unperturbed frequencies are not degenerate. The effect of the perturbation is to lift the degeneracy so that the corresponding frequencies are not exactly zero but have small values. In consequence, there appear in the solution terms that have small frequencies, that is, long periods. The corresponding angle variables are known as “slow” variables, in contrast to the angle variables with nondegenerate frequencies, which are therefore called the “fast” variables. Long-period terms may appear as secular terms over restricted time intervals; for example,  $\sin 2\pi\nu t$  can be taken as a linear function of  $t$  so long as  $\nu t \ll 1$ .

When there is exact degeneracy, a transformation is first made to the  $(\omega_0, \mathbf{J}_0)$  set. The unperturbed Hamiltonian will be a function only of the nondegenerate

$J_0$  variables; in all other respects Eq. (12.82) still represents the complete Hamiltonian. We now carry through the canonical transformation of the perturbation calculation, but only for the nonperturbed variables, leaving the degenerate variables unchanged. The new Hamiltonian, Eq. (12.62), now has the form

$$\alpha(\mathbf{J}, \mathbf{J}'_0, \mathbf{w}'_0, \epsilon) = \alpha_0(\mathbf{J}) + \epsilon\alpha_1(\mathbf{J}, \mathbf{J}'_0, \mathbf{w}'_0) + \epsilon^2\alpha_2(\mathbf{J}, \mathbf{J}'_0, \mathbf{w}'_0) + \dots$$

Here  $\mathbf{w}'_0$  stands for the  $m$  (degenerate) variables that in the unperturbed problem have zero values and  $\mathbf{J}'_0$  for their conjugate momenta. The transformed nondegenerate momenta are represented by  $\mathbf{J}$ . The result of the canonical transformation is thus to eliminate the “fast” variables, but to leave in terms with the “slow” variables. Note that since  $\alpha$  is cyclic in  $w$ , the transformed  $\mathbf{J}$  momenta are true constants of the motion, and  $\alpha(\mathbf{J}, \mathbf{J}'_0, \mathbf{w}'_0, \epsilon)$  can be considered as a Hamiltonian of a system with  $m$  degrees of freedom. Further, since  $\alpha_0(\mathbf{J})$  is a constant, independent of the remaining variables, it doesn’t matter for the equations of motion of  $(\mathbf{J}'_0, \mathbf{w}'_0)$  and can be dropped from  $\alpha$ . Thus, the new effective Hamiltonian is now of order  $\epsilon$ ; in effect, the “unperturbed Hamiltonian” is  $\epsilon\alpha_1(\mathbf{J}, \mathbf{J}'_0, \mathbf{w}'_0)$ , and in this unperturbed problem  $\mathbf{w}'_0$  no longer consists of zero values. If there is only one degeneracy condition, the effective problem is of only one degree of freedom and is in principle immediately integrable. With more degeneracy conditions, we can seek a second canonical transformation to eliminate the “slow” variable terms just as was done for the “fast” variables. In practice, the procedure obviously becomes quite complicated.

It has already been pointed out, in connection with Eq. (12.79), that even with nondegenerate frequencies, small values of the divisor  $\mathbf{j} \cdot \nu_0$  will inevitably occur as the indices  $\mathbf{j}$  become larger and larger. This phenomenon is referred to as *resonance*, implying that the amplitude of some particular term in the Fourier expansions becomes very large. It would seem therefore that the problems of degeneracy will always be with us, no matter what the unperturbed frequencies are! The situation is not all as bad as that, in part because of the nature of the perturbation Hamiltonians encountered in practice. From Eq. (12.79), it will be noted that what counts is not so much the value of  $\mathbf{j} \cdot \nu_0$  as the ratio

$$\frac{C_{\mathbf{j}}}{\mathbf{j} \cdot \nu_0},$$

where  $C_{\mathbf{j}}$  is the Fourier series expansion of the perturbation Hamiltonian  $H_1$ , cf. Eq. (12.77). It turns out that in celestial mechanics, at least, most perturbation Hamiltonians have what is called the *D'Alembert characteristic*. While the formal mathematical definition of the property is complicated, what it says, roughly, is that when the values of the integers in the  $\mathbf{j}$  indices are larger than the exponent of  $\epsilon$  in the Hamiltonian, the magnitudes of  $C_{\mathbf{j}}$  fall rapidly (generally exponentially) with increasing values of the indices. The ratios in Eq. (12.79) then do not become too large, and the expansion process actually can be proved to converge when the frequencies  $\nu_0$  are incommensurate.

Resonant behavior in the presence of the D'Alembert characteristic, or generally when  $C_j / (\mathbf{j} \cdot \mathbf{v}_0) < O(\epsilon^{1/2})$ , is described as a *shallow resonance*. In principle, at least, shallow resonances may not upset the perturbation expansion process and can be tolerated without introducing new methods. There are situations where the ratio  $C_j / (\mathbf{j} \cdot \mathbf{v}_0)$  becomes large, at least larger than order  $\epsilon^{1/2}$ , and these are referred to as *deep resonances*. Special methods have to be devised to handle deep resonances, such as the so-called Bohlin expansion in powers of  $\epsilon^{1/2}$  rather than in powers of  $\epsilon$ .

## 12.5 ■ ADIABATIC INVARIANTS

At the first Solvay Conference in 1911, which grappled with the problems of introducing quantum notions into physics, a deceptively simple problem in classical mechanics was raised. Consider a bob on a string oscillating as a plane pendulum, with the string passing through a small hole in the ceiling. Now imagine that the string is either pulled up or let down slowly, so slowly that there is little change in the length of the pendulum during one period of oscillation. What happens to the frequency of oscillation during this process? Note that the energy of the pendulum is not conserved, for work is done on the system (or extracted from it) as the length of the string is altered. By elementary means it was demonstrated that for very slow change of the ratio  $E/v$  would be constant. It will be recognized that this ratio is precisely the action variable  $J$ . The *adiabatic invariance* of the action variables under slow change of parameters was a very satisfying property to physicists developing quantum mechanics. For simplicity, we shall examine only periodic systems with one degree of freedom, although the extension to many degrees of freedom normally is not difficult in the absence of degeneracy. We consider a system that initially has no dependence on the time, and that involves a parameter  $a$ . Implicit in the method is a picture of the system as initially conservative with  $a$  constant. Time dependence of  $a$  is then "switched on," and  $a$  varies slowly over a long time, eventually reaching a constant value. When  $a$  is constant, the motion is periodic, and the slow change in the parameter does not alter the periodic nature of the motion. Although the changes in the motion are small in any one period, over a long interval of time the properties of the motion can accumulate large quantitative changes. The switching on of the time dependence is thus in the nature of a small perturbation, and we are looking for secular changes in the motion.

When the parameter  $a$  is constant, the system will be described by action-angle variables  $(J_0, w_0)$  such that the Hamiltonian is  $H = H(J_0, a)$ . It will be useful to consider these variables as derived from an original canonical set  $(q, p)$  via an  $F_1$  generating function  $W^*(q, w_0, a)$ . The usual Hamilton–Jacobi equation of course leads to an  $F_2$  generating function of the form  $W(q, J_0, a)$ , but these two generating functions are normally connected by a Legendre transformation (cf. Eq. (9.19)):

$$W^*(q, w_0, a) = W(q, J_0, a) - J_0 w_0. \quad (12.89)$$

When  $a$  is allowed to vary with time,  $(w_0, J_0)$  of course remain as valid canonical variables, but the generating function is now an explicit function of time through the time dependence of  $a$ . Hence, the appropriate Hamiltonian for the  $(w_0, J)$  set is now

$$\begin{aligned} K(w_0, J_0, a) &= H(J_0, a) + \frac{\partial W^*}{\partial t} \\ &= H(J_0, a) + \dot{a} \frac{\partial W^*}{\partial a}. \end{aligned} \quad (12.90)$$

Since  $J_0$  is no longer a constant and  $w_0$  does not vary linearly with time, the second term in the Hamiltonian is a perturbation. The time dependence of  $J_0$  is governed by the equation of motion

$$J_0 = -\frac{\partial K}{\partial w_0} = -\dot{a} \frac{\partial}{\partial w_0} \left( \frac{\partial W^*}{\partial a} \right), \quad (12.91)$$

where of course the derivative in parenthesis is expressed, as is  $K$ , in terms of  $J_0$ ,  $w_0$ , and  $a$ . In the spirit of a first-order perturbation theory, we look for a secular term, the average of  $J_0$  over the period of the unperturbed motion *for the appropriate  $a$* . Since  $a$  varies slowly,  $a$  can be taken as constant during this time interval, and the average can be written as

$$\begin{aligned} \bar{J}_0 &= -\frac{1}{\tau} \int_{\tau} \dot{a} \frac{\partial}{\partial w_0} \left( \frac{\partial W^*}{\partial a} \right) dt \\ &= -\frac{\dot{a}}{\tau} \int_{\tau} \frac{\partial}{\partial w_0} \left( \frac{\partial W^*}{\partial a} \right) dt + O(\dot{a}^2, \bar{a}). \end{aligned} \quad (12.92)$$

It will be remembered from Eq. (10.17) that  $W$  is given by the indefinite integral

$$W = \int p dq.$$

In one period of  $w_0$ , the generating function,  $W$ , therefore increases by  $J_0$ . At the same time,  $J_0 w_0$  also increases by  $J_0$ , since  $w_0$  increases by unity. Hence, by Eq. (12.89),  $W^*$  is a periodic function of  $w_0$ , and both it and the derivative with respect to  $a$  can be expressed as a Fourier series:

$$\frac{\partial W^*}{\partial a} = \sum_k A_k(J_0, a) e^{2\pi i k w_0}. \quad (12.93)$$

The average,  $\bar{J}_0$ , therefore has the form

$$\bar{J}_0 = -\frac{\dot{a}}{\tau} \int_{\tau} \sum_{k \neq 0} 2\pi i k A_k(J_0, a) e^{2\pi i k w_0} dt + O(\dot{a}^2, \bar{a}).$$

Since the integrand has no constant term, the integral vanishes,

$$\overline{\dot{J}_0} = 0 + O(\dot{a}^2, \ddot{a}), \quad (12.94)$$

and  $\overline{\dot{J}_0}$  has no secular variation to first order in  $\dot{a}$ , proving the desired property of adiabatic invariance.

Let us see how this derivation would work in detail for the problem of the harmonic oscillator:

$$H = \frac{1}{2m}(p^2 + m^2\omega^2q^2),$$

where  $\omega$  may be an explicit function of time. The equations of the canonical transformation from the  $(q, p)$  set to the  $(J_0, w_0)$  set are given by Eqs. (10.21) and (10.97), which can be written so as to facilitate the evaluation of  $W^*$ :

$$\begin{aligned} J_0 &= \pi m\omega q^2 \csc^2 2\pi w_0 = -\frac{\partial W^*}{\partial w_0}, \\ p &= m\omega q \cot 2\pi w_0 = \frac{\partial W^*}{\partial q}. \end{aligned} \quad (12.95)$$

To within constant (and therefore irrelevant) terms,  $W^*$  is found by integration of Eqs. (12.95) to be

$$W^*(q, w_0, \omega) = \frac{m\omega q^2}{2} \cot 2\pi w_0. \quad (12.96)$$

The derivative with respect to  $\omega$  is

$$\frac{\partial W^*}{\partial \omega} = \frac{mq^2}{2} \cot 2\pi w_0,$$

or, using Eq. (10.96) as a function of  $w_0$ ,  $J_0$ , and  $\omega$ ,

$$\frac{\partial W^*}{\partial \omega} = \frac{J_0}{4\pi\omega} \sin 4\pi w_0. \quad (12.97)$$

Thus,  $\dot{J}_0$  is given by the one-term Fourier expansion

$$\dot{J}_0 = -\frac{\dot{\omega}}{\omega} J_0 \cos 4\pi w_0, \quad (12.98)$$

which, as predicted, has no constant term. So far, Eq. (12.98) is rigorous. Similarly the rigorous connection between  $w_0$  and time is determined by the  $w_0$  equation of motion

$$\dot{w}_0 = \frac{\partial K}{\partial J_0} = \frac{\partial H}{\partial J_0} + \dot{\omega} \frac{\partial}{\partial J_0} \left( \frac{\partial W^*}{\partial \omega} \right) = \frac{\omega}{2\pi} + \frac{\dot{\omega}}{4\pi\omega} \sin 4\pi w_0. \quad (12.99)$$

In order to calculate an average of  $\dot{J}_0$  over a period, including at least the first correction term, we begin to make approximations. First we shall assume that over a particular period of the perturbed motion the ratio

$$\frac{\dot{\omega}}{\omega} \equiv \epsilon \quad (12.100)$$

is a constant, and one such that  $\epsilon t \leq 1$ . Equation (12.100) corresponds to a variation

$$\omega = \omega_0 e^{\epsilon t} \approx \omega_0(1 + \epsilon t), \quad (12.101)$$

where  $t$  is measured from the start of the period interval, at which time  $\omega(0) = \omega_0$ . Equation (12.99) now looks like

$$\dot{\omega}_0 = \frac{\omega}{2\pi} + \frac{\epsilon}{4\pi} \sin 4\pi \omega_0. \quad (12.99')$$

The zeroth-order solution is

$$2\pi \omega_0^{(0)} = \omega_0 t,$$

where the constant term has been set zero by suitable choice of the initial phase. To first order in  $\epsilon$ , Eq. (12.99') becomes

$$\dot{\omega}_0^{(1)} = \frac{\omega_0(1 + \epsilon t)}{2\pi} + \frac{\epsilon}{4\pi} \sin 2\omega_0 t, \quad (12.102)$$

with the solution

$$2\pi \omega_0^{(1)} = \omega_0 t + \frac{\epsilon}{2} \left( \omega_0 t^2 + \frac{1 - \cos 2\omega_0 t}{2\omega_0} \right). \quad (12.103)$$

Correspondingly the equation for  $\dot{J}_0$  correct to second order in  $\epsilon$  can be written as

$$\frac{d \ln J_0}{dt} = -\epsilon \cos \left[ 2\omega_0 t + \epsilon \left( \omega_0 t^2 + \frac{1 - \cos 2\omega_0 t}{2\omega_0} \right) \right].$$

Expanding the cosine, treating the term in  $\epsilon$  as a small quantity to first order, the derivative reduces to

$$\frac{d \ln J_0}{dt} = -\epsilon \cos 2\omega_0 t + \epsilon^2 \left( \omega_0 t^2 + \frac{1 - \cos 2\omega_0 t}{2\omega_0} \right) \sin 2\omega_0 t.$$

To find the secular behavior, this equation can be averaged over the period of the motion as it is at  $t = 0$ , that is, over an interval  $\tau = 2\pi/\omega_0$ . In the averaging, almost all terms on the right drop out, except the first inside the parentheses, involving  $t^2$ . The final result is

$$\frac{d \ln J}{dt} = \frac{\pi \epsilon^2}{\omega_0} = \frac{\omega_0 \delta^2}{4\pi}, \quad (12.104)$$

where  $\delta = \epsilon \tau$ , that is, fractional change in  $\omega$  over the period  $\tau$ . Correspondingly, the fractional secular change in  $J$  over the period is

$$\frac{\Delta J}{J} = \frac{\delta^2}{2}. \quad (12.105)$$

As expected from the more general considerations, the secular change in the action variable has no term in first order in  $\epsilon$ . Only by retaining quantities of the order  $\epsilon^2 = (\dot{\omega}/\omega)^2$  do we find any nonvanishing long-term change in  $J$ .

The adiabatic invariance of the action variables has proven to be especially useful in applications involving the motion of charged particles in electromagnetic fields. One of the simplest instances, and one with important practical consequences, concerns the motion of electrons in a uniform (or nearly uniform) constant magnetic field. As is well known, the charged particle in such a situation circles around the magnetic field lines. At the most basic level, this can be shown from Newton's equations of motion. The Lorentz force in a constant magnetic field  $\mathbf{B}$  is  $(\mathbf{v} \times q\mathbf{B})$ ; hence, the equation of motion, Eq. (1.4), is

$$\frac{d\mathbf{v}}{dt} = \mathbf{v} \times \frac{q\mathbf{B}}{m}. \quad (12.106)$$

Equation (12.106) says the velocity vector  $\mathbf{v}$  rotates, without change of magnitude, about the direction of the magnetic field, with an angular frequency

$$\omega_c = -\frac{qB}{m}. \quad (12.107)$$

The frequency, called the cyclotron frequency, has a value twice the Larmor frequency of Eq. (5.104) (cf. Eq. (7.154)).

An equivalent derivation can be formulated in terms of Lagrangian mechanics. It was shown, in Section 5.9, that the Lagrangian in this case can be written as

$$L = \frac{mv^2}{2} + \mathbf{M} \cdot \mathbf{B}, \quad (12.108)$$

where  $\mathbf{M}$  is magnetic moment of the moving particle defined in terms of its angular momentum  $\mathbf{L}$  by

$$\mathbf{M} = \frac{q\mathbf{L}}{2m}. \quad (12.109)$$

(Cf. Eq. (5.108).) In cylindrical coordinates with the  $z$  axis along the direction of  $\mathbf{B}$ , the component of  $\mathbf{M}$  parallel to  $\mathbf{B}$  is

$$M_z = \frac{qr^2\dot{\theta}}{2}, \quad (12.110)$$

and the Lagrangian is

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2) + \frac{q}{2}Br^2\dot{\theta}. \quad (12.111)$$

Since  $\theta$  is cyclic in the Lagrangian, the corresponding canonical momentum  $p_\theta$ ,

$$p_\theta = mr^2\dot{\theta} + \frac{qBr^2}{2}, \quad (12.112)$$

is a constant of the motion. Further, the radial equation of motion is

$$m\ddot{r} - r\dot{\theta}(m\dot{\theta} + qB) = 0. \quad (12.113)$$

A steady-motion solution to Eqs. (12.112) and (12.113) corresponds to  $r$  and  $\dot{\theta}$  constant, with  $\dot{\theta}$  having the cyclotron value

$$\dot{\theta} = \omega_c \equiv -\frac{qB}{m}, \quad (12.114)$$

in agreement with Eq. (12.107). In this case,  $p_\theta = -(qBr^2/2)$  and the action variable corresponding to  $\theta$  is

$$J_\theta = \oint p_\theta d\theta = -\pi qBr^2. \quad (12.115)$$

By (12.110), we can write

$$qr^2 = \frac{2M}{\omega_c}$$

(as  $M_z$  is equal to  $M$  for this motion), and therefore  $J_\theta$  can also be written as

$$J_\theta = -\frac{2\pi MB}{\omega_c} = \frac{2\pi m}{q}M. \quad (12.116)$$

The adiabatic invariance theorem implies that under sufficiently slow variation of the magnetic field  $J_\theta$  remains constant. Equation (12.116) says that the magnetic moment is similarly invariant adiabatically. An alternative statement, on the basis of Eq. (12.115), is that  $B$  times the area  $\pi r^2$  of the orbit (that is, the number of lines of force threading through the orbit) remains constant.

An adiabatic variation of  $B$  might arise if the magnetic field configuration remained static but was slightly nonuniform. If then the particle had a small  $z$  component of velocity, the resultant drift would move the particle slowly into regions of different  $B$  values. From Eqs. (12.114), (12.115), and (12.116), it follows simply that the kinetic energy of motion around the lines of  $B$  is

$$T_{(\theta)} = \frac{mr^2\dot{\theta}^2}{2} = MB. \quad (12.117)$$

Suppose a charged particle drifts in the direction of increasing  $B$ ; by Eq. (12.117), the kinetic energy of rotation increases. As the total kinetic energy is conserved, the kinetic energy of longitudinal drift  $m\dot{z}^2/2$  along the lines of force must decrease. Eventually, the drift velocity  $\dot{z}$  goes to zero and the motion reverses in direction. If it can be arranged that  $B$  eventually increases in the other direction, the charged particle will remain confined, drifting back and forth between the two ends—the principle of the so-called mirror confinement. The mirror principle is used to contain hot plasmas for thermonuclear energy generation. The complete story is of course more complicated, but the significance of the adiabatic invariance of  $M$  is clearly demonstrated.

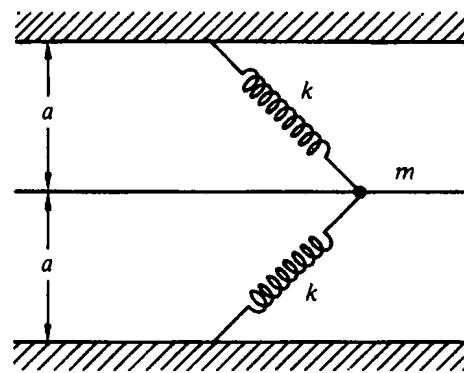
We have seen that almost all phenomena of small oscillations about steady-state or steady motion can be described in terms of harmonic oscillators. In consequence, there is a good deal of practical interest in questions of the invariance of  $J$  for a harmonic oscillator under slow, and not so slow, variations of a parameter. The study of oscillations in charged particle accelerators, for example, has led to a number of new insights.

It has been possible to sketch here only the highlights of the subject of adiabatic invariants. The ramifications of the field go into many areas of classical and quantum physics and of mathematics.

## EXERCISES

1. By the method of time-dependent perturbation theory, carry the solution for the linear harmonic oscillator (in which the potential is considered a perturbation on the free particle motion) out through *third*-order terms, assuming the initial condition  $\beta_0 = 0$ . Find expressions for both  $x$  and  $p$  as functions of time and show that they agree with the corresponding terms in the expansion of the usual harmonic solutions.
2. A mass point  $m$  hangs at one end of a vertically hung Hook's-law spring of force constant  $k$ . The other end of the spring is oscillated up and down according to  $z_1 = a \cos \omega_1 t$ . By treating  $a$  as a small quantity, obtain a first-order solution to the motion of  $m$  in time, using time dependent perturbation theory. What happens as  $\omega_1$  approaches the unperturbed frequency  $\omega_0$ ?
3. (a) A linear harmonic oscillator of force constant  $k$  has its mass suddenly increased by a fractional amount  $\epsilon$ . Use first-order time-independent perturbation theory, to find the resultant shift in the frequency of the oscillator to first order in  $\epsilon$ . Compare your results with the exact solution and discuss.  
 (b) Repeat part (a), for the effect of increasing  $k$  by a fractional amount  $\epsilon$ .
4. Carry out a consistent second-order perturbation calculation (using whichever method you choose) of the correction to the frequency of a plane pendulum as the result of a finite amplitude of oscillation. All terms of order  $\lambda^2$  should be retained in the Hamiltonian and in the perturbation treatment.
5. A mass particle is constrained to move in a horizontal straight line and is attached to the ends of two ideal springs of equal force constants, as shown in the diagram. The

unstretched length of each spring is  $b \leq a$ . Use perturbation theory to first-order to find the lowest order correction to the frequency of oscillation for finite amplitude of oscillation. What happens as  $a$  approaches  $b$  in magnitude?



6. (a) Show that to lowest order in correction terms the relativistic (but noncovariant) Hamiltonian for the one-dimensional harmonic oscillator has the form

$$H = \frac{1}{2m}(p^2 + m^2\omega^2 q^2) - \frac{1}{8} \frac{p^4}{m^3 c^2}.$$

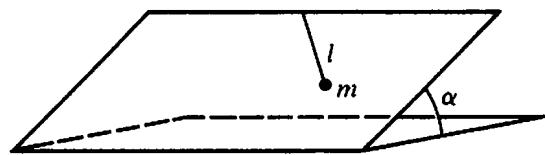
- (b) Use first order perturbation theory to calculate the lowest-order relativistic correction to the frequency of the harmonic oscillator. Express your result as a fractional change in the frequency.

7. A plane isotropic harmonic oscillator is perturbed by a change in the Hamiltonian of the form

$$\epsilon H_1 = bp_x^2 p_y^2$$

where  $b$  is a constant. Use time-independent perturbation theory to first order find the shift in the frequencies.

8. A model of the atomic Stark effect can be made by taking the Kepler elliptic orbit in a plane and perturbing it by a potential  $\Delta V = -Kx$ . Use perturbation theory to first order to determine what happens to the frequencies of motion. This model can also be used as a first approximation to the effect of the light pressure of solar radiation on the orbit of an Earth satellite.
9. By considering the work done to alter adiabatically the length  $l$  of a plane pendulum, prove by elementary means the adiabatic invariance of  $J$  for the plane pendulum in the limit of vanishing amplitude.
10. Consider the system described in Exercise 13 of Chapter 10. Suppose the parameter  $F$  is slowly varied from an initial value. What happens to the energy of the particle? The amplitude of oscillation? The period?



- 11.** A plane pendulum of small amplitude is constrained to move on an inclined plane, as shown in the accompanying figure. How does its amplitude change when the inclination angle  $\alpha$  of the plane is changed slowly?

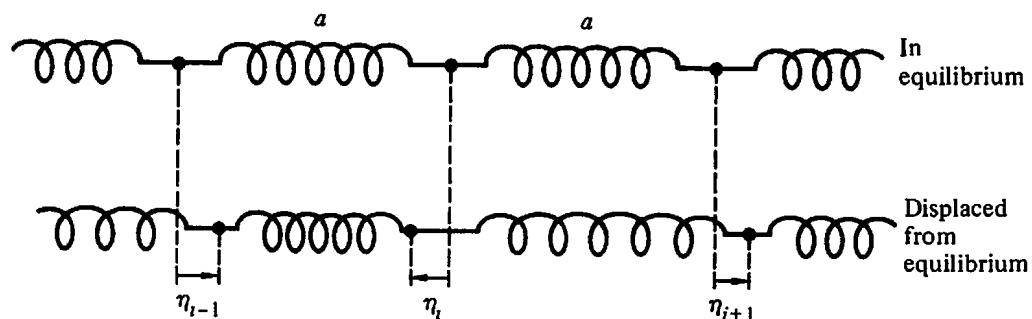
# CHAPTER 13

## Introduction to the Lagrangian and Hamiltonian Formulations for Continuous Systems and Fields

All the formulations of mechanics discussed thus far have been devised for treating systems with a finite or at most a denumerably infinite number of degrees of freedom. There are some mechanical problems, however, that involve continuous systems, as, for example, the problem of a vibrating elastic solid. Here each point of the continuous solid partakes in the oscillations, and the complete motion can only be described by specifying the position coordinates of *all* points. It is not difficult to modify the previous formulations of mechanics so as to handle such problems. The concepts of field theory can be developed by approximating the continuous system with a discrete system, solving that problem, and taking the continuous limit.

### 13.1 ■ THE TRANSITION FROM A DISCRETE TO A CONTINUOUS SYSTEM

We shall apply this procedure to an infinitely long elastic rod that can undergo small longitudinal vibrations, that is, oscillatory displacements of the particles of the rod parallel to the axis of the rod. A system composed of discrete particles that approximates the continuous rod is an infinite chain of equal mass points spaced a distance  $a$  apart and connected by uniform massless springs having force constants  $k$  (cf. Fig. 13.1). It will be assumed that the mass points can move only along the length of the chain. The discrete system will be recognized as an extension of the linear polyatomic molecule discussed in Section 6.4. We can therefore



**FIGURE 13.1** A discrete system of equal mass points connected by springs, as an approximation to a continuous elastic rod.

obtain the equations describing the motion by the customary techniques for small oscillations. Denoting the displacement of the  $i$ th particle from its equilibrium position by  $\eta_i$ , the kinetic energy is

$$T = \frac{1}{2} \sum_i m \dot{\eta}_i^2, \quad (13.1)$$

where  $m$  is the mass of each particle. The corresponding potential energy is the sum of the potential energies of each spring as the result of being stretched or compressed from its equilibrium length (cf. Section 6.4):

$$V = \frac{1}{2} \sum_i k(\eta_{i+1} - \eta_i)^2. \quad (13.2)$$

Combining Eqs. (13.1) and (13.2), the Lagrangian for the system is

$$L = T - V = \frac{1}{2} \sum_i [m \dot{\eta}_i^2 - k(\eta_{i+1} - \eta_i)^2], \quad (13.3)$$

which can also be written as

$$L = \frac{1}{2} \sum_i a \left[ \frac{m}{a} \dot{\eta}_i^2 - ka \left( \frac{\eta_{i+1} - \eta_i}{a} \right)^2 \right] = \sum_i a L_i, \quad (13.4)$$

where  $a$  is the equilibrium separation between the points (cf. Fig. 13.1). The resulting Lagrange equations of motion for the coordinates  $\eta_i$  are

$$\frac{m}{a} \ddot{\eta}_i - ka \left( \frac{\eta_{i+1} - \eta_i}{a^2} \right) + ka \left( \frac{\eta_i - \eta_{i-1}}{a^2} \right) = 0. \quad (13.5)$$

The particular form of  $L$  in Eq. (13.4), and of the corresponding equations of motion, has been chosen for convenience in going to the limit of a continuous rod as  $a$  approaches zero. It is clear that  $m/a$  reduces to  $\mu$ , the mass per unit length of the continuous system, but the limiting value of  $ka$  may not be so obvious. For an elastic rod obeying Hooke's law, it will be remembered that the extension of the rod *per unit length* is directly proportional to the force or tension exerted on the rod, a relation that can be written as

$$F = Y\xi,$$

where  $\xi$  is the elongation per unit length and  $Y$  is Young's modulus. Now the extension of a length  $a$  of a discrete system, per unit length, will be  $\xi = (\eta_{i+1} - \eta_i)/a$ . The force necessary to stretch the spring by this amount is

$$F = k(\eta_{i+1} - \eta_i) = ka \left( \frac{\eta_{i+1} - \eta_i}{a} \right),$$

so that  $ka$  must correspond to the Young's modulus of the continuous rod. In going from the discrete to the continuous case, the integer index  $i$  identifying the particular mass point becomes the continuous position coordinate  $x$ ; instead of the variable  $\eta_i$  we have  $\eta(x)$ . Further, the quantity

$$\frac{\eta_{i+1} - \eta_i}{a} = \frac{\eta(x+a) - \eta(x)}{a}$$

occurring in  $L_i$  obviously approaches the limit

$$\frac{d\eta}{dx},$$

as  $a$ , playing the role of  $dx$ , approaches zero. Finally, the summation over a discrete number of particles becomes an integral over  $x$ , the length of the rod, and the Lagrangian (13.4) appears as

$$L = \frac{1}{2} \int \left[ \mu \dot{\eta}^2 - Y \left( \frac{d\eta}{dx} \right)^2 \right] dx. \quad (13.6)$$

In the limit as  $a$  goes to zero, the last two terms in the equation of motion (13.5) become

$$\lim_{a \rightarrow 0} -\frac{Y}{a} \left[ \left( \frac{d\eta}{dx} \right)_x - \left( \frac{d\eta}{dx} \right)_{x-a} \right],$$

which clearly defines a second derivative of  $\eta$ . Hence, the equation of motion for the continuous elastic rod is

$$\mu \frac{d^2\eta}{dt^2} - Y \frac{d^2\eta}{dx^2} = 0, \quad (13.7)$$

the familiar wave equation in one dimension with the propagation velocity

$$v = \sqrt{\frac{Y}{\mu}}. \quad (13.8)$$

Equation (13.8) is the well-known formula for the velocity of longitudinal elastic waves.

This simple example is sufficient to illustrate the salient features of the transition from a discrete to a continuous system. The most important fact to grasp is the role played by the position coordinate  $x$ . It is *not* a generalized coordinate; it serves merely as a continuous index replacing the discrete  $i$ . Just as each value of  $i$  corresponds to a different one of the generalized coordinates,  $\eta_i$ , of the system, so here for each value of  $x$  there is a generalized coordinate  $\eta(x)$ . Since  $\eta$  depends also upon the continuous variable  $t$ , we should perhaps write more accurately  $\eta(x, t)$ , indicating that  $x$ , like  $t$ , can be considered as a parameter entering into the Lagrangian. If the continuous system were three-dimensional,

rather than one-dimensional as here, the generalized coordinates would be distinguished by three continuous indices  $x, y, z$ , and would be written as  $\eta(x, y, z, t)$ . Note that the quantities  $x, y, z$ , and  $t$  are completely independent of each other, and appear only as explicit variables in  $\eta$ . Derivatives of  $\eta$  with respect to any of them can therefore always be written as total derivatives without any ambiguity. Equation (13.6) also shows that the Lagrangian appears as an integral over the continuous index  $x$ ; in the corresponding three-dimensional case the Lagrangian would have the form

$$L = \int \int \int \mathcal{L} dx dy dz, \quad (13.9)$$

where  $\mathcal{L}$  is known as the *Lagrangian density*. For the longitudinal vibrations of the continuous rod the Lagrangian density is

$$\mathcal{L} = \frac{1}{2} \left[ \mu \left( \frac{d\eta}{dt} \right)^2 - Y \left( \frac{d\eta}{dx} \right)^2 \right], \quad (13.10)$$

corresponding to the continuous limit of the quantity  $L_i$ , appearing in Eq. (13.4). It is the Lagrangian density, rather than the Lagrangian itself, that will be used to describe the motion of the system.

## 13.2 ■ THE LAGRANGIAN FORMULATION FOR CONTINUOUS SYSTEMS

It will be noted from Eq. (13.9) that  $\mathcal{L}$  for the elastic rod, besides being a function of  $\dot{\eta} \equiv \partial\eta/\partial t$ , also involves a spatial derivative of  $\eta$ , namely,  $\partial\eta/\partial x$ ;  $x$  and  $t$  thus play a similar role as parameters of the Lagrangian density. If there were local forces present in addition to the nearest neighbor interactions, then  $\mathcal{L}$  would be a function of  $\eta$  itself as well as of the spatial gradient of  $\eta$ . Of course, in the general case  $\mathcal{L}$  might well be an explicit function of  $x$  and  $t$  also. So the Lagrangian density for any one-dimensional continuous system would appear as a function of the form

$$\mathcal{L} = \mathcal{L} \left( \eta, \frac{d\eta}{dx}, \frac{d\eta}{dt}, x, t \right). \quad (13.11)$$

The total Lagrangian, following Eq. (13.10), is then the integral of  $\mathcal{L}$  over the range of  $x$  defining the system, and Hamilton's principle, Eq. (2.2), in the limit of the continuous system appears as

$$\delta I = \delta \int_1^2 \int \mathcal{L} dx dt = 0. \quad (13.12)$$

If Hamilton's principle for the continuous system is to have any usefulness, it must be possible to derive the continuous limit of the equation of motion, for ex-

ample, Eq. (13.7), directly by variation of the double integral of  $\mathcal{L}$  in Eq. (13.12). We can carry out this variation by methods that differ only slightly from those used in Chapter 2 for a discrete system. The variation is only on  $\eta$  and its derivatives; the parameters  $x$  and  $t$  are not affected by the variation either directly or in the ranges of integration. Just as the variation of  $\eta$  is taken to be zero at the end points  $t_1$  and  $t_2$ , so the variation of  $\eta$  at the limits  $x_1$  and  $x_2$  of the integration in  $x$  is also to be zero. As in Section 2.2, a suitable varied path of integration in the  $\eta$  space can be obtained, for example, by choosing  $\eta$  from a one-parameter family of possible  $\eta$  functions:

$$\eta(x, t; \alpha) = \eta(x, t; 0) + \alpha \zeta(x, t). \quad (13.13)$$

Here  $\eta(x, t; 0)$  stands for the correct function that will satisfy Hamilton's principle, and  $\zeta$  is any well-behaved function that vanishes at the end points in  $t$  and  $x$ . If  $I$  is considered as a function of  $\alpha$ , to be an extremum for  $\eta(x, t; 0)$  the derivative of  $I$  with respect to  $\alpha$  vanishes at  $\alpha = 0$ . By straightforward differentiation,

$$\frac{dI}{d\alpha} = \int_{t_1}^{t_2} \int_{x_1}^{x_2} dx dt \left[ \frac{\partial \mathcal{L}}{\partial \eta} \frac{\partial \eta}{\partial \alpha} + \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} \frac{\partial}{\partial \alpha} \left( \frac{d\eta}{dt} \right) + \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} \frac{\partial}{\partial \alpha} \left( \frac{d\eta}{dx} \right) \right]. \quad (13.14)$$

Because the variation of  $\eta$ , that is,  $\alpha \zeta$ , vanishes at the end points, integration by parts in  $x$  and  $t$  yields the relations

$$\int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} \frac{\partial}{\partial \alpha} \left( \frac{d\eta}{dt} \right) dt = - \int_{t_1}^{t_2} \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} \right) \frac{\partial \eta}{\partial \alpha} dt,$$

and

$$\int_{x_1}^{x_2} \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} \frac{\partial}{\partial \alpha} \left( \frac{d\eta}{dx} \right) dx = - \int_{x_1}^{x_2} \frac{d}{dx} \left( \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} \right) \frac{\partial \eta}{\partial \alpha} dx.$$

Hamilton's principle can therefore be written as

$$\int_{t_1}^{t_2} \int_{x_1}^{x_2} dx dt \left[ \frac{\partial \mathcal{L}}{\partial \eta} - \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} \right) - \frac{d}{dx} \left( \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} \right) \right] \left( \frac{\partial \eta}{\partial \alpha} \right)_0 = 0, \quad (13.15)$$

and by the same arguments as in Section 2.2 the arbitrary nature of the varied path implies the vanishing of the expression in the brackets:

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} \right) + \frac{d}{dx} \left( \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} \right) - \frac{\partial \mathcal{L}}{\partial \eta} = 0. \quad (13.16)$$

The Euler-Lagrange equations (13.16) (cf. Eq. (2.18)) is the appropriate form of the equation of motion as derived from Hamilton's principle, Eq. (13.12).

A system of  $n$  discrete degrees of freedom will have  $n$  Lagrange equations of motion; for the continuous system with an infinite number of degrees of freedom we seem to obtain only one Lagrange equation! It must be remembered, however, that the equation of motion for  $\eta$  is a differential equation involving the time only, and in that sense Eq. (13.15) furnishes a separate equation of motion for each value of  $x$ . The continuous nature of the indices  $x$  appears in that Eq. (13.15) is a partial differential equation in the two variables  $x$  and  $t$ , yielding  $\eta$  as  $\eta(x, t)$ .

For the specific instance of longitudinal vibrations in an elastic rod, it is seen from the form of the Lagrangian density, Eq. (13.10), that

$$\frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} = \mu \frac{d\eta}{dt}, \quad \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} = -Y \frac{d\eta}{dx}, \quad \frac{\partial \mathcal{L}}{\partial \eta} = 0.$$

Thus, as desired, Eq. (13.16), reduces properly to the equation of motion, Eq. (13.7).

The Lagrangian formulation developed here for one-dimensional continuous systems needs obviously to be extended to two- and three-dimensional situations, for example, a general elastic solid. Further, instead of one field quantity  $\eta$  there may be several; for example, displacement from an equilibrium position would be described by a spatial vector  $\eta$  with three components. There is no difficulty in carrying out the mathematical steps for the more general situation in close parallelism to the one-component one-dimensional case. However, the formulas become lengthy and cumbersome if written in the same manner, especially in view of the two tiers of derivatives. Considerable gain in notational simplicity can be achieved by noticing that time  $t$  and the spatial coordinates  $x, y, z$  play the same type of mathematical role in Hamilton's principle. The field quantities are functions of the coordinates of both time and space that are to be treated as independent variables. No variation of the field quantities occurs at the limits of integration in Hamilton's principle over both time and space.

It is mathematically convenient to think in terms of a four-dimensional space with coordinates  $x^0 = ct, x^1 = x, x^2 = y, x^3 = z$ . No physical significance is implied for this space. The  $c$  in  $x^0$  is the speed of light used only to convert the units of  $x^0$  to the same as those used for  $x^i$ . The entire tensor formalism developed in Chapter 7 applies. The metric tensor  $g$  will have a Euclidean metric with the Galilean transformation group as the allowed coordinate transformations on the space components of the metric tensor restricted by  $g_{i0} = g_{0i} = 0$ . A Roman letter superscript refers only to the three coordinates of the physical space, a Greek letter superscript or subscript refers to all four coordinates. Use of the summation convention with respect to repeated indices will be resumed for the rest of the chapter. The various components of the field quantities will be symbolized by a subscript  $\rho$ , which may cover a multitude of forms. At times, it will stand for a single index having two, three, four, or more values. Or it may stand for multiple indices. Thus, if the field quantity is a spatial tensor of second rank, then  $\rho$  really refers to two subscript indices. Finally, a derivative of the field quantities with respect to any one of the four coordinates  $x^\mu$  will be denoted by the subscript  $v$

separated from  $\rho$  by a comma. Where there is only one field quantity the index does not appear. Examples are

$$\eta_{\rho,\nu} \equiv \frac{d\eta_\rho}{dx^\nu}; \quad \eta_{,j} \equiv \frac{d\eta}{dx^j}; \quad \eta_{i,\mu\nu} = \frac{d^2\eta_i}{dx^\mu dx^\nu}. \quad (13.17)$$

Only the derivatives of the field quantities will be symbolized in this manner.

In this notation, the most general form of the Lagrangian density to be considered here is written as

$$\mathcal{L} = \mathcal{L}(\eta_\rho, \eta_{\rho,\nu} x^\nu). \quad (13.18)$$

The total Lagrangian is then an integral over three-space:

$$L = \int \mathcal{L}(dx^i), \quad (13.19)$$

but it rarely occurs explicitly. Hamilton's principle appears as an integral over a region in 4-space:

$$\delta I = \delta \int \mathcal{L}(dx^\mu) = 0, \quad (13.20)$$

where the variation of the  $\eta_\rho$  vanishes at the bounding surface  $S$  of the region of integration. The derivation of the corresponding Euler–Lagrange equations of motion proceeds symbolically as before. We consider a one-parameter set of varied functions that reduce to  $\eta_\rho(x^\nu)$  as the parameter  $\alpha$  goes to zero. As previously, a possible suitable set can be constructed, for example, by adding to  $\eta_\rho$  the product  $\alpha \zeta_\rho$ , where  $\zeta_\rho(x^\nu)$  are convenient arbitrary functions vanishing on the bounding surface. The vanishing of the variation of  $I$  is equivalent to setting the derivative of  $I$  with respect to  $\alpha$  equal to zero:\*

$$\frac{dI}{d\alpha} = \int \left( \frac{\partial \mathcal{L}}{\partial \eta_\rho} \frac{\partial \eta_\rho}{\partial \alpha} + \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{\partial \eta_{\rho,\nu}}{\partial \alpha} \right) (dx^\mu).$$

Integration by parts yields

$$\begin{aligned} \frac{dI}{d\alpha} &= \int \left[ \frac{\partial \mathcal{L}}{\partial \eta_\rho} - \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \right) \right] \frac{\partial \eta_\rho}{\partial \alpha} (dx^\mu) \\ &\quad + \int (dx^\mu) \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{\partial \eta_\rho}{\partial \alpha} \right). \end{aligned} \quad (13.21)$$

The second integral vanishes in the limit as  $\alpha$  goes to zero, as can be seen in various ways. We can examine it term by term: carrying out the integration for the particular  $x^\nu$  of each derivative term, which then vanishes because the derivative with respect to  $\alpha$  is zero at the end points. Or the integral can be transformed by

\*Unless otherwise noted, the summation convention will be used in the remainder of this chapter, for all types of subscript-superscript pairs.

a four-dimensional divergence theorem into an integral over the surface bounding the region of integration in 4-space. The surface integral again vanishes because the variation of  $\eta_\rho$  in the vicinity of the correct field functions is zero on the surface. Equation (13.21) in the limit as  $\alpha$  goes to zero therefore reduces to

$$\left( \frac{dI}{d\alpha} \right)_0 = \int (dx^\mu) \left[ \frac{\partial \mathcal{L}}{\partial \eta_\rho} - \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \right) \right] \left( \frac{\partial \eta_\rho}{\partial \alpha} \right)_0. \quad (13.22)$$

Again, the arbitrary nature of the variation of each  $\eta_\rho$  means that Eq. (13.22) is satisfied only when each of the square brackets vanishes:

$$\frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \right) - \frac{\partial \mathcal{L}}{\partial \eta_\rho} = 0. \quad (13.23)$$

Equations (13.23) represent a set of partial differential equations for the field quantities, with as many equations as there are different values of  $\rho$ . It may be worth repeating that since the space coordinates  $x^i$  are indices for the field quantities, each of Eqs. (13.23) in effect corresponds to an entire set of Lagrange differential equations of motion in the discrete case.

For a one-dimensional continuous system, where  $\nu$  takes on only the values 0 and 1, Eq. (13.23) expands to the same form as Eq. (13.16). The compactness of the notation is evident even in so simple an example. Although we have used covariant notation, the use of a four-dimensional space for symbolic convenience in no way requires covariant behavior (in the physicist's sense of the word) of any of the quantities in that space.

For discrete systems, the Lagrangian is uncertain to a total time derivative of an arbitrary function of the generalized coordinates and time. With continuous systems, the corresponding statement is that  $\mathcal{L}$  is uncertain to any "4-divergence," that is, to a term of the form

$$\frac{dF_\nu(\eta_\rho, x^\mu)}{dx^\nu} \quad (13.24)$$

where the  $F_\nu$  are any four (differentiable) functions of the field quantities  $\eta_\rho$  and the coordinates  $x^\mu$ . That such a term makes no contribution to the variation of the action integral is obvious. Application of the divergence theorem in 4-space converts the volume integral into an integral over the bounding surface where the variation of  $F_\nu$  is zero. In symbols, the relevant variation can be written

$$\delta \int (dx^\mu) \frac{dF_\nu(\eta_\rho, x^\mu)}{dx^\nu} = \delta \int F_\nu(\eta_\rho, x^\mu) d\sigma^\nu = 0, \quad (13.25)$$

where  $d\sigma^\nu$  represents the components of an element of surface (in Euclidean 4-space) oriented along the direction of the outward normal.

The Lagrangian formulation for a continuous set of generalized coordinates has been developed in order to treat continuous mechanical systems such as an

elastic solid in longitudinal oscillation, or a gas vibrating in such a manner as to set up acoustic waves. As has been implied, the formulation may also be used, even in the absence of a mechanical system, to describe the equations governing a *field*. Mathematically, a field is no more than a set of one or more independent functions of space and time, and the generalized coordinates fit this definition. There is no requirement that the field be related to some underlying mechanical system. In thus breaking the connection between the Lagrangian field description and purely mechanical motion, we are merely recapitulating the history of physics. For example, the electromagnetic field was long thought of in terms of the elastic vibrations of a mysterious ether. Only in recent times was it generally realized that the ether had no other role than being the subject of the verb “to undulate.” We recognize equally well that the variational procedures developed here also stand independent of the notion of a continuous mechanical system, and that they serve to furnish the equations describing any spacetime field. Hamilton’s principle then becomes in effect a convenient and compact description of the field, one that upon expansion leads to the field equations.

In addition to implying the field equations, the Lagrangian density has more to tell us about the physical nature of the field. As with systems of a discrete number of degrees of freedom, the structure of the Lagrangian also contains information on conserved properties of the system. One such set of conservation theorems is discussed in the next section.\*

### 13.3 ■ THE STRESS-ENERGY TENSOR AND CONSERVATION THEOREMS

An analog to the conservation of Jacobi’s integral in point mechanics found in Section 2.6, can be derived here, and in much the same manner. All we have to remember is that the treatment of time must be extended in parallel fashion to the  $x^i$  since they are all independent parameters in  $\mathcal{L}$ . Thus, instead of the time derivative of  $L$ , we seek to evaluate the total derivative of  $\mathcal{L}$  with respect to  $x^\mu$ :

$$\frac{d\mathcal{L}}{dx^\mu} = \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\mu}} + \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_{\rho,\mu\nu} + \frac{\partial \mathcal{L}}{\partial x^\mu}. \quad (13.26)$$

By the equations of motion, Eq. (13.23), this becomes (with a slight change in notation),

$$\begin{aligned} \frac{d\mathcal{L}}{dx^\mu} &= \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \right) \eta_{\rho,\mu} + \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{d\eta_{\rho,\mu}}{dx^\nu} + \frac{\partial \mathcal{L}}{\partial x^\mu} \\ &= \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_{\rho,\mu} \right) + \frac{\partial \mathcal{L}}{\partial x^\mu}. \end{aligned} \quad (13.27)$$

\*A more general attack on the conservation properties inherent in the Lagrangian will be found in Section 13.7 on Noether’s theorem.

Combining total derivatives, this can be written

$$\frac{d}{dx^\nu} \left[ \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_{\rho,\mu} - \mathcal{L} \delta_{\mu\nu} \right] = -\frac{\partial \mathcal{L}}{\partial x^\mu}. \quad (13.28)$$

Let us suppose, now, that  $\mathcal{L}$  does not depend explicitly upon  $x^\mu$ . This usually means that  $\mathcal{L}$  represents a free field, that is, contains no external driving sources or sinks that interact with the field at explicit space points and with given time dependence. In effect, this means no interaction between the field and point particles moving in space and time through the field. Under this condition, Eq. (13.28) takes on the form of a set of divergence conditions,

$$\frac{dT_\mu{}^\nu}{dx^\nu} = T_\mu{}^\nu,_\nu = 0 \quad (13.29)$$

on a quantity with the form of a 4-tensor of the second rank:

$$T_\mu{}^\nu = \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_{\rho,\mu} - \mathcal{L} \delta_\mu{}^\nu. \quad (13.30)$$

That these equations have only the *form* of tensor equations in 4-space is emphasized because as yet the 4-space has no transformation properties—space and time are still distinct—and there is no transformation requirement on  $T_\mu{}^\nu$ . However, the space portions of these quantities do behave like vectors and tensors in ordinary space; that is,  $T_{ij}$  are the components of a three-dimensional tensor of the second rank. Before considering the possible transformations, we will determine the physical meaning of  $T_\mu{}^\nu$ .

The similarity between  $T_\mu{}^\nu$  and Jacobi's integral, Eq. (2.54), is obvious. It becomes especially clear for the component  $T_0^0$ :

$$T_0^0 = \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\rho} \dot{\eta}_\rho - \mathcal{L}. \quad (13.31)$$

In mechanical systems, the Lagrangian density often has the form  $\mathcal{L} = \mathcal{T} - \mathcal{V}$ , the difference between a kinetic energy density and a potential energy density. This is the case, for example, with the Lagrangian densities for the elastic rod, with the kinetic energy density having the form of one-half the mass density times a square of the displacement velocity:

$$\mathcal{T} = \frac{1}{2} \mu \dot{\eta}_\rho \dot{\eta}_\rho.$$

By the same arguments as used in discrete mechanics,  $T_0^0$  can then be identified as a total energy density.

The corresponding identification tags to be placed on the other elements of  $T_\mu{}^\nu$  can be suggested by writing the set of Eqs. (13.29) as

$$\frac{dT_\mu{}^0}{dt} + \frac{dT_\mu{}^j}{dx} = 0, \quad (13.32)$$

or

$$T_{\mu,\nu}^v = \frac{dT_\mu^0}{cdt} + \frac{dT_\mu^i}{dx^i} = \frac{dT_\mu^0}{cdt} + \nabla \cdot T_\mu = 0 \quad (13.33)$$

where  $T_\mu$ , whose components are  $T_\mu^\nu$ , are a set of 4-space vectors. In either form, Eqs. (13.32) or (13.33) appear as equations of continuity, which is to say that the time rate of change of some density plus the divergence of some corresponding flux or current density vanishes. In turn, the equations of continuity imply the conservation of some integral quantities providing the field volume is finite; that is, the field can be contained within a volume beyond which the field quantities are zero, defined, in such a case, integral quantities  $R_\mu$  by

$$R_\mu = \int T_\mu^0 dV, \quad (13.34)$$

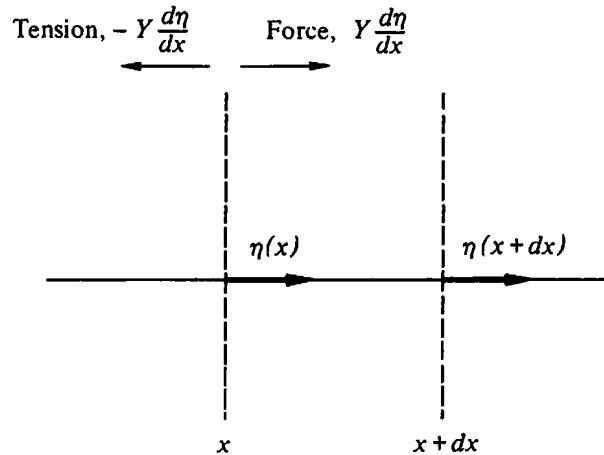
where the volume integral extends beyond the region containing the field. Then, by Eqs. (13.33),

$$\frac{dR_\mu}{dt} = \int \nabla \cdot T_\mu dV = \int T_\mu \cdot dA = 0. \quad (13.35)$$

It is because of these conservation theorems, derived from Eq. (13.29), that the four arrays  $T_\mu^\nu$ ,  $\mu = 0, 1, 2, 3$  are known as conserved currents, in analogy with the conservation equations for electromagnetic current.

We should therefore expect  $T_0$  to play the role of the components of an energy current density. That this is reasonable can be seen again from considerations of the longitudinal vibration field in an elastic rod. Imagine the rod divided by an imaginary cut at point  $x$  (cf. Fig. 13.2). From the considerations that led to the Lagrangian, Eq. (13.6), the force exerted by the part of the rod on the right to extend the part that is to the left of the cut is

$$Y \frac{d\eta}{dx}. \quad (13.36)$$



**FIGURE 13.2** Diagram illustrating calculation of energy current density in elastic rod.

Hence, there is a tension at  $x$  in the left-hand portion of equal magnitude but of opposite direction. Further, the left-hand portion is being stretched by an amount that at  $x$  is  $\eta$ , and the rate at which this extension changes in time is  $\dot{\eta}$ . Hence, the rate of work being done by the tension at the cut is

$$-\dot{\eta} Y \frac{d\eta}{dx}, \quad (13.37)$$

which is thus the rate at which energy is being transferred to the right per unit time. Comparison shows that this is exactly  $T_0^1$  for the appropriate Lagrangian density of Eq. (13.10). If  $T_0^0$  is an energy density then the quantity,  $R_0$ , of Eq. (13.34) can be identified as the total energy in the field. The fourth component of the conservation equation (13.35) therefore says that the total field energy is conserved if  $T_0^i$  vanishes on the bounding surface, that is, if the system does not radiate energy to the outside.

Physical meaning for the  $T_i^0$  components can be suggested similarly by turning once more to the vibrations of the elastic rod. If the particles in the rod move by the same amount all along the rod, the motion will be that of a rigid body, that is, no oscillatory disturbances. The net change of mass in a length  $dx$  of the rod as a result of the motion would clearly be zero, since as much mass moves past  $x + dx$  as past  $x$ . There would still be a net momentum density  $\mu\dot{\eta}$  for this case of rigid-body motion. When wave motion takes place, a net mass change in the length  $dx$  exists, amounting at any given time to (cf. Fig. 13.2)

$$\mu[\eta(x) - \eta(x + dx)] = -\mu \frac{d\eta}{dx} dx. \quad (13.38)$$

The additional momentum in the interval resulting from the wave motion is therefore

$$-\mu\dot{\eta} \frac{d\eta}{dx} dx.$$

Thus, an additional momentum density, above and beyond that of the steady-state motion, can be identified as the wave or field momentum density:

$$-\mu\dot{\eta} \frac{d\eta}{dx}. \quad (13.39)$$

This quantity is just  $-T_1^0$  for the Lagrangian density given by Eq. (13.10). Thus, we are led to identify  $-T_i^0$  as the components of field momentum density and  $-R_i$ , as the total (linear) momentum of the field, at least in this four-dimensional convention.

The equations of continuity, Eqs. (13.33), then suggest that  $-T_i$  must represent the vector flux density for the  $i$ th component of the field momentum density. We ascribe a vector property to  $T_i$  because there can be, for example, a flow in the  $y$ -direction of the  $x$ -component of the momentum density, as measured by  $-T_x^y$ . An alternative interpretation of  $T_i^j$  comes from considering the displacement field

of an elastic solid. It is well known that in such a solid there are also shear forces (besides the compression forces normal to a surface) along a surface element. The entire assemblage of forces can be described by saying that the force  $d\mathbf{F}$  acting on an element of area  $d\mathbf{A}$  is expressed in terms of a *stress tensor*  $T$  such that

$$d\mathbf{F} = T \cdot d\mathbf{A}. \quad (13.40)$$

Hence, the net force, say in the  $x$ -direction, on a rectangular volume element  $dx dy dz$  has a contribution from the forces on the surfaces in  $yz$  planes given by (cf. Fig. 13.3) (where 1 indicates the  $x$  component, 2 the  $y$ , etc.)

$$[T_1^1(x + dx) - T_1^1(x)] dy dz = \frac{dT_1^1}{dx} dx dy dz, \quad (13.41)$$

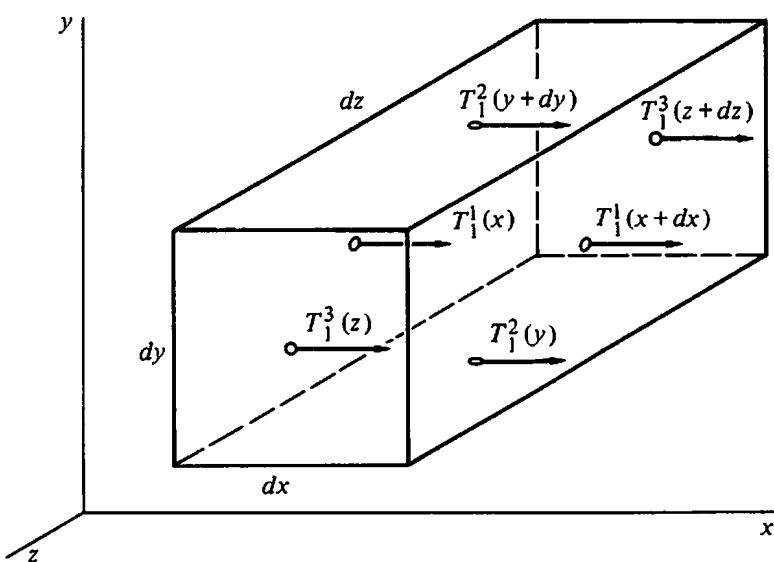
but there is also a contribution from the surfaces in the  $xz$  plane;

$$[T_1^2(y + dy) - T_1^2(y)] dx dz = \frac{dT_1^2}{dy} dx dy dz, \quad (13.42)$$

and similarly from the  $xy$  planes. Newton's equations of motion here correspond to saying that the time rate of change of the momentum density in the  $x$  direction,  $-T_1^0$ , is equal to the  $x$ -component of the force on a unit volume element:

$$-\frac{dT_1^0}{cdt} = \frac{dT_1^1}{dx} + \frac{dT_1^2}{dy} + \frac{dT_1^3}{dz}, \quad (13.43)$$

which is precisely the  $x$ -component of Eq. (13.33). For this particular field  $T_i^j$  can be identified as the elements of the three-dimensional stress tensor, hence the origin of the name "stress-energy tensor" for  $T_{\mu}^{\nu}$ .



**FIGURE 13.3** Force in  $x$  direction on a volume element  $dx dy dz$  of an elastic solid.

By considerations of a continuous mechanical system, we have thus been able to attach physical identifications, or associations, to each of the components of the stress-energy tensor. Thus, the components are

$T_0^0$	field energy density divided by $c$ ,
$T_0$ , with components $T_0^j$	field energy current density,
$-T_i^0$	field momentum density, $i$ th component,
$-T_i$ , with components $T_i^0$	current density for the $i$ th component of the field momentum density,
$T_i^j$	three-dimensional stress tensor

where, as we saw discussed following Eqs. (13.33) and (13.35),  $T_0$  and  $T_i$  form 4-space vectors each of which is conserved and thus identified, in analogy with the charge-current vector of electromagnetic theory as a “4-current”. All such conserved objects are called currents in field theory.

In almost all cases the three-dimensional tensor  $T$  is symmetric. This is not only physically desirable, but almost necessarily a characteristic for the spatial portion of the stress-energy tensor.

It must be remembered that although the example of mechanical systems gave birth to the procedures and nomenclature, the formalism can be applied to any field irrespective of its nature or origin. A classical theory of fields can be constructed not only for vibrations of an elastic solid, but also for the electromagnetic field, for the “field” of the Schrödinger wave function, or for the relativistic field describing a “scalar” meson, among others. We shall examine some of these examples in more detail later on.

Recalling the identification of  $R_i$ , the conservation equations, Eq. (13.35), say that for a closed noninteracting system the total linear momentum of the field is conserved. We would expect no less. But there should be a corresponding conservation theorem for the total angular momentum of the field. It is simple to construct a quantity that should act as an angular momentum density. Since angular momentum is an axial vector. We expect that the components of the angular momentum density are the elements of an antisymmetric tensor of the second rank. A suitable form for this tensor is

$$\mathcal{M}^{ij} = -(x^i T^{j0} - x^j T^{i0}), \quad (13.44)$$

with the total angular momentum of the field given by

$$M^{ij} = \int \mathcal{M}^{ij} dV. \quad (13.45)$$

In as much as  $t$  and  $x^i$  are completely independent variables, the time rate of change of  $M^{ij}$  is

$$\frac{dM^{ij}}{dt} = - \int \left( x^i \frac{dT^{j0}}{dt} - x^j \frac{dT^{i0}}{dt} \right) dV, \quad (13.46)$$

or, from the continuity conditions, Eqs. (13.32),

$$\frac{dM^{ij}}{dt} = - \int \left( x^i \frac{dT^{jk}}{dx^k} - x^j \frac{dT^{ik}}{dx^k} \right) dV. \quad (13.47)$$

Integration by parts converts this expression to

$$\frac{dM^{ij}}{dt} = - \int \frac{d}{dx^k} (x^i T^{jk} - x^j T^{ik}) dV + \int (T^{ij} - T^{ji}) dV. \quad (13.48)$$

The first integral on the right is in the form of a volume integral of a divergence. It is therefore equal to an integral over the bounding surface, which vanishes for a closed nonradiating system. Finally, if  $T^{ij} = T^{ji}$ , the second integral is also zero. Thus, the total angular momentum of the field is conserved if  $T$  is symmetric.

If the stress tensor is not symmetric, we can often make use of the ambiguity in defining the stress tensor to restore this symmetry. Just as for the Lagrangian, the form of the stress-energy tensor, Eq. (13.30), was chosen to satisfy divergence conditions (cf. Eq. (13.29)). Therefore  $T_{\mu}^{\nu}$  is indeterminate by any function whose 4-divergence vanishes. Usually it is possible to find such a quantity to “symmetrize” the stress-energy tensor.

### 13.4 ■ HAMILTONIAN FORMULATION

It is possible to obtain a Hamiltonian formulation for systems with a continuous set of coordinates much as was done in Chapter 8 for discrete systems. To indicate the method of approach, we return briefly to the linear chain of mass points discussed in Section 13.1. Conjugate to each field component,  $\eta_i$ , there is a canonical momentum

$$p_i = \frac{\partial L}{\partial \dot{\eta}_i} = a \frac{\partial L_i}{\partial \dot{\eta}_i}. \quad (13.49)$$

The Hamiltonian for the system is therefore

$$H \equiv p_i \dot{\eta}_i - L = a \frac{\partial L_i}{\partial \dot{\eta}_i} \dot{\eta}_i - L,$$

or

$$H = a \left( \frac{\partial L_i}{\partial \dot{\eta}_i} \dot{\eta}_i - L_i \right). \quad (13.50)$$

It will be remembered that in the limit of the continuous rod, when  $a$  goes to zero,  $L_i \rightarrow \mathcal{L}$  and the summation in Eq. (13.50) becomes an integral:

$$H = \int dx \left( \frac{\partial \mathcal{L}}{\partial \dot{\eta}} \dot{\eta} - \mathcal{L} \right). \quad (13.51)$$

The individual canonical momenta  $p_i$ , as given by Eq. (13.49), vanish in the continuous limit, but we can define a *momentum density*,  $\pi$ , that remains finite:

$$\lim_{a \rightarrow 0} \frac{p_i}{a} \equiv \pi = \frac{\partial \mathcal{L}}{\partial \dot{\eta}_i}. \quad (13.52)$$

Equation (13.51) is in the form of a space integral over a *Hamiltonian density*,  $\mathcal{H}$ , defined by

$$\mathcal{H} = \pi \dot{\eta} - \mathcal{L}. \quad (13.53)$$

While a Hamiltonian formulation can thus be introduced in a straightforward manner for classical fields, note that the procedure singles out the time variable for special treatment. It is therefore in contrast to the development we have given for the Lagrangian formulation where the independent variables of time and space were handled symmetrically. For this reason the Hamiltonian approach, at least as introduced here, lends itself less easily to incorporation in a relativistically covariant description of fields. The Hamiltonian way of looking at fields has therefore not proved as useful as the Lagrangian method, and a rather brief description should suffice here.

The obvious route for generalizing to a three-dimensional field described by field quantities  $\eta_\rho$  is to define, analogously to Eq. (13.52), the canonical momentum densities

$$\pi^\rho(x^\mu) = \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\rho}. \quad (13.54)$$

The quantities  $\eta_\rho(x^i, t)$ ,  $\pi^\rho(x^i, t)$  together define the infinite-dimensional *phase space* describing the classical field and its time development. A conservation theorem can be found for  $\pi_\rho$  that is roughly similar to that for the canonical momentum in discrete systems. If a given field quantity  $\eta_\rho$  is cyclic in the sense that  $\mathcal{L}$  does not contain  $\eta_\rho$  explicitly (as in the case of Eq. (13.10)), then the Lagrange field equation looks like an existence statement for a conserved current:

$$\frac{d}{dx^\mu} \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\mu}} = 0, \quad (13.55)$$

or

$$\frac{d\pi^\rho}{dt} + \frac{d}{dx^i} \frac{\partial \mathcal{L}}{\partial \eta_{\rho,i}} = 0. \quad (13.56)$$

It follows that if  $\eta^\rho$  is cyclic, there is an integral conserved quantity

$$\Pi^\rho = \int dV \pi^\rho(x^i, t).$$

The obvious generalization of Eq. (13.53) for a Hamiltonian density is

$$\mathcal{H}(\eta^\rho, \eta_{\rho,i}, \pi_\rho, x^i) = \pi^\rho \dot{\eta}_\rho - \mathcal{L}, \quad (13.57)$$

where it is assumed that functional dependence upon  $\dot{\eta}_\rho$  can be eliminated by inversion of the defining equations (13.49). From this definition it follows that

$$\frac{\partial \mathcal{H}}{\partial \pi^\rho} = \dot{\eta}_\rho + \pi^\lambda \frac{\partial \dot{\eta}_\lambda}{\partial \pi^\rho} - \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\lambda} \frac{\partial \dot{\eta}_\lambda}{\partial \pi^\rho} = \dot{\eta}_\rho, \quad (13.58)$$

by Eq. (13.51). The other half of the canonical field equation is more cumbersome. When expressed in terms of the canonical variables,  $\mathcal{H}$  is a function of  $\eta_\rho$  through the explicit dependence of  $\mathcal{L}$ , and through  $\dot{\eta}_\rho$ . Hence,

$$\frac{\partial \mathcal{H}}{\partial \eta_\rho} = \pi^\lambda \frac{\partial \dot{\eta}_\lambda}{\partial \eta_\rho} - \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\lambda} \frac{\partial \dot{\eta}_\lambda}{\partial \eta_\rho} - \frac{\partial \mathcal{L}}{\partial \eta_\rho} = -\frac{\partial \mathcal{L}}{\partial \eta_\rho}. \quad (13.59)$$

Using the Lagrange equations, this can be written

$$\frac{\partial \mathcal{H}}{\partial \eta_\rho} = -\frac{d}{dx^\mu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\mu}} \right) = -\dot{\pi}^\rho - \frac{d}{dx^i} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,i}} \right). \quad (13.60)$$

Because of the appearance of  $\mathcal{L}$ , we still do not have a useful form. By an exactly parallel derivation, however, we find that

$$\frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} = \pi^\lambda \frac{\partial \dot{\eta}_\lambda}{\partial \eta_{\rho,i}} - \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\lambda} \frac{\partial \dot{\eta}_\lambda}{\partial \eta_{\rho,i}} - \frac{\partial \mathcal{L}}{\partial \eta_{\rho,i}} = -\frac{\partial \mathcal{L}}{\partial \eta_{\rho,i}}. \quad (13.61)$$

Hence, we can write as the second half of the canonical equations

$$\frac{\partial \mathcal{H}}{\partial \eta_\rho} - \frac{d}{dx^i} \left( \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \right) = -\dot{\pi}^\rho. \quad (13.62)$$

Equations (13.58) and (13.62) can be put in a notation more closely approaching Hamilton's equations for a discrete system by introducing the notion of a *functional derivative* defined as

$$\frac{\delta}{\delta \psi} = \frac{\partial}{\partial \psi} - \frac{d}{dx^i} \frac{\partial}{\partial \psi_i}. \quad (13.63)$$

Since  $\mathcal{H}$  is not a function of  $\pi_{,i}^\rho$ , Eqs. (13.58) and (13.62) can be written as

$$\dot{\eta}_\rho = \frac{\delta \mathcal{H}}{\delta \pi^\rho}, \quad \dot{\pi}^\rho = -\frac{\delta \mathcal{H}}{\delta \eta_\rho}. \quad (13.64)$$

Note that in the same symbolism the Lagrange equations, Eqs. (13.23), take the form

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\rho} \right) - \frac{\delta \mathcal{L}}{\delta \eta_\rho} = 0. \quad (13.65)$$

About the only advantage of the functional derivative, however, is that of the resultant similarity with discrete system. It suppresses, on the other hand, the parallel treatment of time and space variables.

There is a way to treat classical fields that provides almost all of the Hamiltonian formulation of discrete mechanics. The main idea behind this treatment is to replace the continuous space variable or index by a denumerable discrete index. We can see how to do this by referring again to the longitudinal oscillations of the elastic rod. Let us suppose the rod is of finite length  $L = x^2 - x^1$ . The requirement that  $\eta$  vanish at the extremities is a boundary condition that could be achieved physically by placing the rod between two perfectly rigid walls. Then the amplitude of oscillation can be represented by a Fourier series:

$$\eta(x) = \sum_{n=0}^{\infty} q_n \sin \frac{2\pi n(x - x_1)}{2L}. \quad (13.66)$$

Instead of the continuous index  $x$ , we have the discrete index  $n$ . We are allowed to use this representation for all  $x$  only when  $\eta(x)$  is a well-behaved function, which most physical field quantities are.

For simplicity in illustrating how the scheme may be carried out, it will be assumed that only one real field quantity,  $\eta$ , can be expanded in a three-dimensional Fourier series of the form

$$\eta(\mathbf{r}, t) = \frac{1}{V^{1/2}} \sum_k q_k(t) e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (13.67)$$

Here  $\mathbf{k}$  is a wave vector that can take on only discrete magnitudes and directions, such that only an integral (or sometimes, half-integral) number of wavelengths fit into a given linear dimension. We say that  $\mathbf{k}$  has a discrete spectrum. The scalar index  $k$  stands for some ordering of the set of integer indices used to enumerate the discrete values of  $\mathbf{k}$ , and  $V$  is the volume of the system, appearing in a normalization factor. Because  $\eta$  is real, we must have  $q_k^* = q_{-k}$ .

The orthogonality of the exponentials over the volume can be stated as the relation

$$\frac{1}{V} \int e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} dV = \delta_{kk'}. \quad (13.68)$$

In effect, the allowed values of  $\mathbf{k}$  are those for which the condition (13.68) is satisfied (as can be seen by looking at the one-dimensional Fourier series). It follows that the coefficients of expansion,  $q_k(t)$ , are given by

$$q_k(t) = \frac{1}{V^{1/2}} \int e^{-i\mathbf{k}\cdot\mathbf{r}} \eta(\mathbf{r}, t) dV. \quad (13.69)$$

In similar fashion, the canonical momentum density can be expanded as

$$\pi(\mathbf{r}, t) = \frac{1}{V^{1/2}} \sum_k p_k(t) e^{-i\mathbf{k} \cdot \mathbf{r}}, \quad (13.70)$$

again with  $p_k^* = p_{-k}$ . Correspondingly, the expansion coefficients,  $p_k(t)$ , are to be found from

$$p_k(t) = \frac{1}{V^{1/2}} \int e^{i\mathbf{k} \cdot \mathbf{r}} \pi(\mathbf{r}, t) dV. \quad (13.71)$$

In a sense we have almost come full circle. We began this chapter with a discrete system employing a denumerable number of generalized coordinates. By then going to the limit of a continuous set of variables, we were able to treat continuous systems. Finally, we have introduced a description of the continuous system in terms of a denumerable, discrete set of coordinates that obey the same type of mechanics as the discrete system we started with. Because of the formal correspondence with the variables of discrete systems, the  $q_k$  and  $p_k$  quantities are the obvious candidates for quantization when we go from classical to quantum field theory. Indeed, the  $q_k$  correspond to what are spoken of as the “occupation numbers” for the field.

We could describe the field in terms of discrete coordinates because the finite size of the system, and the boundary conditions, permitted a discrete Fourier expansion. Equivalently, we can say that the expansion is made over a discrete spectrum of plane waves. Since the wave vector  $\mathbf{k}$  is in quantum mechanics directly proportional to the momentum of the particle associated with the plane wave, the expansions used here are often spoken of as the *momentum representation*. We

**TABLE 13.1** Comparison of Minkowski 4-dimensional spacetime and symplectic structure (after Misner, Thorne & Wheeler, *Gravitation*, San Francisco: Freeman, 1973)

Comparison item	Hamiltonian symplectic structure	Minkowski spacetime metric structure
Canonical coordinates	$q^1, q^2, p_1, p_2$	$ct, x, y, z$
Canonical structure	$\Theta = dp_1 \wedge dq^1 + dp_2 \wedge dq^2$	$ds^2 = c^2 dt \otimes dt - dx \otimes dx - dy \otimes dy - dz \otimes dz$
Nature of “metric”	antisymmetric	symmetric
Name for “metric” structure	canonically (or dynamically) conjugate coordinates	Lorentz coordinates
Field equations	$\nabla \Theta = 0$ satisfied automatically	$R_{\alpha\beta\gamma\delta} = 0$ : flat spacetime
4-dimensional manifold	phase space	spacetime
Coordinate free description	$\nabla \Theta = 0$	$Riemann = 0$

need not be restricted to plane wave expansions. A denumerable set of coordinates can be found whenever the field functions can be expanded in terms of a discrete set of orthonormal eigenfunctions.

One final comment. The Hamiltonian or symplectic structure can be expressed in tensor notation. Table 13.1 compares the metric structure of 4-dimensional Minkowski spacetime with the symplectic structure of a Hamiltonian with coordinates  $q^1, q^2, p_1$ , and  $p_2$ .

### 13.5 ■ RELATIVISTIC FIELD THEORY

We saw in Chapter 7 that there is considerable difficulty in constructing relativistically covariant Lagrangian and Hamiltonian descriptions of particle mechanics. Part of the problem can be traced to the separate roles played by space and time coordinates. For point particles, the space coordinates are mechanical variables while time is a monotonic parameter. But in classical field theory there is a natural similarity in handling space and time coordinates. They are all parameters, together defining a point in the spacetime continuum at which the field variables are to be determined. While the four-dimensional spacetime system has been used so far only for reasons of notational simplicity, the easy and natural way it fits into the formulation suggests that a relativistically covariant description is quite feasible for classical fields. Indeed, only relatively minor tinkering has to be done to the formulation already presented so that it can handle relativistic fields in a manner that is manifestly Lorentz covariant.

Three points require specific attention: (1) the nature (and metric) of the four-dimensional space used; (2) the Lorentz transformation properties of the field quantities, Lagrangian densities, and related functions; and (3) the covariant description of the limits of integration. The simple Cartesian, 4-space with coordinates  $t, x, y, z$  that we have implicitly used so far in this chapter is not convenient for exhibiting Lorentz invariance. We will use the notation and conventions adopted in Chapter 7 as well as the results of that chapter. Accordingly, the Greek letter indices will still run from 0 to 3, with  $x^0 = ct$ . Note that the Lagrange equations (13.23) are unaffected by this change. Indeed, the term

$$\frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \right)$$

remains unaltered by a scale change of any of the  $x^\nu$ , and the other term in the Lagrange equation does not involve the coordinates at all. Further, the change in space does not affect the formulation of Hamilton's principle in Eq. (13.20), since it only introduces a multiplicative constant.

All of the quantities related to the field and associated equations must now have some definite Lorentz covariant properties. The field quantities must therefore consist of 4-tensors of some given rank—scalar, 4-vector, and so on. In principle,  $\eta_\rho$  need not be restricted to any one of these categories but may stand for a set of such, for example, two scalars. The Lagrangian and Hamiltonian densities must

also be covariant. In Hamilton's principle, the volume element ( $dx^\nu$ ) of 4-space is invariant under Lorentz transformation. Since we usually think of the action  $I$  as a scalar, this means that the Lagrangian density (and therefore  $\mathcal{H}$ ) should be scalars. That is to say, they must be functions of the field quantities (possibly along with external covariant quantities) in such manner as to form scalars under Lorentz transformations. It then follows that the stress-energy tensor  $T_{\mu\nu}$ , as defined by Eq. (13.30) is automatically a 4-tensor of the second rank. The change in the 4-space however means that the components of  $T_{\mu\nu}$  may be altered in value.

In tensor notation, the stress-energy tensor,  $\mathbf{T}$ , is a linear, symmetric "functional" with slots for two vectors. It has the following properties:

1. If we insert the 4-velocity  $u$  of the observer into one of the slots and leave the other slot empty, the output is

$$\mathbf{T}(u, \dots) = \mathbf{T}(\dots, u) = - \left( \text{density of 4-momentum, } \frac{d\mathbf{p}}{dV} \right) \quad (13.72)$$

The right-hand side is the negative of the 4-momentum per unit three-dimensional volume as measured in the observer's frame at the event where  $\mathbf{T}$  is measured. In component notation,

$$T^\alpha{}_\beta u^\beta = T_\beta{}^\alpha u^\beta = - \left( \frac{dp^\alpha}{dV} \right) \quad (13.73)$$

2. If we insert the 4-velocity  $u$  of the observer into one of the slots and an arbitrary unit vector  $n$  into the other slot, the output is

$$T(u, n) = T(n, u) = - \left( n \cdot \frac{d\mathbf{p}}{dV} \right) \quad (13.74)$$

The right-hand side is the negative of the component of the 4-momentum density along the  $n$  direction. In component notation

$$T_{\alpha\beta} u^\alpha n^\beta = T_{\beta\alpha} u^\beta n^\alpha = -n_\mu \frac{dp^\mu}{dV}. \quad (13.75)$$

3. If we insert the 4-velocity of the observer into both slots, the output is

$$\mathbf{T}(u, u) = (\text{mass energy per unit volume}) \quad (13.76)$$

as measured in the frame with 4-velocity  $u$ .

In component notation,

$$T_{\alpha\beta} u^\alpha u^\beta = T_{\beta\alpha} u^\beta u^\alpha = u_\mu \frac{dp^\mu}{dV}. \quad (13.75')$$

4. If we pick a frame and insert two spacelike basis vectors  $e_i$  and  $e_k$  in that frame, the output is

$$\begin{aligned}
T_{ik} &= T_{ki} = T(e_i, e_k) = T(e_k, e_i) \\
&= i\text{-component of force acting from side } x^k - \delta \text{ to side } x^k + \delta \text{ across} \\
&\quad \text{a unit surface area perpendicular to direction } e_k \\
&= k\text{-component of force acting from side } x^i - \delta \text{ to side } x^i + \delta \text{ across} \\
&\quad \text{a unit surface area perpendicular to direction } e_i
\end{aligned} \tag{13.77}$$

For example, if we assume the Lorentz transformations apply and consider a perfect fluid moving with a 4-velocity  $u$ , which may vary in spacetime, we can describe the fluid in terms of its mass density,  $\rho$ , and an isotropic pressure,  $p$ , both in the rest frame of the fluid element. The stress-energy tensor is given by

$$T = (\rho + p)u \otimes u + pg \tag{13.78}$$

or in component form

$$T_{\alpha\beta} = (\rho + p)u_\alpha u_\beta + p g_{\alpha\beta}. \tag{13.79}$$

Insert the 4-velocity into one slot giving

$$T^\alpha{}_\beta u^\beta = [(\rho + p)u^\alpha u_\beta + p\delta^\alpha{}_\beta]u^\beta = \rho u^\alpha. \tag{13.80}$$

In the rest frame of the fluid, this becomes

$$T^0{}_\beta u^\beta = \rho c \tag{13.81}$$

and

$$T^i{}_\beta u^\beta = \frac{dp^i}{dV} = \text{momentum density} = 0, \tag{13.82}$$

where the last equality follows from the choice of the rest frame. Finally

$$T_{ik} = T(e_i, e_k) = p\delta_{ik}. \tag{13.83}$$

The Lagrangian density is of course uncertain to a multiplicative constant factor. It is customary to choose the factor such that  $T_{00}$  (or its symmetrized form) directly represents the energy density in the field. In the chosen 4-space the quantities  $R_\mu$ , Eq. (13.34), are now defined as

$$R_\mu = \int T_\mu{}^0 dV. \tag{13.84}$$

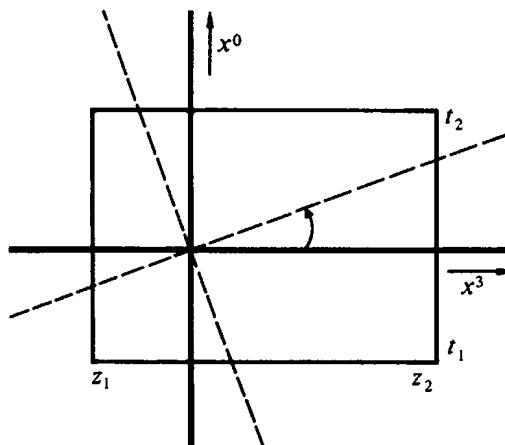
Let us consider a related set  $P^\mu$  defined as

$$P^\mu = \frac{1}{c} R^\mu. \tag{13.85}$$

It follows then, from Eqs. (13.72) to (13.76) and the interpretation given above for  $T_{i0}$ , that  $P^i$  represents the components of the total linear momentum of the field,

and  $P^0$  is  $E/c$ , where  $E$  is the total energy in the field. This suggests that we can interpret  $P^\mu$  as the 4-momentum of the field. However, we still have to show that  $R^\mu$  and  $P^\mu$  transform like 4-vectors under a Lorentz transformation. To prove this property, we shall examine what is meant by an integration over three-space in a covariant formulation and indeed how the integration limits are to be treated in general.

The first instance where the covariance of the limits of integration may be questioned is in Hamilton's principle. In Eq. (13.20), the integral appears manifestly covariant, but the limits of integration derived from Eq. (13.12) are not. The spatial integration is over some fixed volume in three-space followed by an integration over time between  $t_1$  and  $t_2$ . But an integration over  $V$  for fixed  $t$  is not a covariant concept, for simultaneity ("constant time") is not preserved under Lorentz transformation. A suitable covariant description is to say the integration is conducted over a hypersurface of three dimensions that is *spacelike*. By a spacelike surface, we mean one in which all 4-vectors lying in it are spacelike. The vectors normal to such a surface are timelike. Now, any vector connecting two points on a surface of constant time is certainly spacelike, for its  $x^0$ -component vanishes. Hence, a surface at constant time is a particular example of a spacelike surface. But such a surface retains its character in all Lorentz frames, because the spacelike or timelike quality of a vector is not affected by the Lorentz transformation. In a similar fashion, what is in one frame an integration over  $t$  at a fixed point can be described covariantly as an integration over a timelike surface. With a system of one dimension (in physical space), the integration in Hamilton's principle as given in Eq. (13.12) is over the rectangle shown in Fig. 13.4. A Lorentz transformation is a rotation in Minkowski space, and the sides of the rectangle will not be parallel to the axes in the transformed space. But we can describe the integration in all Lorentz frames as being over a region in 4-space contained between two spacelike hypersurfaces and bounded by intersecting timelike surfaces.



**FIGURE 13.4** Regions of integration in Hamilton's principle for a system extending in only one space dimension.

The appropriate covariant description of integral quantities such as  $P^\mu$  is then given as

$$P^\mu = \frac{1}{c} \int_S T^{\mu\nu} dS_\nu, \quad (13.86)$$

where the integration is over a region on a spacelike hypersurface for which the 1-form elements of surface, in the direction of the surface normal, are  $dS^\nu$  (a gradient). As  $T^{\mu\nu}$  is a 4-tensor of the second rank, it is obvious that  $P^\mu$  so defined is a 4-vector. But now we can show that the components of  $P^\mu$  given by (13.86) reduce to a volume integral in ordinary three-space, *providing* it is divergenceless, that is, satisfies Eq. (13.29). Imagine a region in 4-space defined by three surfaces:  $S_1$  and  $S_2$  that are spacelike, and  $S_3$  that is timelike (cf. Fig. 13.5). By a four-dimensional divergence theorem, a volume integral of a divergence can be replaced by a surface integral:

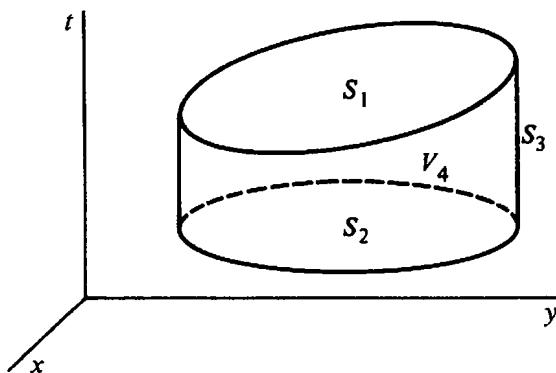
$$\int_{V_4} \frac{dT^{\mu\nu}}{dx^\nu} (dx^4) = \int_{S_1 + S_2 + S_3} T^{\mu\nu} dS_\nu, \quad (13.87)$$

where  $dx^4$  is the invariant 4-volume,  $\sqrt{|g|} c dt dx dy dz$ . The integration over  $S_3$  corresponds to an integration over  $t$  at constant  $\mathbf{r}$ . By allowing the volume to expand sufficiently, the integral over this surface will involve  $\mathbf{r}$  outside the system, where all field quantities vanish. Because of the assumed divergenceless property of  $T^{\mu\nu}$ , the integral on the left-hand side also vanishes. Therefore, if the normals to the spacelike surfaces are taken in the same sense,

$$\int_{S_1} T^{\mu\nu} dS_\nu = \int_{S_2} T^{\mu\nu} dS_\nu. \quad (13.88)$$

If  $S_1$  is any arbitrary spacelike surface, and  $S_2$  is a particular surface for which  $x^0$ , or  $t$ , is constant, then by Eq. (13.88),

$$\int_{S_1} T^{\mu\nu} dS_\nu = \int V T^{\mu 0} dV. \quad (13.89)$$



**FIGURE 13.5** Schematic integration volume in 4-space

The 4-vector transformation property of the left-hand side is obvious; hence, the right-hand side, i.e.,  $R^\mu$  according to Eq. (13.84), also transforms as a 4-vector. Further, if both  $S_1$  and  $S_2$  are surfaces at constant  $t$ , say  $t_1$  and  $t_2$ , respectively, then Eq. (13.88) is equivalent to

$$R^\mu(t_1) = R^\mu(t_2), \quad (13.90)$$

which is thus the covariant way proving that  $R^\mu$  is conserved in time.

With some care, therefore, the conserved integral quantities can still be used within the framework of a relativistic theory of classical fields. We shall not always carry through the detailed correspondence but will let it suffice in most instances that the volume integration refers to a particular Lorentz frame in which the spacelike hypersurface is a region in three-space at constant  $t$ . For the angular momentum density, note that the covariant analog of  $\mathcal{M}^{ij}$ , Eq. (13.44), is a 4-tensor of third rank:

$$\mathcal{M}^{\mu\nu\lambda} = \frac{1}{c}(x^\mu T^{\nu\lambda} - x^\nu T^{\mu\lambda}), \quad (13.91)$$

which is antisymmetric in  $\mu$  and  $\nu$ . The corresponding global or integral quantity is

$$M^{\mu\nu} = \int \mathcal{M}^{\mu\nu\lambda} dS_\lambda, \quad (13.92)$$

where the integration is over a spacelike hypersurface. If the Lorentz frame is chosen such that the surface is one at constant  $t$ , then

$$M^{\mu\nu} \rightarrow \int \mathcal{M}^{\mu\nu 0} dV, \quad (13.93)$$

which corresponds to the previous definition. The rest of the argument on the conservation of  $M^{ij}$  for symmetrical stress-energy tensors then can be carried out as before by considering this particular Lorentz frame. All of this follows from Chapter 7.

As constructed in the previous section, the Hamiltonian formulation sharply distinguishes between the time coordinate and the space coordinates. This is not to say that it is necessarily nonrelativistic, merely that the formulation is not manifestly covariant. We must imagine the Hamiltonian framework as constructed in terms of the time as seen by each particular observer. Providing the field quantities and derived functions have suitable transformation properties, this construction for each Lorentz frame is not in violation of special relativity.

One further point needs to be made here. By allowing  $\eta_\rho$  to stand for a set of covariant field quantities, we allow for the possibility that the system consists of two or more fields that interact with each other. The complete Lagrangian density may consist of a sum of Lagrangian densities representing the free fields plus terms that describe the interactions between the fields. It will be remembered that

one of the difficulties of relativistic point mechanics was the problem of considering interactions between particles that necessarily implied action-at-a-distance. However, interactions between fields can be at a point and, therefore, consistent with special relativity. We can often go further and treat the interaction between a field and a particle at a given point in spacetime. There is thus the possibility of considering relativistically a system consisting of a continuous field, a discrete particle, and the interaction between them. How this can be done in a specific case will be shown in the next section, which provides illustrations of relativistic field theories.

### 13.6 ■ EXAMPLES OF RELATIVISTIC FIELD THEORIES

We shall consider three examples, of increasing complexity.

A. *Complex scalar field.* Any complex field will be described by two independent parts, which can be expressed either as the real and imaginary part of the field or as the complex field itself and its complex conjugate. We shall follow the latter alternative. Accordingly, the Lagrangian density and associated functions will here be given in terms of two independent field variables,  $\phi$  and  $\phi^*$ , each of which are 4-scalars.\* For this particular example, we choose the Lagrangian density

$$\mathcal{L} = c^2 \phi_{,\lambda} \phi^{*,\lambda} - \mu_0^2 c^2 \phi \phi^* \quad (13.94)$$

where  $\mu_0$  is a constant and  $\phi_{,\lambda} = \frac{\partial \phi}{\partial x^\lambda}$ ,  $\phi^{,\lambda} = g^{\lambda\nu} \frac{\partial \phi}{\partial x^\nu}$  as given in Eq. (13.17). Notice, that as required,  $\mathcal{L}$  is a world scalar. Expressed in terms of space and time variables,  $\mathcal{L}$  is written as (where  $\dot{\phi} = \partial \phi / \partial t$ )

$$\mathcal{L} = \dot{\phi} \dot{\phi}^* - c^2 \nabla \phi \cdot \nabla \phi^* - \mu_0^2 c^2 \phi \phi^*. \quad (13.95)$$

To obtain the field equation for which  $\eta_\rho = \phi^*$ , note that

$$\frac{\partial \mathcal{L}}{\partial \phi^{*,\nu}} = c^2 \phi_{,\nu}, \quad \frac{\partial \mathcal{L}}{\partial \phi^*} = -\mu_0^2 c^2 \phi. \quad (13.96)$$

Hence, the Lagrange-Euler field equation is

$$\phi_{,\nu}^\nu + \mu_0^2 \phi = 0, \quad (13.97)$$

or, in equivalent form,

$$\sum_\nu \frac{d^2 \phi}{(dx^\nu)^2} + \mu_0^2 \phi = 0 \quad (13.98)$$

\*As shall be seen in the next section, complex fields lead naturally to an associated charge and current density, and this is the main reason for their introduction in physical theories.

and

$$-\nabla^2\phi + \frac{1}{c^2} \frac{d^2\phi}{dt^2} + \mu_0^2\phi = 0. \quad (13.97')$$

In terms of the D'Alembertian (cf. Section 7.5), the field equation can also be written covariantly as

$$(\square^2 + \mu_0^2)\phi = (\nabla^2 + \mu_0^2)\phi = 0. \quad (13.99)$$

Similarly, from the symmetry of  $\mathcal{L}$ , the field equation obtained when  $\eta_\rho = \phi^*$  is

$$(\square^2 + \mu_0^2)\phi^* = (\nabla^2 + \mu_0^2)\phi^* = 0. \quad (13.100)$$

This basic field equation satisfied by both  $\phi$  and  $\phi^*$  is known as the Klein–Gordon equation and, as given here, represents the relativistic analog of the Schrödinger equation for a charged zero-spin particle of rest mass energy  $\mu_0$ .

The stress-energy tensor defined by Eq. (13.30) has components

$$T_{\mu\nu} = c^2\phi_{,\mu}\phi^{*,\nu} + c^2\phi^{*,\mu}\phi_{\nu} + c^2(\phi_{,\lambda}\phi^{*,\lambda} + \mu_0^2\phi\phi^*)g_{\mu\nu} \quad (13.101)$$

and is clearly symmetrical. As the Lagrangian density describes a free field, without interactions with the outside world,  $\mathcal{L}$  does not contain  $x$  explicitly and the conservation theorem (13.29) holds for  $T_{\mu\nu}$ , as can be verified directly. To introduce the Hamiltonian formulation, we must distinguish between the time and space coordinates in some particular Lorentz frame. The conjugate momenta, according to Eq. (13.54), are then (cf. Eq. (13.95))

$$\pi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \dot{\phi}^*, \quad \pi^* = \frac{\partial\mathcal{L}}{\partial\dot{\phi}^*} = \dot{\phi}. \quad (13.102)$$

It follows that the Hamiltonian density (which has the same magnitude as  $T_{00}$ ) takes the form

$$\begin{aligned} \mathcal{H} &\equiv \pi\dot{\phi} + \pi^*\dot{\phi}^* - \mathcal{L}, \\ &= \pi\pi^* + c^2\nabla\phi \cdot \nabla\phi^* + \mu_0^2c^2\phi\phi^*. \end{aligned} \quad (13.103)$$

For the moment, all that we shall do here is illustrate the transformation to the momentum representation. The expansions (13.67) and (13.70) can be introduced into the Hamiltonian density. Since the field is not real, we do *not* have that  $q_k^* = q_{-k}$ . In effect,  $q_k$  and  $q_k^*$  now stand for two independent sets of discrete coordinates, one representing  $\phi$  and the other  $\phi^*$ . The total Hamiltonian is a sum of volume integrals over the three terms in Eq. (13.103). As a typical example, let us consider

$$\mu_0^2 \int \phi\phi^* dV = \frac{\mu_0^2}{V} \sum_{k,k'} \int q_k q_{k'}^* e^{i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{r}} dV, \quad (13.104)$$

which by Eq. (13.68) reduces to

$$\mu_0^2 q_k q_k^*.$$

The only other term requiring any special note at all is that involving the divergences, which introduce a factor  $(i\mathbf{k}) \cdot (-i\mathbf{k}')$  in the integrand. The final form for  $H$  can be written as

$$H = p_k p_k^* + \omega_k^2 q_k q_k^*, \quad (13.105)$$

where  $\omega_k$  is related to  $k$  through the dispersion relation

$$\omega_k^2 = c^2(k^2 + \mu_0^2). \quad (13.106)$$

Each term of the summation in Eq. (13.105) is in the form of a harmonic oscillator of unit mass with frequency  $\omega_k$ . This can be seen explicitly by evaluating Hamilton's equations of motion. In the momentum or plane wave representations, the fields  $\phi$  and  $\phi^*$  are thus replaced by discrete systems of harmonic oscillators, much in the same manner that the sound field in a solid is looked on as a collection of "phonons." The discrete spectrum of "vibrations" of our scalar charged field is given by Eq. (13.106). Quantization of the field (that is, the so-called second quantization) is done most simply via the momentum representation. In effect, the motion of each harmonic oscillator is quantized as would be done for an actual harmonic oscillator. But this subject certainly lies outside our province.

**B. The Sine–Gordon equation and associated field.** If the scalar field in the previous example were taken as real (that is,  $\phi^* = \phi$ ) and to exist in only one spatial dimension, then the obvious corresponding Lagrangian density along the model of Eq. (13.95) would be

$$\mathcal{L} = \frac{c^2}{2} \left[ \frac{\dot{\phi}^2}{c^2} - \left( \frac{\partial \phi}{\partial x} \right)^2 - \mu_0^2 \phi^2 \right]. \quad (13.107)$$

(The factor of  $\frac{1}{2}$  is introduced for convenience; it clearly does not affect the form of the equations of motion.) The associated field equation (cf. Eq. (13.16))

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = \mu_0^2 \phi, \quad (13.108)$$

is the one-dimensional Klein–Gordon equation. Note that it is linear in the field  $\phi(x, t)$ .

We can look upon the Lagrangian density of Eq. (13.107) as a small-field approximation to a Lagrangian density of the form

$$\mathcal{L} = \frac{c^2}{2} \left[ \frac{\dot{\phi}^2}{c^2} - \left( \frac{\partial \phi}{\partial x} \right)^2 \right] - \mu_0^2 c^2 (1 - \cos \phi), \quad (13.109)$$

which has the corresponding field equation

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = \mu_0^2 \sin \phi. \quad (13.110)$$

Inevitably, if perhaps frivolously, Eq. (13.110) has come to be known as the sine-Gordon equation. If the Klein–Gordon equation, Eq. (13.99), is reminiscent of the harmonic oscillator, then the “potential” term in the Lagrangian equation (13.109) recalls the potential term of the linear pendulum. Indeed, Eq. (13.110) has also been called, perhaps more appropriately, the pendulum equation.

In this one-dimensional world, the stress-energy tensor has only four components. As  $x$  and  $t$  again do not appear explicitly in  $\mathcal{L}$ , the elements of the tensor satisfy conservation equations, which are here two in number. Details will be left to the exercises, but of particular interest is the energy density  $T_{00}$ :

$$T_{00} = \frac{1}{2} \left[ \dot{\phi}^2 + c^2 \left( \frac{\partial \phi}{\partial x} \right)^2 \right] + \mu_0^2 c^2 (1 - \cos \phi), \quad (13.111)$$

which is of course the same in magnitude as the Hamiltonian density

$$\mathcal{H} = \frac{1}{2} \left[ \pi^2 + c^2 \left( \frac{\partial \phi}{\partial x} \right)^2 \right] + \mu_0^2 c^2 (1 - \cos \phi), \quad (13.112)$$

where the conjugate momentum is

$$\pi(x, t) = \dot{\phi}. \quad (13.113)$$

The momentum representation for the Klein–Gordon field as the sum over harmonic oscillators means that in the one-dimensional case the field can be built up as a superposition of plane waves of the form

$$q_k(t) e^{ikr} = A_0(k) e^{i(kr - \omega_k t)}, \quad (13.114)$$

where  $k$  and  $\omega_k$  are related by the *dispersion relation*, Eq. (13.106). For the field obeying the sine-Gordon equation, it is much more difficult to apply a momentum representation, because of the presence of the  $\cos \phi$  term in  $\mathcal{H}$ . But we can still solve the sine–Gordon equation by something resembling a traveling wave. A solution for  $\phi$  in Eq. (13.110) that has the form of a disturbance traveling with a speed  $v$ , but otherwise keeping its shape, must be a function only of  $\tau = t - x/v$ . In that case, Eq. (13.110) reduces to

$$\frac{d^2 \phi}{d\tau^2} - A \sin \phi = 0, \quad (13.115)$$

where

$$A = \frac{\mu_0^2 c^2 v^2}{c^2 - v^2}. \quad (13.116)$$

In terms of the variable  $\tau$ , the equation of motion is indeed that for a simple pendulum of finite amplitude. For very small amplitude, we know that  $\phi$  is a simple harmonic motion in  $\tau$  with  $\omega$  given by Eq. (13.106) for a wave number  $k = \omega/v$ , independent of the amplitude. With finite amplitude, we also know from our study of the pendulum, that while  $\phi$  will still be periodic, the frequency  $\omega$  will also depend upon the amplitude. That is to say, the dispersion relation will be amplitude dependent. This is a characteristic of course of nonlinear equations, of which the sine–Gordon equation is one example. The Klein–Gordon equation is linear, but the dispersion equation, Eq. (13.106), is said to be nonlinear; that is,  $\omega_k$  is not a linear function of  $k$ . It becomes linear only when  $\mu_0 \rightarrow 0$ , reducing the Klein–Gordon equation the usual linear wave equation.

We can thus describe the sine–Gordon equation as being nonlinear, with a nonlinear amplitude-dependent dispersion relation. Further examination reveals that it can have solutions with properties shared by only a few other nonlinear equations. These solutions are traveling wave disturbances that can interact with each other—pass through each other—and emerge with unchanged shape except perhaps for a phase shift. Such solutions are also found, for example, for the nonlinear Korteweg–deVries equation,

$$\frac{\partial \phi}{\partial t} + \alpha \phi \frac{\partial \phi}{\partial x} + \nu \frac{\partial^3 \phi}{\partial x^3} = 0, \quad (13.117)$$

where  $\alpha$  and  $\nu$  are constants. These solitary waves that preserve their shape even through interactions have been termed “solitons” and have found many applications throughout physics, from elementary particles through solid-state physics. The pendulum sine–Gordon equation, for example, has been used to describe families of elementary particles, and it also shows up in connection with the theory of the Josephson junction.

**C. The Electromagnetic Field.\*** The formalism and field equations for the electromagnetic field were developed in Section 7.5. It remains to express these ideas in terms of the Lagrangian formalism. If the components  $A^\mu$  of the electromagnetic potential are treated as the field quantities, then a suitable Lagrangian density for the electromagnetic field is

$$\mathcal{L} = -\frac{F_{\lambda\rho} F^{\lambda\rho}}{4} + j_\lambda A^\lambda. \quad (13.118)$$

To obtain the Euler–Lagrange equations, we note that

$$\frac{\partial \mathcal{L}}{\partial A^\mu} = j_\mu; \quad \frac{\partial \mathcal{L}}{\partial A_{\mu,\nu}} = -\frac{F_{\lambda\rho}}{2} \frac{\partial F^{\lambda\rho}}{\partial A_{\mu,\nu}}$$

\*Part of the difficulty in handling the electromagnetic field arises from the fact that the components  $A^\mu$  are not entirely independent; to be unique, they must be connected through some gauge condition, such as Eq. (7.66). However, it will be sufficient for our present purposes if we treat the gauge condition as a “weak” constraint.

Now, from the defining equations (7.71), the derivative of  $F_{\lambda\rho}$  vanishes except when  $\lambda = \mu$ ,  $\rho = \nu$  and  $\lambda = \nu$ ,  $\rho = \mu$ . Hence,

$$\frac{\partial \mathcal{L}}{\partial A_{\mu\nu}} = \frac{F_{\mu\nu}}{2} - \frac{F_{\nu\mu}}{2} = F_{\mu\nu}, \quad (13.119)$$

and the Euler–Lagrange equations are

$$\frac{dF^{\mu\nu}}{dx^\nu} - \sqrt{\frac{\mu_0}{\epsilon_0}} j^\mu = 0. \quad (13.120)$$

Finally, it has already been noted that  $\mathcal{L}$  for an electromagnetic field consists of a free-field Lagrangian density plus a term describing the interaction of a continuous charge and current density with the field. It is tempting to see how far we can go toward introducing field–particle interactions, by localizing the charge to a point. This is most easily done by considering the physical situation in some particular Lorentz frame, that is, as seen by a particular observer. Manifest covariance is thereby abandoned, but the result still conforms to special relativity, as it derives from a clearly relativistic theory. The current density is a measure of the motion of the charges, and in any given system  $\mathbf{j}$  is defined in terms of the charge density  $\rho$  by the relation

$$\mathbf{j}(\mathbf{r}, t) = \rho(\mathbf{r}, t)\mathbf{v}(\mathbf{r}, t).$$

Here  $\mathbf{v}$  is the velocity “field” of the continuous charge distribution. The localization can be carried out through the use of the well-known Dirac  $\delta$ -function. In three-dimensional form, the  $\delta$ -function has the property that if  $f(\mathbf{r})$  is any function of space, then

$$\int dV f(\mathbf{r})\delta(\mathbf{r} - \mathbf{s}(t)) = f(\mathbf{s}), \quad (13.121)$$

where  $\mathbf{s}(t)$  is the spatial position, say, of a particle at time  $t$  (so long as  $\mathbf{s}$  is inside the volume of integration). Thus, the spatial charge and current density corresponding to a particle of charge  $q$  at point  $\mathbf{s}$  is

$$\rho = q\delta(\mathbf{r} - \mathbf{s}) \quad (13.122)$$

and

$$\mathbf{j} = q\delta(\mathbf{r} - \mathbf{s})\mathbf{v}(\mathbf{r}). \quad (13.123)$$

If we write  $\mathcal{L}$  of Eq. (13.118) as the sum of a free-field term  $\mathcal{L}_0$  and an interaction term, the Lagrangian as seen in the given Lorentz frame is

$$L = \int dV \mathcal{L}_0 - \int dV \rho\phi + \int dV \mathbf{A} \cdot \mathbf{j} = \int dV \mathcal{L}_0 - q\phi + q\mathbf{A} \cdot \mathbf{v}. \quad (13.124)$$

The interaction terms in Eq. (13.124) are exactly the same as those in Eq. (7.141) for the Lagrangian of a single particle in an electromagnetic field. This suggests that a single Lagrangian can be formed for the complete system of particle and field that, analogous to Eq. (7.141), would look like

$$L = -mc^2\sqrt{1 - \beta^2} - q\phi + q\mathbf{v} \cdot \mathbf{A} + \int dV \mathcal{L}_0. \quad (13.125)$$

Considered as a function of the field tensor or potentials, this Lagrangian implies the field equations; considered as a function of the particle coordinates,  $L$  leads to the particle equations of motion. The mechanical descriptions of the continuous field and the discrete particle have in effect been put under one wing, expressed in a common formalism!

An important branch of modern physics is concerned with the construction of fields to represent various types of elementary particles. Of course, all such theories are quantum-mechanical, but many features of quantum field theories will have concomitant or nearly corresponding classical analogs. There is little *a priori* physical guidance in the construction of possible Lagrangian densities and interaction terms for the various particles. Some constraint on the form of these functions comes from covariance limitations. For example, the terms in  $\mathcal{L}$  must be combinations of field and other quantities in such a manner as to produce a 4-scalar. Usually,  $\mathcal{L}$  is also restricted to the field quantities or their first derivatives, although Lagrangian densities with higher derivatives have also been explored. Additional requirements on the form of the terms are also provided, or suggested, by conservation and invariance properties, implicit in the Lagrangians. These properties go beyond the conservation conditions contained in the stress-energy tensor and are usually to be found by the application of a powerful procedure known as Noether's theorem, which forms the subject of the next and last section.

### 13.7 ■ NOETHER'S THEOREM

A recurring theme throughout this text has been that symmetry properties of the Lagrangian (or Hamiltonian) imply the existence of conserved quantities. Thus, if the Lagrangian does not contain explicitly a particular coordinate of displacement, then the corresponding canonical momentum is conserved. The absence of explicit dependence on the coordinate means the Lagrangian is unaffected by a transformation that alters the value of that coordinate; it is said to be invariant, or symmetric, under the given transformation. Similarly, invariance of the Lagrangian under time displacement implies conservation of energy. The formal description of the connection between invariance or symmetry properties and conserved quantities is contained in Noether's theorem. It is in the 4-space of classical field theory that the theorem attains its most sophisticated and fertile form. For that reason, explicit discussion of the theorem has been reserved for the treatment of fields, although a discrete-system version can also be derived.

Symmetry under coordinate transformation refers to the effects of an infinitesimal transformation of the form

$$x^\mu \rightarrow x'^\mu = x^\mu + \delta x^\mu, \quad (13.126)$$

where the infinitesimal change  $\delta x^\mu$  may be a function of all the other  $x^\nu$ . Noether's theorem also considers the effect of a transformation in the field quantities themselves, which may be described by

$$\eta_\rho(x^\mu) \rightarrow \eta'_\rho(x'^\mu) = \eta_\rho(x^\mu) + \delta\eta_\rho(x^\mu). \quad (13.127)$$

Here  $\delta\eta_\rho(x^\mu)$  measures the effect of both the changes in  $x^\mu$  and in  $\eta_\rho$  and may be a function of all the other field quantities  $\eta_\lambda$ . Note that the change in one of the field variables at a particular point in  $x^\mu$  space is a different quantity  $\bar{\delta}\eta_\rho$ :

$$\eta'_\rho(x^\mu) = \eta_\rho(x^\mu) + \bar{\delta}\eta_\rho(x^\mu). \quad (13.128)$$

The description of the transformations in terms of infinitesimal changes from the untransformed quantities indicates we are dealing only with *continuous* transformations. Thus, symmetry under inversion in three dimensions (parity symmetry) is not one of the symmetries for which Noether's theorem can be applied. As a consequence of the transformations of both the coordinates and the field quantities the Lagrangian appears, in general, as a different function of both the field variables and the spacetime coordinates:

$$\mathcal{L}(\eta_\rho(x^\mu), \eta_{\rho,\nu}(x^\mu), x^\mu) \rightarrow \mathcal{L}'(\eta'_\mu(x'^\mu), \eta'_{\rho,\nu}(x'^\mu), x'^\mu). \quad (13.129)$$

The version of Noether's theorem that we shall present here is not the most general form possible, but is such as to facilitate the derivation without significantly restricting the scope of the theorem or the usefulness of the conclusions. Three conditions will be assumed to hold. The first two are

1. The 4-space is flat; that is, either it is Euclidean, or in the form of Eq. (7.171),  $R^\alpha{}_{\beta\gamma\sigma} = 0$ .
2. The Lagrangian density displays the same functional form in terms of the transformed quantities as it does of the original quantities, that is,

$$\mathcal{L}'(\eta'_\mu(x'^\mu), \eta'_{\rho,\nu}(x'^\mu), x'^\mu) = \mathcal{L}(\eta'_\mu(x'^\mu), \eta'_{\rho,\nu}(x'^\mu), x'^\mu). \quad (13.130)$$

This type of condition has not previously entered our discussions of conserved quantities, mainly because it has been automatically satisfied under the transformations considered. When cyclic coordinates are transformed by displacement, the functional dependence of the Lagrangian on the variables is unaltered by the implied shift in origin. But in our present extended types of transformation, it becomes a symmetry property that needs study. Thus, the *free-field* version of the Lagrangian density for the electromagnetic field,

Eq. (13.118), retains its functional form when  $A^\mu$  is subject to a gauge transformation, while other forms may not. Note also that Eq. (13.130) ensures that the equations of motion have the same form whether expressed in terms of the old or the new variables (form invariance). The condition of form-invariance is not the most general circumstance under which this is true; the original and transformed Lagrangian densities may also differ by a 4-divergence without modifying the equations of motion. Indeed, it is possible to carry out the derivation of Noether's theorem with such an extended version of form-invariance because the volume integral of the 4-divergence term vanishes. But for simplicity we shall restrict ourselves to Eq. (13.130). The third condition is

3. The magnitude of the action integral is invariant under the transformation, that is to say, (cf. Hamilton's principle Eq. (2.1))

$$\begin{aligned} I' &= \int_{\Omega'} (dx'^4) \mathcal{L}'(\eta'_\rho(x'^\mu), \eta'_{\rho,\nu}(x'^\mu), x'^\mu) \\ &= \int_\Omega (dx^4) \mathcal{L}(\eta_\rho(x^\mu), \eta_{\rho,\nu}(x^\mu), x^\mu), \end{aligned} \quad (13.131)$$

where  $dx^4$  is the invariant volume element is equal to  $\sqrt{|g|} dx^0 dx^1 dx^2 dx^3$  and  $\sqrt{|g|} = \sqrt{|\det(g)|}$  is the square root absolute value of the determinant of  $g$ .

Again, Eq. (13.131) represents an extension of, and includes, our previous symmetry properties such as cyclic coordinates. The Lagrangian does not change numerically under translation of a cyclic coordinate, nor does the value of the action integral. Equation (13.131) will be called the condition of *scale-invariance*. Our second and third conditions thus represent generalizations of the symmetry or invariance conditions that led to the existence of conserved quantities for discrete systems.

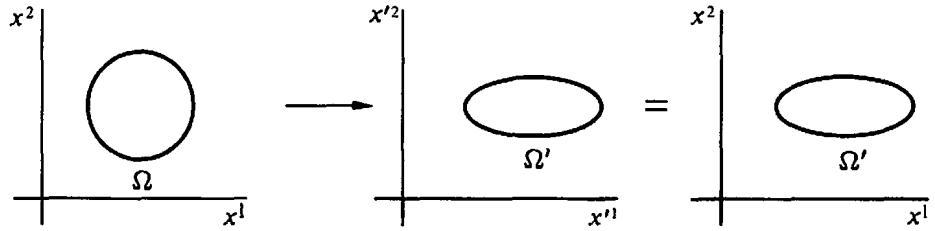
Combining Eqs. (13.130) and (13.131) gives the requirement

$$\int_{\Omega'} \mathcal{L}(\eta'_\rho(x'^\mu), \eta'_{\rho,\nu}(x'^\mu), x'^\mu) dx'^4 - \int_\Omega \mathcal{L}(\eta_\rho(x^\mu), \eta_{\rho,\nu}(x^\mu), x^\mu) dx^4 = 0. \quad (13.132)$$

In the first integral,  $x'^\mu$  now represents merely a dummy variable of integration and can therefore be relabeled  $x^\mu$ . But of course there remains a change in the domain of integration, so the condition becomes

$$\int_{\Omega'} \mathcal{L}(\eta'_\rho(x^\mu), \eta'_{\rho,\nu}(x^\mu), x^\mu) dx^4 - \int_\Omega \mathcal{L}(\eta_\rho(x^\mu), \eta_{\rho,\nu}(x^\mu), x^\mu) dx^4 = 0. \quad (13.133)$$

The sequence of transformations of space and of integration region is illustrated in Fig. 13.6 for a space of two dimensions. Equation (13.133) says that if in the action integral over  $(x^\mu)$  space we replace the original field variables by the



**FIGURE 13.6** Schematic illustration of the transformation of the invariant action integral.

transformed quantities, and transform the region of integration, then the action integral remains unaltered.

Under the infinitesimal transformations of Eqs. (13.126) and (13.127), the first-order difference between the integrals in Eq. (13.133) thus consists of two parts, one being an integral over  $\Omega$  and the other an integral over the difference volume  $\Omega' - \Omega$ . An example in one-dimension will show how the terms are to be formed. Consider the difference of two integrals:

$$\begin{aligned} \int_{a+\delta a}^{b+\delta b} (f(x) + \delta f(x)) dx - \int_a^b f(x) dx &= \int_a^b \delta f(x) dx \\ &\quad + \int_b^{b+\delta b} (f(x) + \delta f(x)) dx \\ &\quad - \int_a^{a+\delta a} (f(x) + \delta f(x)) dx. \end{aligned} \tag{13.134}$$

To first order in small quantities, the last two terms on the right can be written as

$$\int_b^{b+\delta b} f(x) dx - \int_a^{a+\delta a} f(x) dx = \delta b f(b) - \delta a f(a).$$

To this approximation, Eq. (13.134) becomes

$$\int_{a+\delta a}^{b+\delta b} (f(x) + \delta f(x)) dx - \int_a^b f(x) dx = \int_a^b \delta f(x) dx + f(x) \delta x \Big|_a^b, \tag{13.135}$$

or

$$= \int_a^b \left[ \delta f(x) + \frac{d}{dx}(\delta x f(x)) \right] dx. \tag{13.136}$$

The multidimensional analog of Eq. (13.135) then says that the invariance condition of Eq. (13.133) takes the form

$$\int_{\Omega'} \mathcal{L}(\eta', x'^{\mu}) dx'^4 - \int_{\Omega} \mathcal{L}(\eta, x^{\mu}) dx^4 = \int_{\Omega} [\mathcal{L}(\eta', x^{\mu}) - \mathcal{L}(\eta, x^{\mu})] dx^4 + \int_S \mathcal{L}(\eta) \delta x^{\mu} dS_{\mu} = 0. \quad (13.137)$$

Here,  $\mathcal{L}(\eta, x^{\mu})$  is shorthand for the full functional dependence,  $S$  is the three-dimensional surface of the region  $\Omega$  (corresponding to the end points  $a$  and  $b$  in the one-dimensional case), and  $\delta x^{\mu}$  is in effect the difference vector between points on  $S$  and corresponding points on the transformed surface  $S'$  (cf. Fig. 13.7). Corresponding to Eq. (13.136), the last integral can be transformed by the four-dimensional divergence theorem, so for the invariance condition we have

$$0 = \int_{\Omega} dx^4 \left\{ [\mathcal{L}(\eta', x^{\mu}) - \mathcal{L}(\eta, x^{\mu})] + \frac{d}{dx} (\mathcal{L}(\eta, x) \delta x^{\nu}) \right\}. \quad (13.138)$$

Now, by Eq. (13.128), the difference term in the square brackets can be written to first order as

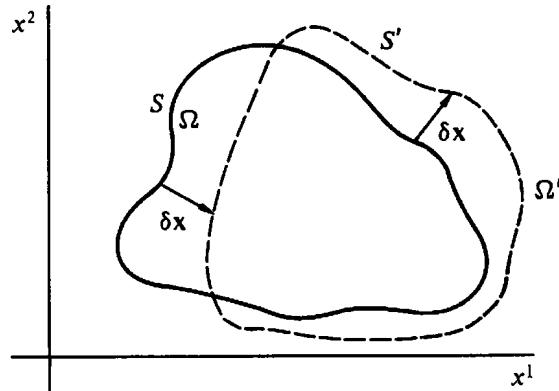
$$\mathcal{L}(\eta'_{\rho}(x^{\mu}), \eta'_{\rho,\nu}(x^{\mu}), x^{\mu}) - \mathcal{L}(\eta(x^{\mu}), \eta_{\rho,\nu}(x^{\mu}), x^{\mu}) = \frac{\partial \mathcal{L}}{\partial \eta_{\rho}} \Big| \bar{\delta} \eta_{\rho} + \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \Big| \bar{\delta} \eta_{\rho,\nu}. \quad (13.139)$$

The important property of the  $\bar{\delta}$  change is that it is a change of  $\eta$  at a fixed point in  $x^{\mu}$  space (unlike the  $\delta$  variation, Eq. (13.127)). Hence, it commutes with the spatial differentiation operator; that is, the order of the quantities

$$\bar{\delta} \quad \text{and} \quad \frac{d}{dx^{\nu}}$$

can be interchanged. Symbolically,

$$\mathcal{L}(\eta', x^{\mu}) - \mathcal{L}(\eta, x^{\mu}) = \frac{\partial \mathcal{L}}{\partial \eta_{\rho}} \bar{\delta} \eta_{\rho} + \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{d \bar{\delta} \eta_{\rho}}{dx^{\mu}}, \quad (13.140)$$



**FIGURE 13.7** The integration regions in two dimensions involved in the transformation of the action integral.

or, using the Lagrange field equations,

$$\mathcal{L}(\eta', x^\mu) - \mathcal{L}(\eta, x^\mu) = \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho, \nu}} \bar{\delta} \eta_\rho \right). \quad (13.141)$$

Hence, the invariance condition, Eq. (13.138), appears as

$$\int (dx^\mu) \frac{d}{dx^\nu} \left\{ \frac{\partial \mathcal{L}}{\partial \eta_{\rho, \nu}} \bar{\delta} \eta_\rho + \mathcal{L} \delta x^\nu \right\} = 0, \quad (13.142)$$

which is a conserved current equation (cf. arguments on pg. 571).

It is helpful however to develop the condition further by specifying the form of the infinitesimal transformation in terms of  $R$  infinitesimal parameters  $\epsilon_r$ ,  $r = 1, 2, \dots, R$ , such that the change in  $x^\nu$  and  $\eta_\rho$  is linear in the  $\epsilon_r$ :

$$\delta x^\nu = \epsilon_r X_r^\nu, \quad \delta \eta_\rho = \epsilon_r \Psi_{r\rho}. \quad (13.143)$$

The functions  $X_r^\nu$  and  $\Psi_{r\rho}$  may depend upon the other coordinates and field variables, respectively. If the transformation symmetry relates to the coordinates only, and corresponds to a displacement of a single coordinate  $x^\nu$ , then these functions are simply

$$X_r^\nu = \delta_r^\nu, \quad \Psi_{r\rho} = 0. \quad (13.144)$$

Thus, the transformations contained in the form of Eq. (13.143) constitute a far more extensive test for symmetries than we have used thus far. From Eqs. (13.127) and (13.128), it follows that to first order  $\delta \eta$  and  $\bar{\delta} \eta$  are related by

$$\delta \eta_\rho = \bar{\delta} \eta_\rho + \frac{\partial \eta_\rho}{\partial x^\sigma} \delta x^\sigma. \quad (13.145)$$

Hence,

$$\bar{\delta} \eta_\rho = \epsilon_r (\Psi_{r\rho} - \eta_{\rho, \sigma} X_r^\sigma). \quad (13.146)$$

Substituting Eqs. (13.143) and (13.146) in the invariance condition, Eq. (13.128), we have

$$\int \epsilon_r \frac{d}{dx^\nu} \left[ \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho, \nu}} \eta_{\rho, \sigma} - \mathcal{L} \delta_\sigma^\nu \right) X_r^\sigma - \frac{\partial \mathcal{L}}{\partial \eta_{\rho, \nu}} \Psi_{r\rho} \right] dx^4 = 0. \quad (13.147)$$

Since the  $\epsilon_r$  parameters are arbitrary, there exist in analogy with Eq. (13.142),  $r$  conserved currents with differential conservation theorems: (integral of divergence = 0)

$$\frac{d}{dx^\nu} \left\{ \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho, \nu}} \eta_{\rho, \sigma} - \mathcal{L} \delta_\sigma^\nu \right) X_r^\sigma - \frac{\partial \mathcal{L}}{\partial \eta_{\rho, \nu}} \Psi_{r\rho} \right\} = 0. \quad (13.148)$$

**Equations (13.148) form the main conclusion of Noether's theorem, which thus says that if the system (or the Lagrangian density) has symmetry properties such that conditions (2) and (3) above hold for transformations of the type of Eqs. (13.143), then there exist  $r$  conserved quantities.**

The conservation of the stress-energy tensor is easily recovered as a special case of Eq. (13.142). If  $\mathcal{L}$  does not contain any of the  $x^\mu$ , then it, and therefore the action integral, will be invariant under transformations such as Eq. (13.144), where  $\lambda$  takes on all the values  $\mu$ . Equation (13.148) then reduces to

$$\frac{d}{dx^\nu} \left[ \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_{\rho,\sigma} - \mathcal{L} \delta_\sigma^\nu \right) \delta_\mu^\sigma \right] = \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_{\rho,\mu} - \mathcal{L} \delta_\mu^\nu \right), \quad (13.149)$$

which is identical with Eqs. (13.29) with  $T_{\mu\nu}$  given by Eq. (13.30).

A large number of other symmetries are covered by transformations of the form of Eq. (13.142). One of the most interesting is a family of transformations of the field variables only, called *gauge transformations of the first kind*,\* such that

$$\delta x = 0, \quad \delta \eta_\rho = \epsilon c_\rho \eta_\rho \quad (\text{no summation on } \rho), \quad (13.150)$$

where the  $c_\rho$  are constants. If the Lagrangian density, and therefore the action integral, is invariant under this transformation, then there is a conservation equation of the form

$$\frac{d\Theta^\nu}{dx^\nu} = 0, \quad (13.151)$$

where

$$\Theta^\nu = c_\rho \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_\rho. \quad (13.152)$$

Equation (13.151) is in the form of an equation of continuity with  $\Theta^\nu$  in the role of a current density  $j^\nu$ . Hence, invariance under a gauge transformation of the first kind leads to the identification of a conserved current that would be appropriate for an electric charge and current density to be associated with the field.

As an illustration, let us consider the first example of Section 13.6, the complex scalar field. A transformation of the type

$$\phi' = \phi e^{i\epsilon}, \quad \phi^{*\prime} = \phi^* e^{-i\epsilon} \quad (13.153)$$

corresponds in infinitesimal form to a gauge transformation of the first type, Eq. (13.150), with

$$c = i, \quad c^* = -i.$$

\*The familiar gauge transformation of the electromagnetic field, which adds a 4-gradient  $\Lambda_{,\mu}$  to  $A_\mu$ , is part of a gauge transformation of the second kind and is not considered here.

It is obvious that the Lagrangian density of Eq. (13.94) is invariant under the transformation (13.153). Hence, there is an associated current density for the Klein–Gordon field that can be given as

$$j_\mu = iq \left( \frac{d\phi}{dx^\mu} \phi^* - \phi \frac{d\phi^*}{dx^\mu} \right), \quad (13.154)$$

which is in agreement with the conventional quantum-mechanical current density. Note that the entire derivation of the conserved charge current density depends upon the fact that the field is complex. Thus, as mentioned above, a real field does not lead to a charge or current density associated with the field. To describe fields associated with charged particles, we must use a pair of complex fields such as  $\phi$  and  $\phi^*$  for the (spin-less) Klein–Gordon particle.

Note that while Noether's theorem proves that a continuous symmetry property of the Lagrangian density leads to a conservation condition, the converse is not true. There appear to be conservation conditions that cannot correspond to any symmetry property. The most prominent examples at the moment are the fields that have soliton solutions, for example, are described by the sine–Gordon equation or the Korteweg–deVries equation.

Consider, for example, the Lagrangian density for the sine–Gordon equation, Eq. (13.107). As  $x$  and  $t$  do not appear explicitly, the Lagrangian density is invariant under translations of space and time in the manner fulfilling the conditions of Noether's theorem. In addition, there is a symmetry under a Lorentz transformation (in  $x, t$  space). No other symmetry is apparent. We would therefore expect no more than three conserved quantities from the application of Noether's theorem. Yet it has been demonstrated, by methods lying outside the Lagrangian description of fields, that there exists an infinite number of conserved quantities. That is to say, an infinite number of distinct functions  $F_i$  and  $G_i$  that are polynomials of  $\phi$ , and derivatives can be found for which

$$\frac{dF_i}{dt} + \frac{dG_i}{dx} = 0, \quad (13.155)$$

so that the volume integrals of the  $F_i$  are constant in time. It appears that the presence of such an infinite set of conserved quantities is a necessary condition in order for the field to describe solitons.

Finally, we can easily deduce the version of Noether's theorem that should apply to discrete systems. Here the four coordinates of spacetime are no longer parametric variables on equal footing—the space coordinates revert to their status as mechanical variables (or functions thereof), and only time remains to fill the role of a parameter. The action integral, instead of being a four-dimensional volume integral,

$$I = \int \mathcal{L} dx^4,$$

is a one-dimensional integral in  $t$  as in Eq. (2.1) which is Hamilton's principle:

$$I = \int L dt.$$

Instead of the continuously indexed field variables  $\eta_\rho(x^\nu)$ , we have the discrete generalized coordinates  $q_k(t)$ . It is straightforward enough to recapitulate with these translations the steps that led to Noether's theorem. We could repeat in this manner the arguments contained in Eqs. (13.126) through (13.148) as applied to discrete systems. But the effect of the conversion is sufficiently obvious and clear, that we can readily see the translation need be done directly only on the final result, Eq. (13.148).

The rules for the translation can be summarized as

$$\begin{aligned} \mathcal{L} &\rightarrow L, \\ x^\mu \text{ or } x^\nu &\rightarrow t, \\ \eta_\rho &\rightarrow q_k, \\ \eta_{\rho,\nu} &\rightarrow \dot{q}_k. \end{aligned} \tag{13.156}$$

Further, all sums over 4-valued Greek indices reduce to one term, in  $t$ . As a result, the transformations, Eq. (13.143), under which the Lagrangian is to exhibit form and scale invariance become

$$\delta t = \epsilon_r X_r, \quad \delta q_k = \epsilon_r \Psi_{rk}. \tag{13.157}$$

Equation (13.148), the statement of the conservation theorems resulting from the invariance, now becomes

$$\frac{d}{dt} \left[ \left( \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k - L \right) X_r - \frac{\partial L}{\partial \dot{q}_k} \Psi_{rk} \right] = 0. \tag{13.158}$$

**Equation (13.158) is the statement of the conclusions of Noether's theorem for a discrete mechanical system.**

The expression in the parentheses in Eq. (13.158) is our old friend the Jacobi integral  $h$  of Eq. (2.53), or equivalently in terms of  $(q, p)$ , the Hamiltonian. Indeed, we can recover the conservation of  $h$  by considering a transformation that involves a displacement of time only:

$$X_r = \delta_{r1}, \quad \Psi_{rk} = 0. \tag{13.159}$$

If the Lagrangian is not an explicit function of time, then clearly the form of the Lagrangian and the value of the action integral are unaffected by this transformation. But Noether's theorem, Eq. (13.148), then says that as a result there is a conservation theorem

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k - L \right) = 0,$$

which is identical with the familiar conclusion of Section 2.6.

Let us suppose further that a particular coordinate  $q_l$  is cyclic. Then the Lagrangian and the action are invariant under a transformation for which

$$X_r = 0, \quad \Psi_{rk} = \delta_{kl} \delta_{rl} \quad (13.160)$$

and Eq. (13.158) immediately implies the single conservation statement

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_l} \right) = 0,$$

or

$$\dot{p}_l = 0.$$

so the canonical momentum is conserved. Thus, the theorems on the conservation both of Jacobi's integral and of the generalized momentum conjugate to a cyclic coordinate are subsumed under Noether's theorem as stated in Eq. (13.158).

The connection between symmetry properties of a mechanical system and conserved quantities has run as a thread throughout formulations of mechanics as presented here. Having come full circle, as it were, and rederived by sophisticated techniques symmetry theorems found in the first chapters, it seems an appropriate point at which to end our discussions.

## EXERCISES

1. (a) The transverse vibrations of a stretched string can be approximated by a discrete system consisting of equally spaced mass points located on a weightless string. Show that if the spacing is allowed to go to zero, the Lagrangian approaches the limit

$$L = \frac{1}{2} \int \left[ \mu \dot{\eta}^2 - T \left( \frac{\partial \eta}{\partial x} \right)^2 \right] dx$$

for the continuous string, where  $T$  is the fixed tension. What is the equation of motion if the density  $\mu$  is a function of position?

- (b) Obtain the Lagrangian for the continuous string by finding the kinetic and potential energies corresponding to transverse motion. The potential energy can be obtained from the work done by the tension force in stretching the string in the course of the transverse vibration.
2. (a) Describe the field of sound vibrations in a gas in the Hamiltonian formalism and obtain the corresponding Hamilton equations of motion.
- (b) Generalizing the momentum expansion to a vector field, express the Hamiltonian for the acoustic modes of a gas in the momentum representation.

3. Obtain Hamilton's equations of motion for a continuous system from the modified Hamilton's principle, following the procedure of Section 8.5.
4. Show that if  $\psi$  and  $\psi^*$  are taken as two independent field variables, the Lagrangian density

$$\mathcal{L} = \frac{\hbar^2}{8\pi^2 m} \nabla\psi \cdot \nabla\psi^* + V\psi^*\psi + \frac{\hbar}{4\pi i} (\psi^*\dot{\psi} - \psi\dot{\psi}^*)$$

leads to the Schrödinger equation

$$-\frac{\hbar^2}{8\pi^2 m} \nabla^2\psi + V\psi = \frac{i\hbar}{2\pi} \frac{\partial\psi}{\partial t},$$

and its complex conjugate. What are the canonical momenta? Obtain the Hamiltonian density corresponding to  $\mathcal{L}$ .

5. Show that

$$G_i = - \int \pi^k \frac{\partial\eta_k}{\partial x^i} dV$$

is a constant of the motion if the Hamiltonian density is not an explicit function of position. The quantity  $G_i$  can be identified as the total linear momentum of the field along the  $x^i$  direction. The similarity of this theorem with the usual conservation theorem for linear momentum of discrete systems should be obvious.

6. (a) In a 4-space that is not Euclidean, the D'Alembertian is defined as

$$\square^2 = \nabla^2 = g^{\mu\nu} \frac{\partial^2}{\partial x^\mu \partial x^\nu}.$$

Here  $g^{\mu\nu}$  is the contravariant metric tensor, which in the flat space of special relativity is indeed the same as  $g_{\mu\nu}$ . For the metric tensor of trace +2 instead of -2 used in Eq. (7.33), find the explicit form of the D'Alembertian so defined.

- (b) A suitable Lagrangian for the charged scalar meson field in this metric is

$$\mathcal{L} = \frac{1}{2} \left( g^{\mu\nu} \frac{\partial\phi}{\partial x^\mu} \frac{\partial\phi^*}{\partial x^\nu} - \mu_0^2 \phi\phi^* \right).$$

Show that one of the corresponding field equations is

$$(\square^2 - \mu_0^2)\phi = (\nabla^2 - \mu_0^2)\phi = 0.$$

Show also that in light of part (a) this equation is actually identical with Eq. (13.99).

7. To the Lagrangian density for the scalar charged meson, Eq. (13.94), add the following term to represent the interaction with an electromagnetic field:

$$j^\lambda A_\lambda$$

where

$$j_\lambda = i(\phi\phi^*,_\lambda - \phi,_\lambda\phi^*).$$

What are the field equations for  $\phi$  and  $\phi^*$ ? What happens to the conserved currents and associated conservation theorems?

8. Suppose the Lagrangian density in Hamilton's principle is a function of higher derivatives of the field quantities  $\eta_\rho$ :

$$\mathcal{L} = \mathcal{L}(\eta_\rho; \eta_{\rho,\mu}; \eta_{\rho,\mu\nu}; x^\lambda).$$

Assuming the vanishing of the variation at the end points, what is the form of the field equations corresponding to such a Lagrangian density?

9. Consider a scalar field quantity  $\eta$  that, for simplicity, is a function only of  $x$  and  $t$ . Suppose now that the Hamiltonian density is a function of higher spatial derivatives of  $\eta$  and  $\pi$ , that is,

$$\mathcal{H} = \mathcal{H}(\eta, \eta_x, \pi, \pi_x, \pi_{xx}).$$

What are the corresponding Hamilton equations of motion?

10. Show that the Korteweg-deVries equation corresponds to the field equation for a scalar field  $\psi$  with Lagrangian density

$$\mathcal{L} = \frac{1}{2}\psi_x\psi_t + \frac{\alpha}{6}\psi_x^3 - \frac{\nu}{2}\psi_{xx}^2,$$

where the subscripts indicate derivatives with respect to the variables indicated, provided  $\psi$  is a potential function for the quantity  $\phi$  of Eq. (13.117):

$$\phi = \frac{\partial \psi}{\partial x}.$$

11. Consider a Hamiltonian density in  $(x, t)$  space:

$$\mathcal{H} = \eta^3 + \frac{1}{2}\eta^2,_x + \pi^3,_x + \frac{1}{2}\pi^2,_xx.$$

Show that the Hamilton equations of motion correspond to a form of the Korteweg-deVries equation, Eq. (13.117), if

$$\eta = \phi(x, t)$$

$$\pi = \int_{-\infty}^{\infty} \phi(x', t) dx'.$$

12. Evaluate explicitly  $T_j^0/c$  and  $T_{ij}$  for the symmetrized stress-energy tensor of the free electromagnetic field as given by

$$T_{\mu\nu}{}_{sym} = T_{\mu\nu} - \frac{A_{\mu,\lambda}F^\lambda_\nu}{4\pi} = -\frac{F_{\lambda\mu}F^\lambda_\nu}{4\pi} + \mathcal{L}g_{\mu\nu}$$

What can be said about the physical meaning of these components?

13. In a 4-space with metric  $g_{\mu\nu}$  of trace +2, evaluate explicitly the elements of the covariant (mathematically speaking) tensor  $F_{\mu\nu}$  of the electromagnetic field. Also give the elements of the matrix with one index lifted and with two indices lifted:

$$F_\rho^\lambda = g^{\lambda\mu} F_{\mu\nu}; \quad F^{\lambda\rho} = g^{\lambda\mu} F_{\mu\nu} g^{\rho\nu}.$$

# APPENDIX A

## Euler Angles in Alternate Conventions and Cayley–Klein Parameters

The Euler angles as defined in Section 4.4 are specified by an initial rotation about the original  $z$  axis through an angle  $\phi$ , a second rotation about the intermediate  $x$  axis through an angle  $\theta$ , and a third roation about the final  $z$  axis through an angle  $\psi$ . This sequence is here denoted as the “ $x$  convention,” referring to the choice of the second rotation. For the  $x$  convention the Cayley–Klein parameters in terms of the Euler angles are

$$\begin{aligned}\alpha &= e^{i(\psi+\phi)/2} \cos \frac{\theta}{2}, & \beta &= ie^{i(\psi-\phi)/2} \sin \frac{\theta}{2}, \\ \gamma &= ie^{-i(\psi-\phi)/2} \sin \frac{\theta}{2}, & \delta &= e^{-i(\psi-\phi)/2} \cos \frac{\theta}{2},\end{aligned}$$

Other conventions are possible, and two in particular have found frequent applications in particular fields. Formulas will be given here for properties of a general rotation in terms of the Euler angles of these two alternate conventions.

### $y$ CONVENTION

The  $y$  convention differs from the  $x$  convention only in that the second rotation is about the intermediate  $y$  axis. Transcription from the  $x$  to the  $y$  convention is particularly simple because  $\theta$  retains its meaning in both conventions and the changes for the other angles are easily obtained. In the  $x$  convention,  $\phi$  is the angle between the line of nodes and the  $x$  axis; in the  $y$  convention, it is the same angle measure to the  $y$  axis. Similarly in the  $x$  convention,  $\psi$  is the angle between the line of nodes and the  $x'$  axis; while in the  $y$  convention, it is the same angle relative to the  $y'$  axis. Temporarily using subscripts to indicate the convention used, these relations imply the connection (cf. Fig. 4.7)

$$\begin{aligned}\phi_x &= \phi_y + \frac{\pi}{2} \\ \psi_x &= \psi_y - \frac{\pi}{2},\end{aligned}\tag{A.1y}$$

or

$$\begin{aligned}\sin \phi_x &= \cos \phi_y & \sin \psi_x &= -\cos \psi_y \\ \cos \phi_x &= -\sin \phi_y & \cos \psi_x &= \sin \psi_y.\end{aligned}\tag{A.2y}$$

With this recipe we obtain the following formulas in terms of the Euler angles in the  $y$  convention:

*Rotation matrix.*

$$\mathbf{A} = \begin{pmatrix} -\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi & \sin \psi \cos \phi + \cos \theta \sin \phi \cos \psi & -\cos \psi \sin \theta \\ -\cos \psi \sin \phi - \cos \theta \cos \phi \sin \psi & \cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi & \sin \psi \sin \theta \\ \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \end{pmatrix} \quad (\text{A.3y})$$

The same result can be obtained by noting that the exchange of  $y$  for  $x$  corresponds to a rotation of the reference frames about the  $z$  axis through an angle of  $-\pi/2$  or  $3\pi/2$ . We can therefore translate the  $\mathbf{A}$  matrix from  $x$  convention to  $y$  convention by a similarity transformation by the orthogonal matrix  $\mathbf{G}$ :

$$\mathbf{G} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{A.4y})$$

again leading to Eq. (A.3y).

*Cayley–Klein parameters.* For this convention the Cayley–Klein parameters are

$$\begin{aligned} \alpha &= e^{i(\frac{\psi+\phi}{2})} \cos \frac{\theta}{2} & \beta &= e^{i(\frac{\psi-\phi}{2})} \sin \frac{\theta}{2} \\ \gamma &= -e^{-i(\frac{\psi-\phi}{2})} \sin \frac{\theta}{2} & \delta &= e^{-i(\frac{\psi+\phi}{2})} \cos \frac{\theta}{2}. \end{aligned} \quad (\text{A.5y})$$

*Euler parameters.* It immediately follows from the definitions of  $e_0 - e_x$  in Section 4.5 and Eq. (A.4y) that in the  $y$  convention the Euler parameters are given by

$$\begin{aligned} e_0 &= \cos \frac{\psi + \phi}{2} \cos \frac{\theta}{2} & e_2 &= \cos \frac{\psi - \phi}{2} \sin \frac{\theta}{2} \\ e_1 &= \sin \frac{\psi - \phi}{2} \sin \frac{\theta}{2} & e_3 &= \sin \frac{\psi + \phi}{2} \cos \frac{\theta}{2}. \end{aligned} \quad (\text{A.6y})$$

*Components of angular velocity.* Either by direct use of the translation equations, (A.2y), or by following through the physical meanings of the component parts of  $\boldsymbol{\omega}$ , we can obtain the following components of  $\boldsymbol{\omega}$  along the body axes in the  $y$  convention:

$$\begin{aligned} \omega_{x'} &= -\dot{\phi} \sin \theta \cos \psi + \dot{\theta} \sin \psi \\ \omega_{y'} &= \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \\ \omega_{z'} &= \dot{\psi} \cos \theta + \dot{\psi}. \end{aligned} \quad (\text{A.7y})$$

Similarly, the components of  $\omega$  along the space axes are

$$\begin{aligned}\omega_x &= -\dot{\theta} \sin \phi + \dot{\psi} \sin \theta \cos \phi \\ \omega_y &= \dot{\theta} \cos \phi + \dot{\psi} \sin \theta \sin \phi \\ \omega_z &= \dot{\psi} \cos \theta + \dot{\phi}.\end{aligned}\quad (\text{A.8y})$$

Finally, note that

$$\cos\left(\frac{\Phi}{2}\right) = e_0 = \cos \frac{\psi + \phi}{2} \cos \frac{\theta}{2} \quad (\text{A.9y})$$

which is the same as Eq. (4.63) for the  $x$  convention.

### xyz CONVENTION

In this convention each rotation is about a differently labeled axis. Obviously, various sequences of rotations are still possible. It appears that most U.S. and British aerodynamicists and pilots prefer the sequence in which the first rotation is the *yaw* angle  $\phi$  about a  $z$  axis, the second is the *pitch* angle  $\theta$  about an intermediary  $y$  axis, and the third is a *bank* or *roll* angle  $\psi$  about the final  $x$  axis (or figure axis of the vehicle). Of the three elementary rotation matrices  $\mathbf{D}$  remains the same as Eq. (4.43),  $\mathbf{C}$  appears as

$$\mathbf{C} = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix}, \quad (\text{A.10xyz})$$

and  $\mathbf{B}$  is the same as Eq. (4.44) (with  $\psi$  in place of  $\theta$ , of course). The product  $\mathbf{BCD}$  gives the following formulas:

*Rotation matrix.*

$$\mathbf{A} = \begin{pmatrix} \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \theta \\ \sin \psi \sin \theta \cos \phi - \cos \psi \sin \phi & \sin \psi \sin \theta \sin \phi + \cos \psi \cos \phi & \cos \theta \sin \psi \\ \cos \psi \sin \theta \cos \phi + \sin \psi \sin \theta & \cos \psi \sin \theta \sin \phi - \sin \psi \cos \phi & \cos \theta \cos \psi \end{pmatrix} \quad (\text{A.11xyz})$$

*Cayley–Klein parameters.* These parameters have the form

$$\begin{aligned}\alpha = \delta^* &= \left( \cos \frac{\psi}{2} \cos \frac{\theta}{2} - i \sin \frac{\psi}{2} \sin \frac{\theta}{2} \right) e^{i\phi/2} \\ \beta = -\gamma^* &= \left( \cos \frac{\psi}{2} \sin \frac{\theta}{2} + i \sin \frac{\psi}{2} \cos \frac{\theta}{2} \right) e^{-i\phi/2}.\end{aligned}\quad (\text{A.12xyz})$$

*Euler parameters.* From Section 4.5 and Eqs. (A.12xyz), it follows that the Euler parameters are

$$\begin{aligned}\cos \frac{\Phi}{2} &= e_0 = \cos \frac{\psi}{2} \cos \frac{\theta}{2} \cos \frac{\phi}{2} + \sin \frac{\psi}{2} \sin \frac{\theta}{2} \sin \frac{\phi}{2} \\ e_1 &= \sin \frac{\psi}{2} \cos \frac{\theta}{2} \cos \frac{\phi}{2} - \cos \frac{\psi}{2} \sin \frac{\theta}{2} \sin \frac{\phi}{2} \\ e_2 &= \cos \frac{\psi}{2} \sin \frac{\theta}{2} \cos \frac{\phi}{2} + \sin \frac{\psi}{2} \cos \frac{\theta}{2} \sin \frac{\phi}{2} \\ e_3 &= -\sin \frac{\psi}{2} \sin \frac{\theta}{2} \cos \frac{\phi}{2} + \cos \frac{\psi}{2} \cos \frac{\theta}{2} \sin \frac{\phi}{2}.\end{aligned}\quad (\text{A.13xyz})$$

Note that the cosine of the total angle of rotation now has a different form from either the  $x$  or the  $y$  convention.

*Components of angular velocity.* Clearly  $\omega_\psi$  lies along the body  $x$  axis,  $\omega_\phi$  along the space  $z$  axis, and  $\omega_\theta$  along the intermediate axis, and therefore in the final  $yz$  plane. The resulting components along body axes are

$$\begin{aligned}\omega_{x'} &= \dot{\psi} - \dot{\phi} \sin \theta \\ \omega_{y'} &= \dot{\theta} \cos \psi + \dot{\phi} \cos \theta \sin \psi \\ \omega_{z'} &= -\dot{\theta} \sin \psi + \dot{\phi} \cos \theta \cos \psi.\end{aligned}\quad (\text{A.14xyz})$$

Similarly, the components of  $\omega$  along the space axes are

$$\begin{aligned}\omega_x &= \dot{\psi} \cos \theta \cos \phi - \dot{\theta} \sin \phi \\ \omega_y &= \dot{\psi} \cos \theta \sin \phi + \dot{\theta} \cos \phi \\ \omega_z &= \dot{\phi} - \dot{\psi} \sin \theta.\end{aligned}\quad (\text{A.15xyz})$$

The previous editions of this work dealt with the Cayley–Klein parameters in more depth.

# APPENDIX

# B

## Groups and Algebras

As we have seen in almost every chapter of this text, invariances in the formulation of classical mechanics display themselves as symmetries in the equations of motion. This property is formally discussed in Section 13.7 as Noether's theorem. Newtonian mechanics was formulated with the explicit assumption that the laws are invariant under any Galilean transformation to another inertial frame. In the special theory of relativity, the laws are formulated to be invariant under Lorentz transformations between inertial frames. The general theory of relativity is formulated to remove the restriction of using inertial frames. These and other invariances and transformation properties that we have discussed can be understood in terms of groups of transformations. In many cases, physicists deal extensively with representations of groups, rather than the groups themselves, so we will put some stress on representations. For example, the set of  $3 \times 3$  rotation matrices with determinant +1, which appear so extensively in the text, is a representation of the special orthogonal group in three dimensions (denoted by SO(3)). Since the reader's knowledge of groups may not be extensive, we will begin with basics by defining a group and give some examples of finite groups. We shall then discuss infinite groups\* and representations.

### PROPERTIES OF GROUPS

A group is a set of objects called elements with a product operation and the following defining properties:

1. Closure—the product of two elements equals a third element in the group. If  $a$  and  $b$  are elements in the group, the product  $ab = c$  where  $c$  is also a member of the group.
2. Multiplication is associative—if  $a$ ,  $b$ , and  $c$  are group members,  $a(bc) = (ab)c$ .
3. The group contains a unit element,  $I$ , called the identity with the property that for all elements of the group,  $a = aI = Ia$ .
4. Each element  $a$  of the group has an inverse element,  $a^{-1}$  with the property  $aa^{-1} = a^{-1}a = I$ .

\*Mathematicians at this point will use a different terminology for infinite groups. We shall follow the physicist's convention of referring to both finite and infinite collections of elements as groups.

**TABLE B.1** Multiplication Table for the Four-Element Cyclic Abelian Group,  $C_4$ 

	1	-1	$i$	$-i$
1	1	-1	$i$	$-i$
-1	-1	1	$-i$	$i$
$i$	$i$	$-i$	-1	1
$-i$	$-i$	$i$	1	-1

A group is *abelian* if the multiplication operation commutes; that is, for all elements  $a$  and  $b$  of the group,  $ab = ba$ . If any of the group elements fail to commute, then the group is *nonabelian*. An example of a finite abelian group is the set of elements  $\{1, -1, i, -i\}$  where 1 is the identity, and  $i = \sqrt{-1}$ . This group has four elements, so it is said to be of order  $h = 4$ . We shall use  $h$  for the group order. This group multiplication table is shown in Table B.1.

Each group element appears once and only once in each row and in each column of the multiplication table. This group can be generated from one element,  $i$ , called the generator, with the property

$$i^2 = -1, \quad i^3 = -1, \quad i^4 = 1, \quad (\text{B.1})$$

so it is called  $C_4$ , the cyclic group of four elements. Any cyclic group,  $C_n$ , of order  $h = n$  elements has a generator element  $A$  with the property that the  $m$ th element of the group,  $A_m$ , is of the form

$$A_m = A^m, \quad (\text{B.2})$$

where

$$A^n = I. \quad (\text{B.3})$$

A *dihedral* group,  $D_n$ , is a group with  $h = 2n$  members and two generators  $A$  and  $F$  with the properties

$$A^n = I \quad \text{and} \quad F^2 = I. \quad (\text{B.4})$$

A *subgroup* is a collection of some of the elements of a larger group that by themselves form a smaller group. For example, in  $C_4$  as we can see from the multiplication table, the elements 1 and -1 form a subgroup. Two elements  $b$  and  $c$  are *conjugate* with respect to each other if for some element of the group,  $a$ ,

$$aba^{-1} = c. \quad (\text{B.5})$$

The collection of all elements "c" conjugate to  $b$  as  $a$  runs through all the elements of the group is called a *class*. All classes are disjoint subsets of the group with each element belonging to one and only one class. For abelian groups, such as the one shown in Table B.1, all elements are their own class. The identity element,  $I$ , always belongs to a class by itself. The class structure is important for nonabelian groups.

There are two groups with six elements, the cyclic group  $C_6$  and the dihedral group  $D_3$ . The elements of  $D_3$  are usually denoted by  $I, A, B, C, D$ , and  $F$ . The generator  $A$  has the property  $A^3 = I$ . It generates the element  $B$

$$AA = A^2 = B, \quad (B.6)$$

and  $A^{-1} = B$  and  $B^{-1} = A$ , since

$$AB = BA = I. \quad (B.7)$$

The element  $F$ , has the property  $F^2 = I$  and generates the remaining two elements  $C$  and  $D$  through multiplications of  $A$  and  $B$ . The elements  $C, D$  and  $F$  are their own reciprocals since  $F^2 = C^2 = D^2 = I$ ; that is,

$$C^{-1} = C, \quad D^{-1} = D, \quad \text{and} \quad F^{-1} = F. \quad (B.8)$$

This is a *nonabelian* group since, for example, the elements  $A$  and  $C$  do not commute

$$AC = F \quad CA = D. \quad (B.9)$$

The group multiplication table is shown in Table B.2.

The subgroups are

- subgroup 1  $\rightarrow I, C$
- subgroup 2  $\rightarrow I, D$
- subgroup 3  $\rightarrow I, F$
- subgroup 4  $\rightarrow I, A, B$ .

The six elements divide into three classes,

- class 1  $I$
- class 2  $A, B$
- class 3  $C, D, F$ .

**TABLE B.2** The Multiplication Table for the Dihedral Group,  $D_3$

	$I$	$A$	$B$	$C$	$D$	$F$
$I$	$I$	$A$	$B$	$C$	$D$	$F$
$A$	$A$	$B$	$I$	$F$	$C$	$D$
$B$	$B$	$I$	$A$	$D$	$F$	$C$
$C$	$C$	$D$	$F$	$I$	$A$	$B$
$D$	$D$	$F$	$C$	$B$	$I$	$A$
$F$	$F$	$C$	$D$	$A$	$B$	$I$

Note that in Table B.2 class 3 appears only in the upper-right and the lower-left quadrant of the multiplication table, while classes 1 and 2 appear only in the upper-left and lower-right. This shows the representations that are possible for  $D_3$ .

## REPRESENTATIONS OF GROUPS

A *representation* of a group is a set of matrices that satisfies the multiplication table of the group.\* By a representation we mean what is more precisely called an inequivalent *irreducible* representation,  $\Gamma_i$ , or a set of  $m \times m$  matrices that cannot be simultaneously decomposed into lower-order matrices. A theorem in group theory states that the number of irreducible representations,  $k$ , is equal to the number of classes and the sum of the squares of the dimensions,  $l_i$ , of the irreducible representations,  $\Gamma_i$  equals the group order,  $h$ . That is,

$$\sum_{i=1}^k l_i^2 = h, \quad (\text{B.10})$$

where  $h$  is the number of elements in the group,  $k$  is the number of irreducible representations, and  $l_i$  is the dimension of the  $i$ th representation. For the group  $D_3$ ,  $k = 3$  and  $h = 6$ , so Eq. (B.10) becomes

$$l_1^2 + l_2^2 + l_3^2 = 6, \quad (\text{B.11})$$

whose only solution is  $l_1 = l_2 = 1$ ,  $l_3 = 2$ . There is, as for all groups, a one-dimensional identity representation,  $\Gamma_1$  in which we map each element onto +1. Another one-dimensional representation of  $D_3$  is the set  $\Gamma_2 = \{1, -1\}$ , where the mapping is  $\{I, A, B\} \rightarrow 1$  and  $\{C, D, F\} \rightarrow -1$  as can be seen from Table B.2. The two-dimensional matrix representation,  $\Gamma_3$ , can be given in terms of the unit matrix and the Pauli matrices:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (\text{B.12})$$

with

$$\begin{aligned} I &= 1, & A &= -\frac{1}{2} (I - i\sigma_2\sqrt{3}), & B &= -\frac{1}{2} (I + i\sigma_2\sqrt{3}), \\ C &= \frac{1}{2} (\sqrt{3}\sigma_1 + \sigma_3), & D &= -\frac{1}{2} (\sqrt{3}\sigma_1 - \sigma_3), & F &= \sigma_3. \end{aligned} \quad (\text{B.13})$$

Notice how the group elements in class 3 involve only  $\sigma_1$  and  $\sigma_3$ . Thus, they are independent of the matrices  $I$  and  $\sigma_2$ , as is expected from the structure of the

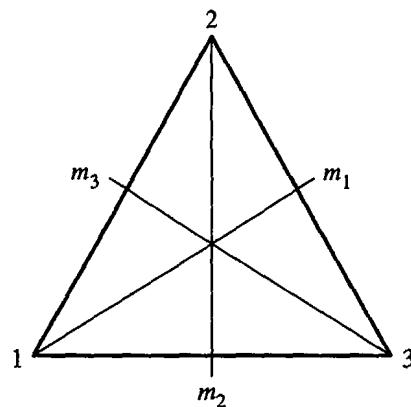
\*Mathematicians always mean matrices when they refer to representations. Some field theorists take a more general meaning.

multiplication table. However, since each representation has an identity element, there is no simple association between classes and representations.

The representation of a group can be *faithful* or *unfaithful*. For a faithful matrix representation, each element in the group is represented by a unique matrix. In an unfaithful matrix representation, more than one element in the group is represented by the same matrix. The representations  $\Gamma_1$  and  $\Gamma_2$  of  $D_3$  are unfaithful, while  $\Gamma_3$  is a faithful representation. A faithful representation is an *isomorphism* or a one-to-one mapping of the group elements onto the matrices of the representation. An unfaithful representation is a *homomorphism* or a many-to-one mapping.

We have discussed the dihedral group  $D_3$  as an abstract entity, that is, as a set of elements that satisfy a group multiplication table, and which has a two-dimensional representation that is a set of matrices also satisfying the same multiplication table. Groups also have mathematical and physical realizations in nature. For example, the permutation group of three numbers (123) is a  $D_3$  group. It has the identity (123), three twofold cycles (213), (132), and (321), which correspond with the elements  $C$ ,  $D$ , and  $F$ , and two threefold cycles, (231) and (312), which correspond to the elements  $A$  and  $B$ . A physical realization of this group is the symmetry operations of an equilateral triangle. The elements  $A$  and  $B$  are  $120^\circ$  and  $240^\circ$  rotations about a centered axis perpendicular to the plane of the triangle, and the reflection planes  $m_1$ ,  $m_2$ , and  $m_3$ , correspond to the elements  $C$ ,  $D$ , and  $F$  of the group. This is sketched in Fig. B.1. We say that the abstract group  $D_3$ , the threefold permutation group and the invariance group of operations on the equilateral triangle are isomorphisms because there is a one-to-one mapping between their elements.

As a further example, let us consider the quaternion group,  $Q$ , which is one of the five groups of order 8 (8 elements). The multiplication table is normally written as shown in Table B.3. This group has 5 classes



**FIGURE B.1** Equilateral triangle showing the three mirror planes  $m_i$ .

**TABLE B.3** The Multiplication Table for the Quaternion Group

	$I$	$-I$	$e_1$	$-e_1$	$e_2$	$-e_2$	$e_3$	$-e_3$
$I$	$I$	$-I$	$e_1$	$-e_1$	$e_2$	$-e_2$	$e_3$	$-e_3$
$-I$	$-I$	$I$	$-e_1$	$e_1$	$-e_2$	$e_2$	$-e_3$	$e_3$
$e_1$	$e_1$	$-e_1$	$-I$	$I$	$e_3$	$-e_3$	$-e_2$	$e_2$
$-e_1$	$-e_1$	$e_1$	$I$	$-I$	$-e_3$	$e_3$	$e_2$	$-e_2$
$e_2$	$e_2$	$-e_2$	$-e_3$	$e_3$	$-I$	$I$	$e_1$	$-e_1$
$-e_2$	$-e_2$	$e_2$	$e_3$	$-e_3$	$I$	$-I$	$-e_1$	$e_1$
$e_3$	$e_3$	$-e_3$	$e_2$	$-e_2$	$-e_1$	$e_1$	$-I$	$I$
$-e_3$	$-e_3$	$e_3$	$-e_2$	$e_2$	$e_1$	$-e_1$	$I$	$-I$

Class 1  $\rightarrow I$ Class 2  $\rightarrow -I$ Class 3  $\rightarrow \pm e_1$ Class 4  $\rightarrow \pm e_2$ Class 5  $\rightarrow \pm e_3$ 

From Eq. (B.10), we have

$$l_1^2 + l_2^2 + l_3^2 + l_4^2 + l_5^2 = 8,$$

which has the solution

$$l_1 = l_2 = l_3 = l_4 = 1, \quad \text{and} \quad l_5 = 2. \quad (\text{B.14})$$

For the one-dimensional representations, all elements can be mapped into +1, or they can be mapped into the one-dimensional representation  $\Gamma = \{1, -1\}$  by  $\{I, -I, e_1, -e_1\} \rightarrow +1$  and  $\{e_2, -e_2, e_3, -e_3\} \rightarrow -1$ . The two-dimensional faithful matrix representation has elements (cf. Eq. (B.12))

$$I = I, \quad -I = -I, \quad \pm e_1 = \mp i\sigma_1, \quad \pm e_2 = \mp i\sigma_2, \quad \text{and} \quad \pm e_3 = \mp i\sigma_3. \quad (\text{B.15})$$

Thus far we have confined our attention to finite groups. However, the rotations in three-space and the Lorentz transformations are infinite dimensional groups since the rotation angles and the boost velocities can take on values from the continuum. The set of all proper (determinant = +1)  $3 \times 3$  rotation matrices are a faithful representation of the special orthogonal group in three dimensions, SO(3). If we add the inversion operation, we include the improper rotations with determinant = -1 and obtain the larger orthogonal group O(3). The group SO(3) is a subgroup of the group O(3). The set of Lorentz transformation matrices in one direction constitutes a group with the O(3) a subgroup. If we allow boosts in two directions, we have a much larger group of inhomogeneous Lorentz transformations.

**TABLE B.4** The Character Table for  $D_3$ 

$D_3$	$C_1$	$2C_2$	$3C_3$
$\Gamma_1$	1	1	1
$\Gamma_2$	1	1	-1
$\Gamma_3$	2	-1	0

The sum of the diagonal elements of a matrix is called the *trace* of the matrix. The trace of the matrix in an irreducible representation,  $\Gamma_i$ , is called the *character*,  $\chi_i$ , of that matrix. The character of a matrix in a representation is determined by the class; that is, all the matrices of a representation that correspond to the same class have the same character. For the dihedral group  $D_3$ , the relation between the classes  $C_i$  of the two-dimensional representation,  $\Gamma_3$ , is given as follows:

Class $C_i$	Elements	Character $\chi_i$
Class 1	$I$	+2
Class 2	$A, B$	-1
Class 3	$C, D, F$	0

For the one-dimensional representations  $\Gamma_1$  and  $\Gamma_2$  of  $D_3$ , the characters are the same as the one-dimensional matrices. This information can be most conveniently expressed in a *character table*. For  $D_3$ , this is shown in Table B.4.

In Table B.4, the headings  $nC_m$  on the columns give the number of elements  $n$  in the class  $C_m$  of that row. The characters in the first row for class  $C_1$  also give the dimensionality of the representation. The rows of the character table are orthogonal to each other, provided we take into account the number of elements in each column. For example, considering  $\Gamma_2$  and  $\Gamma_3$ , we have  $1 \times 2 + 2 \times (1 \times -1) + 3 \times (-1 \times 0) = 0$ . As an application, in quantum mechanics the  $\Gamma_i$ 's can represent energy levels split from a parent atomic state by an electric field environment of  $D_3$  symmetry.

## LIE GROUPS AND ALGEBRAS

The terms Lie group and the associated idea of Lie algebra are used in several chapters. A Lie group is a manifold, which is also a group. A manifold is a continuous geometric object; for example, Euclidean space, the spacetime of the special theory, and a circle of radius 1 in the complex plane are all manifolds. Most of the manifolds considered in physics are continuous manifolds.\* For a manifold to be a Lie group, there must exist a group operation (termed multiplication) for

\*A continuous manifold is a manifold with the concept of nearness. That is, for every point,  $P$ , in the manifold, there exist other points in the manifold that are as close to  $P$  as desired. As the mathematicians would say, for every point,  $P$ , in the manifold and given any  $\varepsilon > 0$ , there exists another point in the manifold that is closer to  $P$  than  $\varepsilon$ , no matter how small  $\varepsilon$ .

all pairs of points in the manifold, which is consistent with the continuous nature of the manifold. Consider four points in the manifold  $a$ ,  $b$ ,  $c$ , and  $d$  and denote the group operation of  $a$  and  $c$  by  $ac$ . Consistent means, if  $a$  and  $b$  are close to each other and  $c$  and  $d$  are also close to each other, then  $ac$ ,  $ad$ ,  $bc$ , and  $bd$  are all close to each other. If we restrict our attention to the Lie groups that physicists are likely to encounter, there are only a few. One set of Lie group elements corresponds to rotations in odd dimensions, for example, the three-dimensional rotation group  $O(3)$ . A second set is the rotations in even dimensions, for example, the Lorentz group in 4 dimensions. Another set involves the unitary groups, for example,  $SU(2)$ , which is the set of  $2 \times 2$  unitary matrices with determinant +1. The final set contains the symplectic groups (See Section 9.4). There are also five special finite groups.

Corresponding to the Lie groups are Lie algebras, which are flat vector spaces with a Lie bracket or commutator defined for a set of vector fields,  $\{\tau_i\}$ , which can serve as the basis vectors of the space. These vectors satisfy

$$[\tau_i, \tau_j] = \tau_i \tau_j - \tau_j \tau_i = c_{ij}^k \tau_k \quad (\text{summation convention}) \quad (B.16)$$

where the  $c_{ij}^k$  (which clearly satisfy  $c_{ij}^k = -c_{ji}^k$ ) are called the *structure constants* of the algebra. All Lie algebras must, by symmetry, satisfy the Jacobi identity

$$J(\tau_i, \tau_j, \tau_k) = [\tau_i, [\tau_j, \tau_k]] + [\tau_j, [\tau_k, \tau_i]] + [\tau_k, [\tau_i, \tau_j]] = 0. \quad (B.17)$$

For example, the Pauli matrices satisfy Eqs. (B.16) and (B.17) with the structure constants  $c_{ij}^k = 2i\epsilon_{ijk}$ , where  $\epsilon_{ijk}$  is the Levi-Civita density symbol. They form a Lie algebra.

There is a distinction between the elements of the Lie group and the elements of the Lie algebra. The manifold of the Lie group is not conceptually identical with the flat vector space of the Lie algebra. The relation between the Lie group and the associated Lie algebra is exponential. The Lie algebra is the logarithm of the Lie group, and conversely the Lie group is the exponential of the Lie algebra in the following sense. Let  $a_m$  be a member of the Lie group, then

$$a_m = e^{(i \sum_k \theta_m^k \tau_k)}, \quad (B.18)$$

where  $\tau_k$  is a basis vector of the Lie algebra. The equal sign is interpreted as a one-to-one uniqueness. For infinite dimensional Lie groups and algebras, the sum in Eq. (B.18) is replaced by an integral and  $m$  is replaced by a continuous index. Each quantity  $\theta_m^k$  is the  $k$ th component (along the basis vector  $\tau_k$ ) of a vector  $\theta_m$  of the algebra associated with the  $m$ th element of the Lie group. The vector  $\theta$  is said to parameterize the Lie group and the Lie algebra.

An example of the group-algebra relationship is provided by the  $SU(2)$  representation of the rotation group. The algebra basis vectors are the unitary Pauli matrices Eq. (B.12) which satisfy Eq. (B.16) (cf. page 412) with the structure constants given above. For a rotation through the angle  $\theta$  about the direction of the

unit vector  $\mathbf{n}$ , we have the rotation matrix  $Q(\theta, \mathbf{n})$  where  $\mathbf{n}$  is a unit vector

$$Q = I \cos \frac{\theta}{2} + i \mathbf{n} \cdot \boldsymbol{\sigma} \sin \frac{\theta}{2}. \quad (\text{B.19})$$

This can be written in the form of Eq. (B.18)

$$Q = e^{[i(\theta/2)\mathbf{n} \cdot \boldsymbol{\sigma}]}. \quad (\text{B.20})$$

This follows from the expansion of the exponential in a power series. An expansion of the scalar product

$$\mathbf{n} \cdot \boldsymbol{\sigma} = n_x \sigma_x + n_y \sigma_y + n_z \sigma_z \quad (\text{B.21})$$

enables us to identify  $\frac{1}{2}n_k\theta$  with the parameter  $\theta_m{}^k$  of Eq. (B.18) and to identify  $n_k$  as  $\sigma_k$ . The matrices  $Q$  are a faithful representation of SU(2). The use of SU(2) was introduced into classical dynamics long before quantum mechanics was devised. It was used because SU(2) notation allows a finite rotation to be described in terms of a single angle and a single direction vector (cf. Eq. (B.21)). For a more extended discussion, see Section 4.5 of the 2nd edition of this text.

Another example of the Lie group–Lie algebra relationship is the Heisenberg algebra which, in one dimension, has the three elements  $x$ ,  $p$  and  $I$ , and the three commutators

$$\begin{aligned} [x, p] &= I = i\hbar/2\pi, \\ [x, I] &= 0 \\ [p, I] &= 0. \end{aligned} \quad (\text{B.22})$$

An associated Lie subgroup comprises the infinite set of elements  $e^{i\alpha p}$  which transform a wavefunction  $|x\rangle$  in the quantum-mechanical coordinate representation as follows:

$$e^{i\alpha p}|x\rangle = |\alpha + x\rangle, \quad (\text{B.23})$$

where  $\alpha$  is a real constant. Another Lie subgroup comprises the  $e^{i\beta x}$  operators which transform a wavefunction  $|p\rangle$  in the quantum-mechanical momentum representation in the following manner:

$$e^{i\beta x}|p\rangle = |\beta + p\rangle, \quad (\text{B.24})$$

where  $\beta$  is real. The overall Heisenberg Lie group is formed by group multiplication of the corresponding subgroup elements  $e^{i\alpha p}$  with  $e^{i\beta x}$ .

For most physical theories, there exists an action that remains unchanged in value for certain continuous changes in the dynamical variables. This is used in Chapters 1, 7, 8, 10, and 13 to derive dynamical equations of the Lagrange and Hamiltonian approaches. We can now see that the set of transformations of the dynamical variables that leave the action integral unchanged form a representation of the invariance group (often a Lie group) of that physical theory.

## CLIFFORD ALGEBRAS

The three Pauli matrices  $\sigma_k$ , their three counterparts  $i\sigma_k$ , the  $2 \times 2$  unit matrix  $\mathbf{I}$  and the matrix  $i\mathbf{I}$  together form another type of algebra called a Clifford algebra. The lowest order Clifford algebra contains the two elements  $i$  and  $1$ . A higher order Clifford algebra is formed from the  $4 \times 4$  Dirac matrices  $\gamma_i$  and their products. The  $\gamma_i$  can be expressed as direct products of Pauli matrices and the unit matrix  $\mathbf{I}$  as follows:

$$\begin{aligned}\gamma_1 &= \begin{bmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{bmatrix} & \gamma_2 &= \begin{bmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{bmatrix} & \gamma_3 &= \begin{bmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{bmatrix} \\ \gamma_4 &= \begin{bmatrix} \mathbf{I} & 0 \\ 0 & -\mathbf{I} \end{bmatrix}\end{aligned}\quad (\text{B.25})$$

In a Pauli matrix Clifford algebra formalism the scalar ( $\mathbf{A} \cdot \mathbf{B}$ ) and cross ( $\mathbf{A} \times \mathbf{B}$ ) products combine into a single operation  $\mathbf{AB}$  called a geometric product:  $\mathbf{AB} = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \times \mathbf{B}$ . The coordinate vector is written in the form  $\mathbf{r} = x\sigma_1 + y\sigma_2 + z\sigma_3$ , so the Pauli matrices act as basis vectors. A quantity  $(S, \mathbf{V}|V_p, S_p)$  defined in this algebra, called a multivector, has one scalar component  $S$ , three vector components  $V_x, V_y, V_z$ , three pseudovector components from  $\mathbf{V}_p$ , and one pseudoscalar component  $S_p$ . Several examples of multivectors and multivector transformations are:

$$\text{energy-momentum 4-vector} \quad (0, 0|\mathbf{p}, E/c) \quad (\text{B.26a})$$

$$\text{electromagnetic field tensor} \quad (0, \mathbf{E}|c\mathbf{B}, 0) \quad (\text{B.26b})$$

$$\text{space rotation} \quad (\cos \theta/2, 0|\mathbf{n} \sin \theta/2, 0) \quad (\text{B.26c})$$

$$\text{special Lorentz transformation} \quad ([\{\gamma - 1\}/2]^{1/2}, -\boldsymbol{\beta}[\{\gamma + 1\}/2]^{1/2}|0, 0) \quad (\text{B.26d})$$

$$\text{identity transformation} \quad (1, 0|0, 0) \quad (\text{B.26e})$$

The first four expressions constitute various ways of combining the nonzero parts of the four terms  $S$ ,  $V$ ,  $V_p$ , and  $S_p$  in pairs. For example, the electromagnetic fields  $\mathbf{B}$  and  $\mathbf{E}$  combine together in a multivector in which  $\mathbf{E}$  is the vector part,  $c\mathbf{B}$  is the pseudovector part, and the scalar and pseudoscalar components are zero. Note that Eq. (B.26c) reduces to (B.26e) in the limit  $\theta \Rightarrow 0$ . In this formalism the product of two successive individual rotations about different axes automatically provides the axis direction  $\mathbf{n}$  and angle  $\theta$  of the equivalent single rotation, information which cannot be readily obtained from the usual rotation matrix product operation. This convenient successive rotation technique involving the use of half angles was described in Section 4.5 of the second edition of the present text, and is omitted in the present third edition to make room for new material. The Clifford algebra approach was developed by Hestenes in his *New Foundations for Classical Mechanics* where he called it *geometric algebra* (see selected bibliography).

## GROUP THEORY CLASSIFICATION OF ELEMENTARY PARTICLES

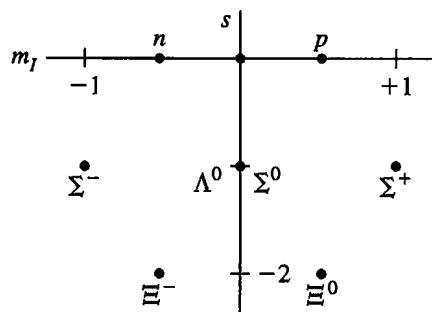
The power of group theory is demonstrated by the simple unitary group  $SU(n)$  classification schemes of elementary particles. We briefly discuss this for baryons. A small submultiplet containing  $N$  baryons is classified in terms of an  $SU(2)$  representation by its isospin number  $I$  where

$$N = 2I + 1. \quad (\text{B.27})$$

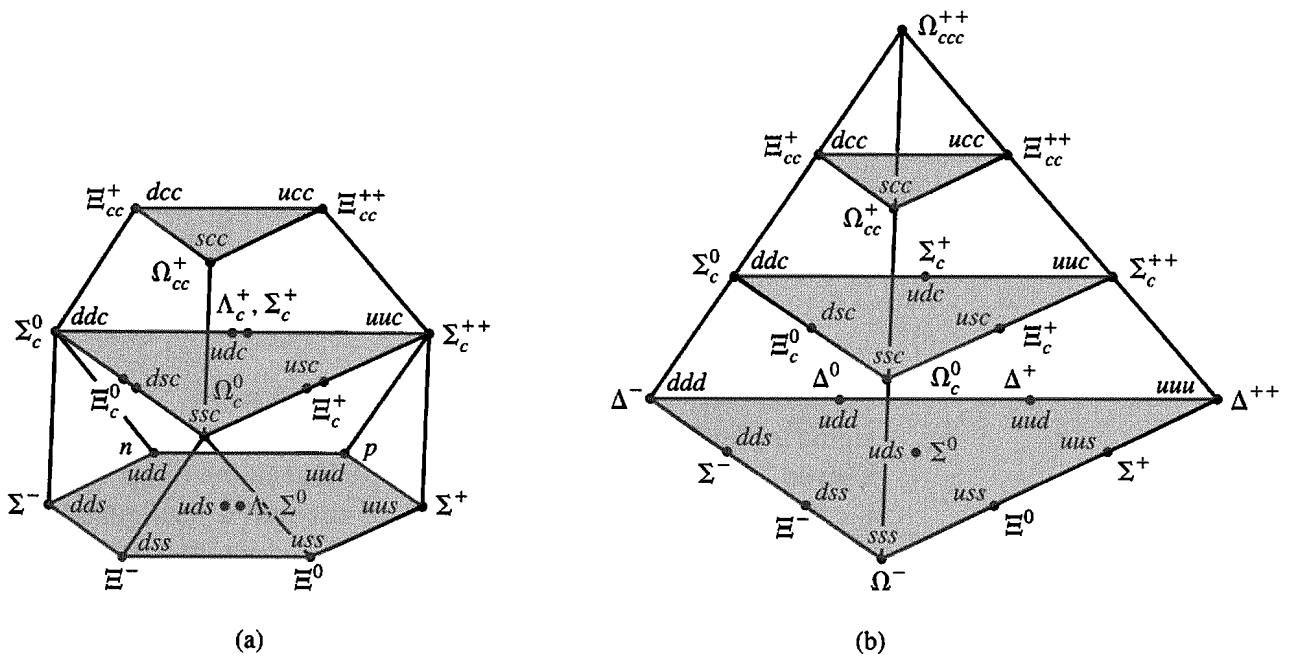
For example  $I = 1/2$  for the neutron, proton pair  $n$  and  $p$ , and  $I = 1$  for the sigma triplet  $\Sigma^-$ ,  $\Sigma^0$ , and  $\Sigma^+$ . Each particle is labeled by its  $m_I$  value, where for a given  $I$  the  $m_I$  values have integer spacings in the range  $-I \leq m_I \leq I$ . When the next higher unitary group  $SU(3)$  is invoked a new quantum number called strangeness,  $s$ , is added, and various  $SU(2)$  submultiplets with different  $s$  values group together in the larger irreducible representations  $\Gamma_i$  of  $SU(3)$ . Each baryon has three quarks called up ( $u$ ), down ( $d$ ) and strange ( $s$ ) for a total of  $3^3 = 27$  combinations (e.g., a proton has the  $uud$  grouping), and the  $SU(3)$  group theory classification divides these 27 into three irreducible representations  $\Gamma_1$ ,  $\Gamma_8$  and  $\Gamma_{10}$ , with  $\Gamma_8$  appearing twice, and the respective dimensionalities of  $\Gamma_i$  add as follows

$$\Gamma_1 + \Gamma_8 + \Gamma_8 + \Gamma_{10} = 1 + 8 + 8 + 10 = 27 \quad (\text{B.28})$$

Figure B.2 presents a plot of  $s$  versus  $m_I$  for the particles of the ground state  $SU(3)$  octet  $\Gamma_8$  which combines four  $SU(2)$  submultiplets:  $(n, p, I = 1/2)$ ,  $(\Lambda^0, I = 0)$ ,  $(\Sigma^-, \Sigma^0, \Sigma^+, I = 1)$ , and  $(\Xi^-, \Xi^0, I = 1/2)$ . A higher order classification of the baryons in terms of the special unitary group  $SU(4)$  takes into account a fourth quark  $c$  called charm, and groups together  $SU(3)$  multiplets in terms of their total charm values. Now there are four types of quarks,  $u$ ,  $d$ ,  $s$ , and  $c$ , corresponding



**FIGURE B.2** Plot of strangeness ( $s$ ) on the ordinate versus isotopic spin ( $m_I$ ) on the abscissa. The strangeness ranges from  $-2$  to  $0$  while the isotopic spin ranges from  $-1$  to  $+1$ . Horizontal lines of constant strangeness contain  $SU(2)$  submultiplets.



**FIGURE B.3** Two of the 20-fold supermultiplets of the SU(4) classification of baryons. Charm ( $c$ ) is plotted vertically and strangeness ( $s$ ) and isotopic spin ( $m_I$ ) are plotted on the horizontal plane. (a) has the uncharmed ground-state octet,  $\Gamma_8$  of Fig. B.2 at the bottom. (b) is the plot of another supermultiplet of SU(4). (See *Phys. Rev.*, D54, Part 1, 1996, p. 100.)

to  $4^3 = 64$  baryon quark combinations. Figure B.3a shows a plot of the 20-fold SU(4) supermultiplet formed by horizontal groupings of SU(3) multiplets, with each particle labeled by its quark composition. In the lowest level we find the ground state uncharmed baryons of Fig. B.2, that is baryons which contain only combinations of the quarks  $u$ ,  $d$ , and  $s$ . The middle level contains singly charmed particles, that is baryons with one  $c$  and two ordinary quarks, and the upper layer contains doubly charmed particles such as  $\Omega_{cc}^+$  with the quark content  $scc$ . Figure B.3b shows another of the SU(4) supermultiplets.

These classification schemes are of more than academic interest because they provide selection rules for predicting elementary particle interactions, such as the conservation of strangeness for strong and electromagnetic interactions, but not for weak interactions. Mesons, each of which contains a quark plus an antiquark, also conform to classification schemes by the simple unitary groups SU( $n$ ).

# Selected Bibliography

## TEXTBOOKS ON CLASSICAL MECHANICS

- V. I. Arnold, *Mathematical Methods of Classical Mechanics*, (Berlin: Springer-Verlag, 1989). A very advanced treatment of the subject.
- A. Arya, *Introduction to Classical Mechanics*, (Upper Saddle River, NJ: Prentice Hall, 1998). Undergraduate text.
- V. Barger and M. Olsson, *Classical Mechanics: A Modern Perspective*, (New York: McGraw Hill, 1995). Undergraduate text that contains some discussion of chaotic dynamics and a unique section on Newtonian Cosmology.
- H. C. Corben and P. Stehle, *Classical Mechanics*, (New York: Dover, 1994). Reprinted as a classic.
- A. L. Fetter and J. D. Walecka, *Theoretical Mechanics of Particles and Continua*, (New York: McGraw Hill, 1980). Has extensive discussion of continuous systems.
- G. Fowles and G. Cassiday, *Analytical Dynamics*, (Ft. Worth, TX: Saunders, 1999). Undergraduate text with many computer problems.
- L. Hand and J. Finch, *Analytical Mechanics*, (Cambridge, England: Cambridge University Press, 1998). Covers standard topics, including Chaos, at a level similar to that of the present text.
- D. Hestenes, *New Foundations for Classical Mechanics*, (Dordrecht, The Netherlands: Kluwer, 1999). An unconventional approach to classical mechanics written in the mathematical language of geometric algebra. It has many keen insights on the subject.
- J. V. Jose and E. J. Saletan, *Classical Dynamics, A Contemporary Approach*, (Cambridge, England: Cambridge University Press, 1998). A good overall coverage of the subject. Much of the theory is expressed in terms of configuration manifolds and tangent bundles.
- L. D. Landau and E. M. Lifshitz, *Mechanics*, Volume I of *Course in Theoretical Physics*, (Oxford, England: Pergamon, 1976). A very economical and pedagogic approach to Mechanics. Contains many popular partially worked out examples.
- J. B. Marion and S. T. Thornton, *Classical Dynamics of Particles and Systems*, (Ft. Worth, TX: Saunders, 1995). A standard undergraduate text recently updated with some chaotic dynamics in the context of nonlinear oscillations.

F. Scheck, *Mechanics: From Newton's Laws to Deterministic Chaos*, (Berlin: Springer-Verlag, 1990). This recent textbook on classical mechanics includes a chapter on the geometric aspects of mechanics which develops the theory in the language of manifolds. There is also a chapter on chaos.

## LAGRANGIAN FORMULATION

### Chapters 1 to 3

- D. Hestenes, *New Foundations for Classical Mechanics*, op cit. It has many good examples from astronomy. The three-body problem, together with the Euler and Lagrange solutions, are explained very well.
- L. D. Landau and E. M. Lifshitz, *Mechanics*, Volume I of *Course in Theoretical Physics*, op cit. A very pedagogic approach to mechanics.
- K. R. Symon, *Mechanics*, (Reading, MA: Addison Wesley, 1971). Discusses the restricted three-body problem.

## RIGID BODIES

### Chapters 4 and 5

- T. L. Chow, *Classical Mechanics*, (New York: Wiley, 1995). Provides an excellent treatment of the spinning top.
- D. Hestenes, *New Foundations for Classical Mechanics*, op cit. Provides new insights into rigid body motion, including the symmetric top.
- L. D. Landau and E. M. Lifshitz, *Mechanics*, op cit. A good, pedagogic discussion of rigid bodies.

## SMALL OSCILLATIONS

### Chapter 6

- L. D. Landau and E. M. Lifshitz, *Mechanics*, op cit. Treatment includes damped, unharmonic, and nonlinear oscillations.
- C. P. Poole, Jr., H. A. Farach and R. J. Creswick, *Superconductivity* (Boston, Academic Press, 1995). Chapter 13 discusses Josephson junctions and their mechanical analogues.

**RELATIVITY****Chapter 7**

- C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, (San Francisco: Freeman, 1973). A complete introduction to the geometric notation as applied to both the special theory and the general theory of relativity.
- B. F. Schutz, *A First Course in General Relativity*, (Cambridge, England: Cambridge University Press, 1985). The first four chapters introduce the reader to the formalism of the special theory of relativity in a form that can carry over to the general theory.
- E. F. Taylor and J. A. Wheeler, *Spacetime Physics*, (San Francisco: Freeman, 1992). An excellent discussion emphasizing the physical thoughts behind and the physical processes of relativity.

**HAMILTONIAN FORMULATION****Chapters 8 to 10**

- I. Percival and D. Richards, *Introduction to Dynamics*, (Cambridge, England: Cambridge University Press, NY, 1982). Good treatment of canonical transformations, Hamilton–Jacobi equation, and action-angle variables.

**CHAOS****Chapter 11**

- H. Bai-Lin, *Chaos*, (Singapore: World Scientific, 1984). The first part of the book consists of ten good introductory chapters that explain chaos. Chapter 2 develops the theory of chaos from Hamilton's equations and Chapter 3 discusses the logistic equation. The second and main part of the book is a collection of 41 reprinted papers, many of them landmark articles in the development of the subject.
- R. H. Enns and G. McGuire, *Nonlinear Physics with Maple for Scientists and Engineers*, (Boston: Birkhäuser, 2000). A good selection of problems with a disk of Maple programs.
- M. Hénon, *Numerical Exploration of Hamiltonian Systems*, Course 2 in *Chaotic Behavior of Deterministic Systems, Les Houches École D'Été de Physique Théoretique*, 1981, ed. G. Iooss, R. H. G. Hellerman and R. Stora, (New York: North Holland, 1983). This 114-page lecture is very readable. It provides one of the best explanations of the Hénon–Heiles Hamiltonian, and covers several other topics included in the present chapter.
- E. A. Jackson, *Perspectives of Nonlinear Dynamics*, (Cambridge, England: Cambridge University Press, 1990). This two-volume set provides a readable pre-

sentation of a variety of complementary approaches to topics in nonlinear dynamics. There are useful discussions of several topics covered in this chapter such as the Hénon–Heiles Hamiltonian, the logistic equation, Liapunov exponents and Poincaré maps. An explanation is given of the KAM theorem.

- S. A. Kauffman, *The Origins of Order*, (New York: Oxford University Press, 1993). An introduction to the concepts of complexity.
- W. Kinzel and G. Reents, *Physics by Computer*, (New York: Springer, 1998). Numerical solutions of linear and nonlinear problems using *Mathematica* and C.
- H. O. Peitgen, H. Jürgens and D. Saupe, *Chaos and Fractals, New Frontiers of Science*, (Berlin: Springer-Verlag, 1992). This volume constitutes one of the best available sources for information on fractals, and it does a good job of explaining the chaotic behavior of the logistic equation. The text is very long and wordy, but it contains many beautiful figures of fractals and trajectories of attractors.
- L. E. Reichl, *The Transition to Chaos*, (Berlin: Springer-Verlag, 1992). The theory is developed from the viewpoint of classical mechanics, using, for example, action-angle variables. There are good discussions of Poincaré sections, Liapunov exponents, the Hénon–Heiles Hamiltonian, and Kolmogorov’s approach for proving the KAM theorem.

## PERTURBATION THEORY

### Chapter 12

- I. Percival and D. Richards, *Introduction to Dynamics*, op cit. Good treatment of perturbation theory.

## CONTINUOUS SYSTEMS AND FIELDS

### Chapter 13

- B. S. Dewitt, Dynamical theory of groups and fields, in B. S. Dewitt and C. DeWitt (eds.) *Relativity, Groups and Topology*, (New York: Gordon and Breach, 1964). An excellent formalistic approach to the use of groups in modern field theory.
- A. L. Fetter and J. D. Walecka, *Theoretical Mechanics of Particles and Continua*, op cit. Has an extensive discussion of continuous systems.
- C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, op cit. This contains an excellent discussion of the concepts of field theory in a Riemannian spacetime. The material is presented in multiple level tracks.

**APPENDIX B**

- H. Goldstein, *Classical Mechanics* (Reading, MA: Addison Wesley, 1st ed., 1950, 2nd Ed., 1980). The first and second editions of the present work have thorough explanations of the correspondence between the  $2 \times 2$  complex unitary matrices of  $SU(2)$  and the  $3 \times 3$  real orthogonal matrices of  $O(3)$ , including the Cayley-Klein parameters. These discussions involve applications of Clifford algebras to classical mechanics.
- M. Hamermesh, *Group Theory*, (Reading, MA: Addison Wesley, 1962). A good coverage of group theory and its applications to physical problems. Lie groups and algebras are discussed in Chapter 8.
- D. Hestenes, *New Foundations for Classical Mechanics*, op. cit, A classical mechanics textbook with extensive sections written in the formalism of Clifford algebra, illustrating the insights to be gained by this approach.
- C. P. Poole, Jr. and H. A. Farach, Pauli-Dirac Matrix Generators of Clifford Algebras, *Found. Phys.*, **12**, 719–738 (1982). This article provides background on the role that Clifford algebras can play in classical mechanics.
- M. Tinkham, *Group Theory and Quantum Mechanics*, (New York, McGraw Hill, 1964). A well written introduction to groups. Chapter 5 covers the rotation group and angular momentum.

# Author Index

- Abel, N. H., 608. *See also* Abelian in subject index  
Arnold, V. I., 484, 487, 489, 617  
Arya, A., 617  
Atwood, G., 27, 28
- Bai-Lin, H., 489, 619  
Baker, G. L., 507, 508, 519  
Barger, V., 617  
Barone, A., 269  
Bernoulli, J., 43  
Bertrand, J., 89  
Bessel, F. W., 127  
Binet, A., 203  
Bohlin, 549  
Bohr, N. H. D., 95, 466  
Boltzmann, L. E., 85, 128  
Boyle, R., 128  
Bryan, 154
- Çambel, A. B., 519  
Cantor, G. F. L. P., 516, 517, 519, 522  
Cassiday, G., 617  
Carathéodory, C., 394  
Cayley, A., 154, 182, 603, 621  
Chandler, S. C., 208, 228  
Chasles, M., 161, 184, 228  
Chow, T. L., 618  
Clausius, R. J. E., 84, 128  
Clifford, W. K., 614, 621  
Corben, H. C., 617  
Coriolis, G. G., 174, 326  
Coulomb, C. A. de, 111  
Cramer, G., 149, 263  
Creswick, R. J., 265, 517, 618
- D'Alembert, J., 16, 18, 296, 313, 548. *See also* D'Alembertian in subject index  
Delaunay, C. E., 477  
Descartes, R., 26. *See also* Cartesian in subject index  
deVries, C., 587, 596  
Dewitt, B. S., 620  
Dirac, P. A. M., 588, 621  
Duffing, G., 523, 524
- Einstein, A., 139, 276, 326, 327, 332, 538  
Enns, R. H., 619  
Euclid, 278, 517. *See also* Euclidean in subject index  
Euler, L., 45, 122, 150, 155, 165, 196, 197, 200, 209, 234, 319, 564, 617
- Farach, H. A., 265, 517, 618, 621  
Faraday, M., 297, 298  
Feigenbaum, M. J., 506–515  
Fermat, P. de, 360  
Fetter, A. L., 617, 620  
Finch, J., 617  
Foucault, J. B. L., 179  
Fourier, Baron J. B. J., 14, 126, 259, 274, 460, 545, 551, 575  
Fowles, G., 617
- Galilei, G., 2. *See also* Galilean in subject index  
Gibbs, J. W., 337
- Goldschmidt, 64  
Goldstein, H., 621  
Gollub, J. P., 507, 508, 519  
Gordon, W., 584, 585, 596  
Gram, 249
- Hamermesh, M., 621  
Hamilton, Sir W. R., 34, 44, 45, 313, 324, 334, 430, 479, 488, 562. *See also* Hamiltonian in subject index  
Hand, L., 617  
Hausdorff, F., 517  
Helmholtz, H. L. F. von, 337  
Hénon, M., 484, 492, 496, 497, 619, 621  
Heiles, C., 484, 492, 496, 497, 619, 621  
Helleman, R. H. G., 497, 619  
Hertz, H. R., 361  
Hestenes, D., 121, 617, 618, 621  
Hooke, R., 52, 91, 317  
Huygens, C., 132
- Iooss, G., 497, 619
- Jackson, E. A., 489, 619  
Jacobi, K. G. J., 61, 334, 338, 361, 390, 398, 426, 430, 479, 488, 566, 597
- Jose, J. V., 617  
Josephson, B. D., 265, 271, 618
- Jürgens, H., 620

- Kauffman, S. A., 620  
 Kelvin, Baron W. T., 85, 338  
 Kinzel, I. W., 620  
 Klein, F., 154, 182, 218, 228, 584,  
     585, 596, 603, 621  
 Kepler, J., 73, 92, 101, 347, 370,  
     414, 445, 466, 470, 484,  
     495, 537  
 Kirchhoff, G. R., 66  
 Kolmogorov, A. N., 484, 487, 489,  
     621  
 Korteweg, D. J., 587, 596, 600  
 Kronecker, L., 138 , 190
- Lagrange, J. L., 14, 123, 198, 200,  
     319, 392, 533, 561, 564,  
     618. *See also* Lagrangian  
     in subject index  
 Landau, L. D., 617, 618  
 Laplace, P. S., marquis de 102,  
     264, 413  
 Larmor, Sir J., 231, 318  
 Legendre, A. M., 224, 334, 539  
 Lenz, H. F. E., 102–104, 413  
 Levi-Civita, T., 169, 410  
 Liapunov, M. A., 484, 492, 512,  
     519, 621  
 Lie, M. S., 171, 385, 392, 411,  
     613, 621  
 Lifshitz, E. M., 617, 618  
 Liouville, J., 418, 483  
 Lissajous J. A., 83, 258, 439, 358,  
     462, 464  
 Lorentz, H. A., 22, 131, 280, 319,  
     328, 415, 580, 612  
 Lorenz, E. N., 523, 576
- MacCullagh, 225  
 Mach, E., 324  
 Marion, J. B., 617  
 Maxwell, J. C., 54
- McGuire, G., 619  
 Minkowski, H., 278, 287, 290,  
     319, 576, 580, 601  
 Misner, C. W., 619, 620  
 Moser, J., 484, 487, 489
- Napier, J. N., 476  
 Newton, Sir I., 1, 5, 101, 132, 199,  
     299, 526. *See also*  
     Newtonian in subject index  
 Nielsen, A. C., 30  
 Noether, A. E., 344, 566, 589, 597
- Olsson, M., 617
- Paterno, G., 269  
 Pauli, W., 610, 614, 621  
 Peitgen, H. O., 512, 513, 515, 620  
 Percival, I., '619, 620  
 Planck, M. K. E. L., 380  
 Poincaré, J. H., 282, 394, 494,  
     524, 545, 621  
 Poinsot, L., 201, 202, 206, 234  
 Poisson, S. D., 225, 388, 398, 411,  
     532. *See also* subject index  
 Poole, C. P., Jr., 265, 517, 618, 621  
 Ptolemy, C., 129
- Raman, Sir C. V., 258  
 Ray, J., 47  
 Rayleigh, Baron R. J. S., 23  
 Reents, G., 620  
 Reichl, L. E., 489, 620  
 Ricci-Curbastro, G , 327  
 Richards, D., 619, 620  
 Riemann, G. F. B., 326, 327, 566,  
     620
- Rössler, O. E., 523  
 Routh, E. J., 56, 347  
 Runge, C., 102–104, 413  
 Rutherford, Baron E., 110, 113
- Saletan, E. J., 617  
 Sander, L. N., 524  
 Saupe, D., 620  
 Scheck, F., 618  
 Schrödinger, E., 54, 571, 584, 599  
 Schutz, B. F., 619  
 Schmidt, 249  
 Schwarzschild, K., 538  
 Shaw, R., 512  
 Sierpinski, W., 517, 519, 522  
 Sommerfeld, A. J. W., 218, 228  
 Staeckel, 447  
 Stehle, P., 617  
 Stokes, Sir G. G., 20, 24, 52  
 Stora, R., 497  
 Symon, K. R., 618
- Tait, P. G., 154  
 Taylor, B., 239  
 Taylor, E. F., 619  
 Thomas, L. H., 282, 330  
 Thomson, Sir W., *see* Baron  
     (Lord) Kelvin  
 Thorne, K. S. , 619, 620  
 Thornton, S. T., 617  
 Tinkham, M., 621
- van der Pol, B., 490, 491  
 Vinti, J., 463
- Walecka, J. D., 617, 620  
 Weber, W. E., 367  
 Wheatstone, Sir C., 66  
 Wheeler, J. A., 324, 619, 620  
 Witten, I. H., 524
- Young, T., 559
- Zeeman, P, 232

# Subject Index

- 1-form, 289  
charge, current, 295  
covariant vector, 290  
definition, 290  
energy, momentum, 295  
figure, 290  
table, 290
- 4-vector  
energy, momentum, 295, 300, 301  
photon momentum, 304  
table, 287  
velocity, 286–288
- 4-velocity, 286–288
- Abbreviated action, 354, 434  
Abelian group, 606  
Acceleration, centripetal, 29, 80  
Acoustics, 53, 237, 239  
Action, 356  
abbreviated, 359, 434  
and reaction, 7  
strong law, 7, 10  
weak law, 5  
at a distance, 323, 583  
integral, 359, 596  
variable, 452  
integral over orbit, 458
- Action-angle variable, 430, 452–478, 619  
celestial mechanics, 456  
chaos, 485  
completely separable, 457–466  
degeneracy, 73, 464, 468  
harmonic oscillator, 456  
Kepler problem, 466–478  
one degree of freedom, 452–457  
periodic motion, 452
- perturbation, 541  
proper variables, 481
- Adiabatic invariance, 549–555
- Algebra, 611  
Clifford, 614, 621  
geometric, 617  
Heisenberg, 613  
Lie, 611, 612. *See also* Lie algebra
- Analogy, structural, 54
- Analytical mechanics, 1
- Angle variable, 455  
Fourier expansion, 460  
libration, 460  
multiply periodic, 460  
quasi-periodic, 461  
rotation, 461  
time dependence, 454, 458, 460
- Angular momentum  
4-vector, 310  
areal velocity, 73  
canonical, 405  
central force problem, 72  
conservation, 3, 72, 73, 571  
total, 7  
definition, 2  
density, total, 571  
eigenvalue, 411  
electromagnetic, 8  
ellipsoid, 203  
mechanical, 8, 405  
Poisson bracket, 408, 411  
relativistic, 309  
rigid body, 188  
spherical symmetry, 72  
spin, 10  
total, 8
- Angular velocity in Euler angles, 602, 615  
Anharmonic oscillator, 545  
Anomalistic year, 131  
Anomaly  
eccentric, 100  
mean, 102  
true, 540  
Antiproton, 304  
Antiquark, 616  
Aphelion, 484  
Approximation, semiclassical, 115  
Apsidal  
distance, 78, 95, 96  
vector, 86  
Areal velocity, 73  
Ascending node, 472  
Astronomy, medieval, 100  
Attitude angle, 154  
Attractor, 489, 516, 620  
regular, 493  
strange, 489, 492, 500  
strange, Hénon-Heiles 500  
Atwood's machine, 27, 28
- Axis  
rigid body, 135  
screw, 161  
semimajor, 95, 475  
semiminor, 101  
symmetry, 161  
Azimuth, 209
- Backward glory, 114  
Bank angle, 154, 603  
Barrier, centrifugal, 112  
Baryon, 615, 616  
Basis vector, 286  
Bertrand's theorem, 89, 92

- Bessel function, 126  
 Biform, 296  
 Bifurcation, 454, 484, 505, 513, 514  
     diagram, 506, 508, 513, 515  
 Bilinear, 219, 388  
     form, 194  
 Binet ellipsoid, 203, 204  
 Biot-Savart law, 7  
 Bivector, 296  
 Black box, 121  
 Bohr  
     quantum mechanics, 466  
     theory, 95  
 Boltzmann  
     constant, 185  
     factor, 128  
 Boost, 280. *See also* Lorentz transformation  
 Bounded motion, 80, 484  
 Boyle law, 128  
     and virial, 84  
 Brachistochrone, 42, 63
- Calculus  
     operational, 275  
     of variations, 36, 43  
         fundamental lemma, 38
- Canonical, 338  
     equations of Hamiltonian, 338  
     extended transformation, 371  
     invariant, 388  
     momentum, 55, 314  
         relativistic, 322, 323  
     perturbation theory, *see*  
     Perturbation theory  
     restricted transformation, 371  
     variables, 335, 377
- Canonical transformation, 348, 368–421, 619  
     active and passive, 400, 405  
     cyclic Hamiltonian, 369, 377, 399, 430, 441  
     degeneracy, 464, 470  
     equations, 368–375  
     examples, 375–377  
     explicit time dependence, 385, 397, 402  
     generated by Hamiltonian, 420
- generating function, 373  
 group, 387  
 harmonic oscillator, 377–381  
 infinitesimal, 385, 402  
 invariant,  
     phase space volume, 393, 420  
     Poisson bracket, 389  
 Jacobi matrix, 382, 394  
 parametric, 385, 405, 408  
 restricted, 371, 381, 382, 387  
 symplectic, 381–388  
 table of, 373
- Cantor set, 516, 519, 522  
 Capacitance, 271  
 Carathéodory theorem, 394  
 Carousel, 183  
 Cartesian coordinates, 25, 141  
 Catenary, 41, 42, 64  
 Cayley-Klein parameters, 154, 182, 601, 602  
 Celestial mechanics, 533  
 Center of  
     energy, 312  
     force, 106  
     gravity, 185  
     mass, 5, 6, 185, 312  
     momentum, 301, 312  
         system, 301
- Central force problem, 70–126.  
*See also* Kepler problem
- Centrifugal  
     barrier, 112  
     effect, 126
- Centripetal acceleration, 29, 175
- Chain rule, 18
- Chandler wobble, 208, 228
- Chaos, 483–522, 617, 619  
     attractor, 489–491  
     bifurcation, 505–509  
     damped harmonic oscillator, 505–509  
     dimensionality, 616–522  
     fractals, 516–522  
     Hénon-Heiles, 496–503, 506  
     Islands, 503–505  
     KAM theorem, 487–489  
     logistic equation, 509–516  
     motion, 491  
     onset, 492, 501, 503
- parametric  
     oscillator, 508  
     resonance, 509
- perturbation theory, 487–489  
 properties of, 491  
 trajectory, 491, 494, 521, 522
- Character table, 611  
 Characteristic  
     equation, 157  
     value, 156
- Charge density, 588  
 Charged particle in  
     electromagnetic field, 23, 317, 553
- Charm, 615, 616
- Chasles' theorem, 161, 184
- Class of group, 607
- Classical mechanics, 1–600
- Clifford algebra, 614, 621
- Closed orbit, 89, 452
- Colliding beam, 304
- Collision  
     elastic, 118, 120, 306  
     inelastic, 118
- C-O-M, center of momentum, 301
- Commensurability, 463  
     condition, 464
- Commensurate, 105, 463  
     completely, 464  
     condition, 464  
     frequency, 462  
     m-fold, 464
- Commutator, 171, 411  
     quantum mechanics, 392, 398  
     relations, 170
- Configuration space, 34, 357  
     point transformation, 370  
     variation, 36
- Congruence transformation, 245, 246, 252
- Conic section, 94, 99
- Conjugate momentum, 55, 335, 351
- Conservation  
     differential theorem, 594  
     energy function, 62  
     momentum, 403
- Conservation theorems, 7, 55, 72, 343, 597

## Subject Index

- angular momentum, 3, 344  
     total, 7  
 canonical momentum, 315, 340  
 energy, 4, 11, 345, 450  
 linear momentum, 2, 6, 344  
     system of particles, 6  
 Noether's theorem, 589  
 Poisson bracket, 396, 402  
 relation to symmetry properties, 54–59
- Conservative system, 4  
 Conserved current, 594, 595  
 Constant of motion, 105, 397, 402, 403, 415  
     algebraic, 418  
     central force, 105  
     Jacobi identity, 397, 411  
     Poisson bracket, 398
- Constraint, 12–16, 24  
     differential, 16  
     equation, 15  
     holonomic, 12  
     nonholonomic, 12  
     nonintegral, 16  
     rheonomous, 13  
     rigid body, 12  
     rolling, 182  
     schleronomous, 13, 25  
     semiholonomic, 46, 49  
     virtual work, 16, 17, 48  
     weak, 321
- Continuity  
     conditions, 572  
     equation, 595
- Continuous system, 265, 558, 568  
     Hamiltonian formulation, 572–577  
     Lagrangian density, 561–566  
     stress energy tensor, 566–572  
     transition from discrete to continuous, 558–561
- Contour integration, 469
- Contraction, 290, 295  
     of tensor, 191
- Contravariant, 289
- Control parameter, 503, 506  
     logistic equation, 510
- Coordinate  
     basis, 286
- Cartesian, 184  
 contraction, 295  
 cyclic, 55, 343, 369, 445  
 generalized, 13, 19, 239  
 internal, 272  
 mass weighted, 241, 258  
 normal, 251, 257, 259  
 polar, 72  
 pseudo-Cartesian, 294  
 rotating, 175
- Coriolis, 174–179  
     acceleration, 176  
     circulation of fluid dynamics, 177  
     deflection, 176–178, 182  
     effect, 126, 174–179, 326  
         on meteorological phenomena, 177  
     force, 175  
     Foucault pendulum, 179  
     hemisphere, 178  
     pressure gradient, 176
- Correspondence principle, 325, 390, 392, 398  
     Poisson bracket, 388, 391, 392, 398
- Cosmological constant, 328  
 Cosmology, 617
- Coulomb  
     field, 109, 111  
     law, 274  
     scattering, 110
- Coupled electrical circuits, 53
- Covariant  
     definition, 277  
     equation, 297  
     Hamiltonian, 349, 352  
     Lagrangian, 318, 321, 322, 350, 352  
     principle, 325  
     relativistic, 577  
     vector, 289
- Cramer's rule, 149, 263
- Cross section  
     high energy limit, 127  
     Rutherford, 110  
     total, 110
- Crossing the T, 8
- Current  
     conserved, 594, 595  
     density, 588  
         elastic rod, 567  
         field flow, 568, 571, 594  
         flow, RL circuit, 51
- Curvature scalar, 327
- Cyclic  
     coordinate, 55, 343, 369  
         Kepler problem, 445  
     group, 606
- Cyclotron  
     frequency, 318, 553  
     resonance, 318
- $\delta$ -function, 588  
 $\delta$ -variation, 38  
 $\delta_{ij}$  Kronecker delta, 138, 181, 190  
 $\Delta$ -variation, 357–359
- D'Alembert  
     characteristic, 548  
     principle, 16–20, 46, 313
- D'Alembertian, 296
- Damping, 519  
     exponential, 262  
     van der Pol equation, 490
- Deflection angle, scattering, 113
- Degeneracy, 244, 465  
     conditions, 465  
     exact, 547  
     Kepler problem, 470, 484  
     proper, 547  
     vibrational modes and frequencies, 257
- Degrees of freedom, 13, 245, 255, 342, 427, 541, 549, 563
- Hamiltonian, 342  
     many, 457  
     molecular vibrations, 256  
     *n* particles, 13  
     oscillator, 264  
     rigid body, 135  
     vibration, 257
- Delaunay variables, 477
- Delta  
      $\delta$ -function, 588  
      $\delta$ -variation, 38  
     Kronecker ( $\delta_{ij}$ ), 138, 181, 190
- Dense quasi-periodic orbits, 491

- Derivative, functional, 574  
 Deterministic, 483  
 Differential equation,  
     inhomogeneous, 259  
 Diffusion, 524  
 Dihedral group, 606, 607  
 Dilation of time, 279  
 Dimension  
     Cantor set, 516  
     fractal, 516, 517  
     Hausdorff, 517  
 Dipole moment  
     gravitational, 226  
     magnetic, 185, 230  
 Dirac  $\delta$ -function, 588  
 Direction cosines, 136  
     orthogonality, 138  
     transformation, 139  
 Dissipation  
     exponential damping, 263  
     forces, 259  
     function, 22–24, 53, 63, 261  
         Rayleigh, 23  
 Disturbing function, 533  
 Divergence, 295  
     4-divergence, 296  
     relativistic, 565  
     theorem, four dimensional, 581, 593  
 Divergenceless, 581  
 Doppler effect, 329  
 Drag force, 24, 52  
 Dual space, 292  
 Duffing oscillator, 523  
     inverted, 524  
 Dumbbell molecule, 347  
 Dynamic steady state, 267  
 $\epsilon_{ijk}$  Levi-Civita density,  
     permutation symbol, 169, 410  
 Earth  
     equatorial bulge, 223  
     figure axis, 226  
     Lagrangian for precession, 227  
     potential, 226, 227  
     precession, 226  
     spinning ring, 229  
     torques, 237  
 Earth-Moon system, 124  
 Eccentric anomaly, 99  
 Eccentricity, 94, 95, 532  
     Sun and Moon, 227  
 Ecliptic, 208, 228  
 Eigenvalue, 156–158  
     angular momentum, 411  
     equation, 157  
         oscillations, 241  
     Euler's theorem, 157  
     inertia tensor, 195, 196  
     linear triatomic molecule, 254  
     problem, 157  
     transformation matrix, 160  
 Eigenvector, 247  
     indeterminacy, 258  
     inertia tensor, 196  
     linear combination, 248  
     orthonormal, 249  
     oscillations, 244  
 Eigenwerte (German for eigenvalue), 156  
 Einstein  
     field equations, 327, 538  
     summation convention, 139  
     tensor, 327  
     velocity addition law, 283, 328  
 Elastic  
     collision, 118, 120, 306  
     scattering, 120  
     solid, 563  
     wave, 560  
 Electric circuit  
     equation, 264  
     Lagrangian, 53  
 Electromagnetic  
     field, 31, 51, 55, 275, 571, 587  
         Lagrangian, 350  
         Lagrangian, covariant, 352  
         potential, 342  
         radiation, 54  
         theory, 276  
     Elementary particle, 51, 54, 300, 615  
 Ellipse, 81  
     figure, 96  
     harmonic oscillator, 377  
     orbit equation, 484  
     phase space plot, 98  
 properties, 97  
 semimajor axis, 95, 475  
 shape, scale, orientation, 105, 473  
 table, 97  
 Ellipsoid  
     Binet, 203, 204  
     inertia, 196, 201  
     kinetic energy, 204, 258  
         rigid body, 185–188  
     moment of inertia, 197  
 Ellipsoidal coordinates for Hamilton-Jacobi equation, 479  
 Elliptic  
     function, 89  
     integral, 234  
     region for chaos, 504  
 Elsewhen, 279  
 Elsewhere, 279  
 Energy  
     center of, 312  
     conservation, 60  
         central force, 74, 77  
         free, Gibbs, 337  
         free, Helmholtz, 337  
     function, 60–63, 314  
         conserved, 61–63  
         hypersurface, 494  
         potential, 4  
 Ensemble, 419  
     microcanonical, 421  
 Enthalpy, 336  
 Equant, 129  
 Equation  
     of motion, 74  
     of state, gas, 85  
 Equilateral triangle group, 609  
 Equilibrium  
     generalized forces, 238  
     indifferent, 240  
     neutral, 240  
     stable, 238  
     statistical, 421  
     unstable, 239  
 Equinox, 539  
     precession, 223–230  
 Equipotential curve  
     gravitation, 125

- Hénon-Heiles, 498  
 Equivalence principle, 324, 346  
 Ergotic hypothesis, 418  
 Escape velocity, 31  
 Ether, 566  
 Euclidean  
     dimension, 518, 519  
     space, 517  
 Euler  
     equations, 198, 199, 234  
         derived from Lagrange's equations, 200  
         heavy symmetrical top, 210  
         symmetric body, 205  
     inhomogeneous function, 86  
     parameters, 155, 182, 602, 603  
     solution of three body problem, 122  
     theorem, 155–161  
         homogeneous functions, 320  
 Euler angles, 150–154, 196, 601  
     angular velocity, 602, 615  
     conventions, 154  
     figure, 152  
     infinitesimal, 165  
     left handed, 152  
     SU(2) rotation, 412  
     time changes, 210  
      $x$ -convention, 154, 601  
      $xyz$ -convention, 154, 603  
      $y$ -convention, 154, 601  
 Euler-Lagrange, 564  
     complex scalar field, 583  
     electromagnetic field, 587  
     equation, 45, 64, 65, 319, 354  
     relativistic equation, 564, 588  
 Event, 279, 311  
 Extremum  
     path, 40  
     problem, 39  
     surface area, 40  
 Faraday tensor, 297, 298  
 Feigenbaum  
     diagram, logistic equation, 510, 513–515  
     number, 514  
     plot, 506  
     point, 511  
 Fermat's principle, 360  
 Field  
     canonical equations, 574  
     classical theory, 571  
     complex, 596  
         scalar, 583  
     definition, 566  
     elastic, 51  
     electromagnetic, 31, 51, 55, 275, 571, 587  
     elementary particle, 51  
     equation, Lagrange-Euler, 583  
     gravitational, 176, 185, 210, 275  
     relativistic, 571  
     scalar 287, 583  
         meson, 571, 599  
     theory, 558–589  
         Hamiltonian formulation, 571–577  
         Noether's theorem, 589–598  
         relativistic, 583–598  
 Schrödinger  
     quantum theory, 576  
     spacetime, 566  
     vector, 286  
     velocity, 588  
     wave function, 571  
 Figure axis, 539  
 Fission, 120  
 Fluid  
     dynamics, 419  
     perfect, 579  
 Flux density, 107  
 Force  
     central, 7, 70  
     centrifugal, 176  
     cut off, 111  
     driving, 259  
     effective, 80, 94, 175  
     electromagnetic, 259  
     external, 5  
     generalized, 19, 21, 58, 238  
     gradient of potential, 10  
     gravitational, 93  
     inertial, 5  
     internal, 5, 11  
     inverse square, 77, 92  
     linear restoring, 83  
     long range, 110  
 Lorentz, 22, 131, 237, 317, 350  
 Minkowski, 299, 322  
 relativistic, 297  
 reversed effective, 18, 80  
 strong, 299  
 weak, 299  
 Foucault pendulum, 179, 183, 184  
 Four-vector, *see* 4-vector  
 Four-velocity, *see* 4-velocity  
 Fourier  
     series, 14, 126, 574  
         convergence, 545  
         multiple, 460  
         transform, 274  
 Fractal, 516, 620  
     area, 521  
     dimension, 490  
         Sierpinski carpet, 519  
     geometry, 491  
     self-similarity, 505, 514  
 Free energy  
     Gibbs, 337  
     Helmholtz, 337  
 Frequency  
     characteristic, 266  
     commensurate, 106  
     critical, 266  
     cyclotron, 318, 553  
     driving, 490  
     imaginary, 244  
     Larmor, 231  
     resonant, 490  
 Friction, 24  
     atmosphere, 32  
     drag, 24  
     electrical, 52  
     oscillating system, 262  
     rolling, 17  
 Functional, 287, 293  
     derivative, 574, 575  
 Future, 279  
 Galactic center, 496  
 Galaxy model, Hénon-Heiles, 496, 497, 516  
 Galilean  
     system, 2  
     transformation, 276–280  
 Gas, equation of state, 85

- Gauge transformation, 595  
 general, 619
- General relativity, 324
- Generalized  
 force 19, 21, 57, 58  
 mechanics, 65
- Generating function, 371, 372  
 canonical transformations, table  
 of, 373
- chaos, 488
- infinitesimal  
 canonical transformation  
 (I.C.T.), 403
- rotation, 404
- Poisson bracket, 404, 406
- symplectic, 394
- table, 373
- Geodesic 40, 324–326, 362  
 deviation, 325
- Geoid, 176
- Geostrophic wind, 178
- Gibbs free energy, 337
- Glory scattering, 114
- Goldschmidt solution, 64, 65
- Gradient, 295
- Gram-Schmidt method, 249
- Gravitational  
 charge, 226  
 field, 176, 185, 210, 275  
 quadrupole moment, 226
- Greek subscript convention, 286
- Group  
 abelian, 606  
 canonical transformation,  
 387  
 class, 607  
 conjugation, 606  
 cyclic, 606  
 definition, 605  
 dihedral, 606, 607  
 generator, 606
- Lorentz, 282, 610  
 multiplication table, 606  
 properties, 387, 605–611  
 quaternion, 610  
 representation, 608  
 rotation, 171  
 symmetry, 412  
 for system, 413
- symplectic, 387, 612  
 theory, 605
- Gyration, radius of, 198
- Gyrocompass, 223
- Gyromagnetic ratio, 230
- Gyroscope  
 inertia, 222  
 torque free mounting, 213
- Hamilton's principle, 34–36,  
 44–50, 313, 324, 355, 562,  
 564
- Lagrange's equations  
 derivation, 44, 45  
 modified, 354, 355, 599  
 nonholonomic systems, 45–50
- Hamiltonian, 334–353  
 as total energy, 339  
 covariant, 349, 352  
 degrees of freedom, 342  
 density, 573, 586  
 formulation  
 continuous systems, 572  
 relativistic mechanics, 349
- generates canonical  
 transformation, 420
- generator of system motion, 399
- Hénon-Heiles, 492, 497, 522
- perturbation, 526
- quantum mechanics, 613
- symplectic, 576
- Hamiltonian formulation, 334–363  
 advantages, 51–54  
 characteristic function, 434,  
 440–444  
 comparison of characteristic and  
 principal functions,  
 442–443  
 conservation theorems,  
 347–349  
 cyclic coordinates, 343–349  
 Hamilton equations of motion,  
 334–363, 368, 397, 402  
 derived from variational  
 principle, 353  
 least action principle,  
 356–363
- Legendre transformation  
 derivation, 334–342
- principal function, 430–434,  
 433, 528  
 compared with characteristic  
 function, 442
- relativistic formulation,  
 349–353
- Routh procedure, 347–349
- symplectic approach, 339–343
- variational principle derivation,  
 353–356
- Hamilton-Jacobi theory and  
 equation, 334, 430–451,  
 488, 528, 549, 619
- central force, 448
- chaos, 485
- completely separable, 444
- cyclic coordinates, 445–451
- ellipsoidal coordinates, 479
- harmonic oscillator, 434–439
- method, 434–439
- new constant coordinates, 432
- Kepler problem, 445–451  
 spherical coordinates, 451
- separation of variables,  
 444–445
- two methods of solution, 442
- Handedness convention, 169
- Harmonic oscillator, 434–440,  
 485
- action-angle variables, 455, 456,  
 485
- adiabatic invariant, 550
- canonical transformation, 377
- constants of motion, 417
- coordinate space plot, 440
- damped, 269  
 driven, 505, 507
- ellipse, 377
- Feigenbaum plot, 508
- Hamilton-Jacobi, 434–440
- isotropic, 82  
 three dimensional, 275
- perturbation, 529, 542
- phase diagram, 380
- Poisson brackets, 417
- relativistic, 316
- two dimensional, 415, 416  
 anisotropic, 437
- Heading angle, 154

- Heisenberg  
algebra, 613  
picture, 408
- Helmholtz free energy, 337
- Hénon-Heiles  
chaos, 484  
equipotentials, 498  
galaxy model, 516  
Hamilton equations, 497  
Hamiltonian, 492, 496, 497, 522  
islands in chaos, 502  
Poincaré map, 499–501  
potential, 497
- Hermitean matrix, 412
- Herpolhode, 202, 203
- Hertz principle of least curvature, 361
- Hierarchy of islands, 504, 505
- High energy physics, 300
- Hodograph, 131
- Holonomic, 12  
constraint, 12  
system, 199
- Homogeneous  
function, 320  
problem, 320, 359
- Homomorphism, 418, 609
- Hooke's law, 52, 92, 317, 559
- Hoop  
rolling, 50  
vertical, 66
- Huygens' waves, 132
- Hydrodynamic derivative, 419
- Hyperbola, 81, 316
- Hyperbolic  
motion, 315  
point, 504  
region, 504
- Hypersurface, 580  
energy, 494  
spacelike, 580
- Hypocycloid, 64
- Hysteresis, 270, 271, 523
- I.C.T. (infinitesimal canonical transformation), 385, 386, 402, 403, 408, 410, 413
- Identity transformation, 146, 156
- Ignorable, *see* cyclic
- Imbedding in chaos, 514–516
- Impact parameter, 107
- Inclination, 532
- Incommensurate, 548  
frequency, period, 462, 489, 548  
oscillator, 521
- Inelastic collision, 118
- Inertia  
ellipsoid, 197, 201  
tensor, 191  
components, 195  
diagonal, 196  
eigenvalue, 195, 196  
eigenvector, 196  
integral, 194  
principal axes, 196  
principal moments, 197  
properties, 195  
similarity transformation, 196
- Inertial  
force, 5  
system, definition, 2
- Infinitesimal  
canonical transformation, 385, 386, 396, 398, 399, 401  
rotation, 163, 166
- Infrared spectroscopy, 258
- Instability, 205
- Integrability breakdown, 502
- Integral  
invariants of Poincaré, 394  
Jacobi, 61  
line, 35  
variation, 44
- Integrating factor, 15
- Invariable plane, 202
- Invariance  
adiabatic, 549  
condition, 594  
group, 613  
logistic equation, 484  
Lorentz, 302  
Poisson bracket, 388  
rotation, 60  
scale, 591  
translation, 60
- Inversion, 150, 181
- Islands  
in chaos, 502, 503
- hierarchy, 504, 505  
various orders, 504
- Isomorphism, 609
- J*-matrix, 342, 382–389, 393  
Poisson bracket, 388
- Jabberwocky, 202
- Jacobi  
determinant, 394  
form of least action principle, 361  
identity, 393, 398, 424, 428  
integral 61, 566, 597  
Lagrange brackets, 424  
Poisson bracket, 390  
matrix of canonical transformation, 426
- Josephson junction, 265, 271, 618
- KAM (Kolmogorov-Arnold-Moser) theorem, 484, 487–492
- Kamiltonian, 370
- Kepler  
equation, 102, 126, 131  
second law, 73  
third law, 101, 470
- Kepler problem, inverse square law potential, 70–126, 347, 415  
action variables, 471  
action-angle variables, 466  
closed orbits, conditions, 89–92  
cyclic coordinate, 445  
equations of motion, 72–76  
equivalent one body problem, 70–71  
equivalent one dimensional problem, 76–83  
inverse square law, 92–96  
Lie algebra, 414  
motion in time, 96  
orbit equation, 86–89, 96–103  
perturbation, 536  
Poincaré map, 495, 496  
scattering, 106–121  
spherical polar coordinates, 467  
symmetry group, 414  
virial theorem, 472

- Kinematics  
 rigid body, 134, 184  
 tools, 184
- Kinetic energy  
 ellipsoid, 203  
 rigid body, 184  
 rotational, 191  
 total, 9
- Kinetic theory, 85, 112
- Kirchhoff junction conditions, 66
- Klein-Gordon  
 equation, 585  
 field, 585  
 particle, 596
- Kolmogorov-Arnold-Moser (KAM) theorem, 484, 487–492
- Korteweg-deVries equation, 596, 600
- Kronecker delta ( $\delta_{ij}$ ), 138, 181, 190
- Laboratory  
 frame, 302  
 system, transformation, 306  
 time, 279
- Lagrange  
 bracket, 392–394  
 fundamental, 393  
 calculus of variations, 36  
 equations, 16, 21–23  
   derivation from Hamilton's principle, 44, 45  
 Euler equation derivation, 200  
 Nielsen form, 30  
 perturbation, 533  
 multipliers, 16, 67  
 point, 124  
 solution of three body problem, 123  
 undetermined multiplier, 198
- Lagrangian  
 applications, 24–29  
 central force, 71  
 conserved quantities, 566  
 covariant, 318, 321, 322, 352  
 definition, 21
- density, 564, 567, 583  
 continuous system, 561–566  
 discrete system, 558–560
- electromagnetic field, 350
- formulation versus Newtonian, 199  
 from Hamilton's principle, 44  
 heavy symmetrical top, 208  
 precession of Earth, 227  
 relativistic, 312  
 rigid body, 185, 199  
 separable, 185
- Laplace transform, 264
- Laplace-Runge-Lenz vector, 102–106, 131, 429
- Larmor  
 frequency, 231  
 precession, 318  
 theorem, 232
- LC circuit, 51
- Least action principle, 356, 362  
 $\Delta$ -variation, 359  
 Jacobi form, 361  
 restrictions, 358
- Legendre  
 polynomial, 539  
 polynomial generating function, 224  
 transformation, 334, 335, 375, 549
- Levi-Civita density,  $(\epsilon_{ijk})$  169, 410
- Liapunov exponent, 491, 519  
 damped pendulum, 519  
 diagram, 520  
 dimension, 521  
 logistic equation, 514, 519  
 negative, 492  
 Sierpinski carpet, 519  
 solar system, 494
- Libration, 452, 455, 460
- Lie  
 algebra, 171, 412–415, 611–613  
 definition, 412  
 Kepler problem, 414  
 Poisson bracket, 392  
 structure constant, 413, 612
- bracket, 171  
 relations, 415
- group, 411, 412, 611–613  
 subgroup, 613
- Light cone, 279, 280
- Lightlike, 278, 304
- Limit cycle, 489  
 figure, 491  
 van der Pol equation, 491
- Line of nodes, 150, 473
- Linear momentum, 1  
 particle, 1  
 system of particles, 6  
 total, 6
- Liouville theorem, 418–421, 428, 483
- Lissajous figure, 83, 258, 439, 458, 462  
 noncommensurate, 464  
 sketch, 440, 463
- Ljapunov, *see* Liapunov
- Logistic equation, 509, 620  
 control parameter, 510  
 Feigenbaum diagram, 510, 513–515  
 fourfold cycle, 510  
 iterations, 510  
 Liapunov exponent, 512, 514  
 self-similarity, 514  
 twofold cycle, 510
- Longitude of ascending node, 474
- Lorentz, 282  
 boost, 284  
 condition, 297  
 force, 22, 131, 237, 350  
 frame, 580  
 group, 282, 610  
 invariance, 302, 577  
 ten constraints, 282  
 transformation, 280–265  
   boost, 282  
   equations for  $ct'$  and  $r'$ , 281  
   general matrix, 281  
   homogeneous, 282  
   inhomogeneous, 282, 610  
   invariance, 302  
   pure, 284  
   scattering, 306
- Lorenz equations, 523
- Lyapunov, *see* Liapunov

- M*-matrix, 382–389, 394  
*MacCullagh* formula, 225  
*Mach's* principle, 324  
**Magnetic**  
 field  
 charge particle motion, 23, 317  
 uniform, 409  
 moment, 230  
 rigidity, 318  
**Manifold**, 576, 611, 618  
**Mapping**, 287  
 quadratic, 503  
**Mass**  
 center of, 312  
 reduced, 71  
 weighted coordinates, 241  
**Matrix**  
 addition, 145  
 antisymmetric, 148, 165  
 cofactor, 340  
 determinant, 159  
 hermitean, 412  
 infinitesimal element, 164  
 inverse, 147  
 $J$ -, 342, 383–389  
 $M$ -, 382–389, 394  
 multiplication, 144  
 orthogonal, 147  
 reciprocal, 147  
 rectangular, 147  
 skew symmetric, 148  
 transpose, 147  
 unitary, 412  
**Maxwell's** equations, 54, 276, 297, 350  
 covariant form, 298  
**Mean anomaly**, 102  
**Mechanics**, *see* Classical mechanics  
**Merry-go-round**, 183  
**Meson**, 331, 616  
 scalar, 571, 599  
**Metric**  
 Minkowski space, 287, 580  
 matrix, 287  
 tensor, 327  
**MeV**, definition, 32  
**Microcanonical ensemble**, 421  
**Million electron volt**, definition, 32  
**Minimum**  
 gravitational coupling principle, 325  
 surface of revolution, 40  
**Minkowski**  
 coordinate, 288  
 force, 299, 322  
 space, 278, 580  
 two dimensional, 287  
**Mixing**, 516  
 property of chaos, 491  
**Mode**, normal, 252  
**Moderator**, 120  
**Molecule**  
 internal coordinates, 272  
 linear triatomic, 272  
 pentatomic, 272  
 polyatomic, 258, 259  
 rotation and vibration, 180  
 triatomic, 275  
 vibrating, 253, 258  
 linear polyatomic, 558  
**Moment**  
 of force, definition, 2  
 of inertia, 191  
 about axis of rotation, 192  
 choice of origin, 193  
 coefficients, 187  
 ellipsoid, 197  
 integral, 194  
 operator, 188  
 parallel axes, 193, 194  
**Momentum**  
 angular, 187, 344  
 canonical, 55, 314  
 center of, 312  
 conjugate, 55, 335, 351  
 conservation, 403  
 density, 569, 573, 579  
 electromagnetic, 55  
 generalized, 55  
 linear, 1, 6, 24, 344  
 representation, 576, 598  
**Monochromatic light**, 259  
**Monogenic**, 34  
**Monopole**, magnetic, 131, 427  
**Motion**  
 bounded, 80, 484  
 chaotic, 491–493  
 equation, 74  
 hyperbolic, 315  
 periodic, 484  
**Multiplet**, 615, 616  
**Multiply periodic**, 458, 461  
**Multivector**, 614  
**Napier's rules**, 476  
**Network**, electrical, 264  
**Neutron scattering**, 120  
**Newtonian**  
 equations of motion, 199  
 formulation versus Lagrangian, 199  
 mechanical corpuscles, 132  
 second law, 1, 299  
 third law, 5  
**Nielsen form of Lagrange's equations**, 30  
**No-interaction theorem**, 324, 353  
**Node**  
 ascending, 472  
 line of, 150, 473  
**Noether's theorem**, 344, 566, 589, 594  
 conditions, 590  
 conserved current, 594  
 conserved quantities, 418  
 discrete, 596, 597  
 statement of, 594, 595, 597  
 symmetry properties, 598  
**Non-Euclidean**, 278  
**Nonabelian group**, 606  
**Noncommensurate**, 464  
**Nonholonomic system**, 45  
**Noninertial system**, 175  
**Normal**  
 behavior in chaos, 515  
 coordinates, 250, 251  
 modes, 252, 256  
**Number theory theorem**, 463  
**Nutation**, 215  
 heavy symmetrical top, 209, 214

- O(3) group, 610  
**Oblateness**  
  Earth, 229  
  Moon, 229  
**Occupation number**, 253  
**One dimensional problem**,  
  equivalent, 76  
**One-form**, *see* 1-form  
**Operational calculus**, 275  
**Optics**  
  geometric, 112  
  meteorological, 114  
**Orbit**  
  bounded, 80  
  chaotic, 522  
  circular, 80, 81, 94  
  closed, 452  
    conditions for, 89  
  commensurate, 106  
  degenerate, 106  
  elliptic, 94, 95, 484  
  equation, 99  
    of state, 86  
    integration, 93  
  hyperbolic, 94, 110  
  inclination, 474  
  open, 452  
  osculating, 531  
  parabolic, 94  
  phase space, 452  
  quasi-periodic, 490  
  reflection symmetry, 87  
  regular, 522  
  satellite, 229  
  shape, scale, orientation, 105,  
    473  
  stable, 90  
  unbounded, 79  
  unstable, 90  
**Orbiting**, 113  
**Orthogonal**  
  matrix, 147  
  transformation, 139  
**Orthogonality condition**, 140  
**Oscillation**, 238–265  
  eigenvalue equation, 241–249  
  forced, 259–265  
  free vibration frequencies,  
    249–253  
  Josephson junction, 271  
  normal coordinates, 249–253  
  pendulum, damped and driven,  
    265–271  
  potential expansion, 238–241  
  principal axis transformation,  
    241–249  
  triatomic molecule, 253–259  
**Oscillator**  
  anharmonic, 545  
  double, 486  
  parametric, 508  
**Parabola**, 81, 94, 128  
**Parametric resonance**, 505, 508,  
  509  
**Parity**, 590  
**Past**, 279  
**Pauli matrices**, 412, 612, 614  
**Pendulum**  
  damped driven, 265  
  double, 14  
  equation, 267  
  hysteresis, 270  
  periodicity, 453  
  perturbation, 533  
  phase angle, 533  
  plane, 234  
  spherical, 83, 428  
**Pentatomic molecule**, 272  
**Periapsis**, 99, 108, 540, 541  
**Periastra**, 474  
**Pericynthion**, 99  
**Perigee**, 474  
**Perihelion**, 99, 100, 474, 477,  
  484  
  Mercury, 332, 538, 539  
**Period doubling**, 516  
**Periodic**  
  frequency, 455  
  motion, 452, 484  
    libration, 452  
    rotation, 452  
  multiply, 458  
  orbits of pendulum, 454  
  quasi, 461  
**Permutation**  
  group, 609  
  symbol ( $\epsilon_{ijk}$ ), 169, 173, 181  
**Perturbation**, 487  
  action-angle variables, 541  
  adiabatic invariance, 549–555  
  degeneracy, 547, 548  
  fast variable, 547  
  first order, 530, 534, 537  
  Hamilton-Jacobi equation, 543  
  Hamiltonian, 526  
  harmonic oscillator, 529  
  Kepler problem, 536  
  n-th order, 530  
  pendulum, 533  
  precession  
    equinoxes, 539  
    Mercury, 538, 539  
    satellite orbits, 539  
  second order, 534, 544  
  secular, 532, 535  
  slow variable, 547  
  solar system, 532  
  theory, 229, 338, 483,  
    526–555  
    quantum, 527  
    time dependent, 527–533  
      examples, 533–541  
      time independent, 541–549  
**Phase space**, 335, 370, 453, 573  
  ellipse, 98  
  harmonic oscillator, 380  
    damped driven, plot of, 507  
    uncoupled, 486, 487  
  Kepler problem, 98  
  orbits, 454  
  point transformation, 370  
  regular orbits, Hénon-Heiles,  
    502  
    trajectory, 354  
**Photomeson production**, 304  
**Photon**, 253  
**Pitch angle**, 154, 603  
**Planck's constant**, 380  
**Poincaré**  
  integral invariants, 394  
  map, (or section), 494, 495  
    Hénon-Heiles, 499–501  
    Kepler problem, 495  
    transformation, 282  
**Poinsot's construction**, 201, 202,  
  206, 234

## Subject Index

- Point  
inflection, 42  
Lagrange, 124  
saddle, 124  
transformation, 31, 370, 422  
  configuration space, 370  
  phase space, 370  
turning, 78
- Poisson  
equation, 225  
theorem, 398
- Poisson bracket, 388–411  
  angular momentum, 408–411  
  applications, 396  
  canonically invariant, 390  
  conservation theorem,  
    402–404  
  correspondence principle, 390,  
    398  
  double, 390  
  equation of motion, 396–398,  
    407  
  fundamental, 389, 411  
  generating function, 402–406  
  infinitesimal canonical  
    transformation (I.C.T.),  
    398–405  
  integral invariants of Poincaré,  
    394  
  invariance, 388  
  Jacobi identity, 390  
  Jacobian determinant, 394  
  Lagrange bracket, 392  
  Lie algebra, 392  
  linear and angular momentum,  
    411  
  nested, 408  
  perturbation theory, 532  
  symmetry groups, 411–418  
  symplectic, 388, 389  
  theorem, 411
- Polar coordinate, 72  
  central force Lagrangian, 73  
  plane, 25  
  spherical, 32
- Polhode, 202
- Polyatomic molecule, 258, 259  
  linear, 558  
  rotation and vibration, 180
- Potential, 4  
  energy, 4  
  equilibrium, 239  
  total, 11  
  equivalent one dimensional,  
    central force, 78  
  generalized, 22  
  gradient, 10  
  Hénon-Heiles, 497, 498  
  hole, 82  
  integrable, 86  
  linear restoring force, 83  
  power law, 86, 87  
  scalar, 20  
  velocity dependent, 22–24
- Power series, 43
- Precession, 206  
  astronomical, 208, 228  
  average frequency, 217  
  Earth, 207, 226  
  equinoxes, 209, 223–229  
  fast and slow, 219  
  force free motion, 207  
  free body, 205  
  heavy symmetrical top, 209  
  Larmor, 231  
  magnetic field, 230  
  Mercury, 332, 538, 539  
  orbital plane, 540  
  pseudoregular, 218  
  regular, 218  
  satellite, 228  
  system of charges, 230  
  Thomas, 282, 330
- Principal axis transformation,  
  241
- Proper time, 279, 310, 321
- Proton-neutron reaction, 304
- Pseudoscalar, 614
- Pseudotensor, 189
- Pseudovector, 168, 614
- Ptolemaic system, 129
- Q* value, 304
- Quadratic  
  forms, diagonalization, 252  
  iterator, 509  
  mapping, 503
- Quadrature, 75, 211
- Quadrupole moment  
  gravitational, 226  
  Sun, 541
- Quantization, 54
- Quantum  
  commutator, 392  
  corrections, 115  
  electrodynamics, 54  
  field theory, 576  
  Hamiltonian, 613  
  Heisenberg picture, 408  
  mechanics, 111  
    Bohr, 466  
  perturbation theory, 526  
  scattering, 120  
  theory, 290  
  transition from classical  
    mechanics, 76
- Quark, 615
- Quasi-  
  periodic, 461, 490  
  static motion, 268
- Quaternion group, 610
- Radius  
  gyration, 198  
  vector, 73
- Rainbow scattering, 114
- Raman spectroscopy, 258
- Randomness, 483
- Rayleigh's dissipation function,  
  23
- Reactance, 53
- Regularity, 488  
  breakdown, 488
- Relativity, 276–328, 619  
  4-vector, 287  
  angular momentum,  
    309–312  
  collisions, 300–309  
  electromagnetism, 297–300  
  force, 297–300  
  general, 324–328, 538  
  Lagrangian, 312–324  
  metric tensor, 287, 288, 291  
  reduced mass, 71  
  spacetime, 278–280  
  special, 265, 276–324  
    postulates, 277

- Representation  
 faithful, 609, 613  
 group, 608  
 irreducible, 608  
 momentum, 576  
 Repulsive centrifugal barrier, 78  
 Residue, 469  
 Resonance, 260, 548  
 deep, 549  
 parametric, 509  
 shallow, 549  
 transients, 260  
 vibrating system, 260  
 Resonant frequency of linear triatomic molecule, 255  
 Reversed effective force, 80  
 Reversible process, 336  
 Rheonomic, 13  
 Ricci tensor, 327  
 Riemann  
 surface, 469  
 tensor, 326, 327  
 Rigid body, 11  
 angular momentum, 185–188  
 definition, 134–138  
 degrees of freedom, 134  
 equations of motion, 184, 198–200  
 Euler  
 equations, 198–200  
 theorem, 155, 156  
 heavy symmetrical top motion, 208–223  
 kinematics, 134, 184  
 Lagrangian, 199  
 motion, 134, 155–174  
 nutating, 209, 214  
 orientation, 169  
 rotation, 155–174  
 finite, 161–163  
 infinite, 163–171  
 solving problems, 198  
 torque free motion, 200–223  
 Rigidity, 318  
 Roll angle, 154, 603  
 Rolling  
 constraint, 14  
 disk, 15  
 hoop, 50  
 Rössler equations, 523  
 Rotation, 141, 452, 455  
 active sense, 143  
 clockwise, 162  
 counterclockwise, 170  
 finite, 161  
 formula, 162, 170  
 generator, 171  
 group, 171  
 infinitesimal, 162, 163  
 instantaneous axis, 172  
 kinetic energy, 191  
 matrix, 142  
 passive, 169  
 sense, 143  
 proper, 158  
 trace, 160  
 vector, 59  
 Routh  
 Kepler problem, 348  
 procedure, 56, 347  
 Routhian, 348  
 Rutherford  
 cross section, 110  
 scattering, 131  
 Satellite  
 artificial, 229  
 close, 229  
 orbiting Earth, 474  
 orbits, 223, 229  
 Scalar, 189, 293  
 curvature, 327  
 field, 287  
 meson, 571  
 field, 599  
 potential, 20  
 product, Minkowski space, 288, 290, 291  
 scale invariance, 591  
 transformation, 370  
 Scattering, 106, 306  
 angle, 112, 308, 309  
 center of mass, 116  
 cross section 107  
 deflection angle, 114  
 differential cross section, 107, 119  
 elastic, 118, 120, 306  
 glory, 114  
 inelastic, 118  
 laboratory coordinates, 115–121  
 neutron, 120  
 rainbow, 112  
 Rutherford, 111, 131  
 Schrödinger equation, 54, 584, 599  
 Schwarzschild solution of Einstein field equations, 538  
 Scleronomous, 13, 25  
 Screening, nucleus, 111  
 Screw  
 motion, 161  
 symmetry axis, 161  
 Secular  
 change, 531  
 equation, 157, 244  
 linear triatomic molecule, 254  
 perturbation, 532, 535  
 Self-similarity, 505, 514  
 fractal, 516–519  
 logistic equation, 513–515  
 Semiclassical approximation, 115  
 Semiholonomic, 46, 48, 49  
 Semimajor axis, 95, 475  
 Semiminor axis, 101  
 Sensitivity to initial conditions, 491  
 Separation constant, 445  
 Siderial  
 day, 175  
 year, 538  
 Sierpinski  
 carpet, 517–519, 522  
 fractal dimension, 518  
 sponge, 522  
 Sigma elementary particle, 615  
 Similarity transformation, 149, 158, 189  
 trace, 160  
 Simultaneity, 580  
 Sine-Gordon  
 equation, 585  
 field, 585  
 SO(3) group, 413, 418, 610  
 SO(4) group, 414  
 SO(n) group, 418  
 SOHO, 126  
 Solar day, 175

- Soliton, 587, 596  
 Sound vibrations in gas, 598  
**Space**  
 configuration, 34, 357  
 dual, 292  
 filling, 521  
 Minkowski, 278, 290  
**Spacelike**, 278, 580  
**Spacetime**, 278  
 interval, 278  
**Special relativity**, 276  
 postulates, 277  
**Spherical triangle**, 181, 476  
**Spin angular momentum**,  
 10  
**Spiraling**, 113  
**Stability**, 205  
 marginal, 493  
**Staeckel conditions**, 446, 447  
**Stationary**  
 path, 37  
 value, 35  
**Steady state, dynamic**, 267  
**Stochastic**, 483  
**Stokes' law**, 24, 52  
**Strange attractor**, 489, 492, 500  
 dimension, 521  
 fractal dimension, 520  
 Hénon-Heiles, 500, 501  
**Strangeness**, 615  
**Stress**  
 energy tensor, 566, 570, 589  
 conservation, 595  
 properties, 578  
 symmetrize, 572, 600  
 tensor, 570  
**Strong**  
 law of action and reaction, 7  
 nuclear force, 299  
**Structure**  
 analogy, 54  
 constant, 412, 413, 612  
**SU(2) group**, 413, 418, 612, 615,  
 616, 621  
**SU(3) group**, 418, 615  
**SU(4) group**, 616  
**SU(n) group**, 418, 615, 616  
**Subgroup**, 606  
**Submultiplet**, 615  
**Summation convention**, 138, 169,  
 186  
**Superconductivity**, 618  
**Supermultiplet**, 615, 616  
**Susceptance**, 53  
**Symmetry**  
 groups, 411–418  
 mechanical systems, 411–418  
 properties, 60  
 spherical, 60, 72  
**Symplectic**, 343, 381  
 approach, 339, 343  
 canonical transformation, 381,  
 382  
 condition, 384, 387, 422  
 generating function, 394  
 group, 387, 612  
 Hamilton's equations, 343  
 matrix, 384  
 Poisson bracket, 388, 397  
**System**  
 continuous, 568  
 discrete, 558  
 vector, 409, 410, 413  
**Tachyon**, 278  
**Tait-Bryan angles**, 154  
**Tardyón**, 278  
**Taylor series**, 239  
 potential expansion, 482  
**Temperature**, definition, 85  
**Tensor**, 188–191  
 alternating, 169  
 Cartesian, 189  
 definition, 293  
 first rank, 189  
 inertia, 191–198  
 isotropic of rank 3, 169  
 metric, 286  
 moment of inertia, 191–198  
 product, 294  
 properties, 188  
 rank, 293  
 second rank, 188  
 slots, 293  
 unit, 190  
 wedge product, 295  
 zero rank, 189  
**Thermodynamics**, 336  
**Thomas**  
 frequency, 285  
 precession, 282, 330  
**Three body problem**, 121–126,  
 617  
 Euler solution, 122  
 Lagrange solution, 123  
 restricted, 124, 133  
**Threshold energy**, 302–305  
**Time dilation**, 279  
**Timelike**, 278  
**Top**  
 Euler equations, 210  
 fast, 215, 221  
 heavy symmetrical, 200, 208,  
 482  
 with one point fixed, 208  
 motion, 208, 212  
 sleeping, 221  
 symmetric, 618  
 tippie, 221  
 uniform, 221  
**Topological dimension**, 518  
**Torque**, 2  
 critical, 266  
 damping, 266  
 gravitational, 223  
 pendulum, 266  
**Torus**, 487, 492  
**Tour de force**, 407  
**Trace of similarity transformation**,  
 160  
**Transformation**  
 active sense, 143  
 canonical, 368–421  
 infinitesimal, 396  
 restricted, 371, 382  
 congruence, 245, 246,  
 252  
 equation, 13  
 extended canonical, 371  
 formal properties, 144  
 Galilean, 281  
 gauge, 595  
 generating function, 371  
 identity, 146, 156, 395  
 improper, 151, 168  
 infinitesimal, 165  
 canonical (I.T.C.), 396

- Transformation (*cont.*)  
 Legendre, 375, 549  
 examples, 375  
 linear, 187  
 Lorentz, 280  
 matrix, 144  
 elements, 140  
 operator, 142  
 orthogonal, 139–150, 184  
 passive sense, 143  
 point, 31, 370, 422  
 principal axis, 241  
 proper, 151  
 restricted canonical, 371, 382  
 rigid body rotation, 139–155  
 scale, 370  
 similarity, 149, 158, 180, 189, 244  
 Transient, 260  
 Translational mode, 272  
 Triatomic molecule, 275  
 Triple cross product, 186  
 Turning angles, 213  
 Twin paradox, 285  
 Ultrarelativistic, 303  
 region, 308  
 Undetermined multipliers of Lagrange, 46, 363  
 Unitary matrix, 412  
 Unstable moment of inertia axis, 205  
 van der Pol  
 equation, 490  
 limit cycle, 491  
 Variable  
 canonical, 335  
 fast, 547  
 slow, 547  
 Variation, 354  
 $\delta$ -type, 38, 44  
 $\Delta$ -type, 357, 359  
 integral, 44  
 line integral, 35  
 Variational  
 Hamiltonian, 353  
 principle, 5, 34–43, 51  
 Vector  
 4-vector  
 energy, momentum, 295, 300, 301  
 photon momentum, 304  
 table, 287  
 velocity, 286–288  
 addition, 163  
 axial, 168  
 conserved, 104  
 covariant, 289  
 field, table, 287  
 first rank tensor, 189  
 flux density, 569  
 Minkowski space, 286  
 polar, 167  
 radius, 73  
 rate of change, 171–174  
 system, 409, 410, 413  
 tangent, 286, 326  
 Velocity  
 addition law, 282  
 angular, 172, 187  
 critical, 221  
 rigid body, 172  
 areal, 73  
 critical angular, 221  
 escape, 31  
 field, 588  
 four-, 286  
 generalized, 25, 319  
 Vibration  
 anharmonic, 255  
 forced, 259, 264  
 free, 250, 253  
 modes, 261  
 linear triatomic molecule, 253  
 longitudinal mode, 257  
 number of normal modes, 255  
 transverse mode, 257  
 Virial  
 Clausius, 84, 128  
 theorem, 83–86, 94, 472  
 Virtual  
 displacement, 16, 20  
 work principle, 17  
 Viscosity, 51, 265  
 Wavefunction, 613  
 Weak nuclear force, 299  
 Weber's electrodynamics, 367  
 Wedge product, 295, 296  
 Wheatstone bridge, 66  
 Witten and Sander diffusion model, 524  
 Wobble, Chandler, 208, 228  
 Work, 9  
 Yaw angle, 154, 603  
 Year, anomalistic, 131  
 Young's modulus, 559, 560  
 Zeeman effect, 232

