Chapter 7 Hamiltonian Mechanics

7.1 Hamilton's Equations

We saw in Chapter 1 that the Lagrange formulation of dynamics enjoyed many advantages in terms of the freedom of choice of coordinates, the ready recognition of conservation laws, and finally in the presentation of the laws of dynamics themselves in terms of the action principle

$$\delta \int_{t_1}^{t_2} L(q, \dot{q}, t) = 0. \tag{7.1}$$

In this Chapter, we will study yet another reformulation of the equations of motion of a system, whereby the n Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) = \frac{\partial L}{\partial q_k}, \quad k = 1, 2, \ldots, n, \tag{7.2}$$

for a system having n degrees of freedom, and which are second-order differential equations in the time variable t, are replaced by $2n$ first-order differential equations, known as *Hamilton's equations of motion*. The simplest way of accomplishing this end is to introduce n auxilliary variables r_k, where

$$\dot{q}_k = r_k, \quad k = 1, 2, \ldots, n. \tag{7.3}$$

Then the Lagrange equations (7.2) revert trivially to a set of n first order differential equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial r_k}\right) = \frac{\partial L}{\partial q_k}, \tag{7.4}$$

which, together with (7.3), make up a set of $2n$ first-order differential equations that are equivalent to the n original Lagrange equations (7.2) that are of second order. However, instead of using the r_k, it proves much more convenient to use the canonical momenta

$$p_k = \frac{\partial L}{\partial \dot{q}_k} = \frac{\partial L}{\partial r_k}, \tag{7.5}$$

defined in (1.72) of Chapter 1. Assuming that this relation can be turned "inside out" to solve for the r_k in terms of q_k and p_k, we can pass from the set of $2n$ variables (q_k, r_k) to the set (q_k, p_k).

Let us retrace the steps leading from the extremum condition (7.1) to the Lagrange equations, but where L is now considered as a function of q_k, and r_k as per (7.3), $L = L(q, r, t)$. Therefore the variation indicated in (7.1) has now to be carried out under the restriction that (7.3) must hold, or that

$$\delta \dot{q}_k - \delta r_k = 0, \tag{7.6}$$

in terms of the variations in $\delta \dot{q}_k = (d/dt)\delta q_k$ and δr_k. Therefore, the $2n$ variables q_k and r_k are not independent. But from the discussion of constraints in Section 1-6 of Chapter 1, we are already prepared to deal with this complication by introducing Lagrange multipliers λ_k, where

$$\sum_k \lambda_k(\dot{q}_k - r_k) = 0. \tag{7.7}$$

Adding the variation of this equation, i.e. the identity

$$\delta \sum_k \lambda_k(\dot{q}_k - r_k) = 0 \tag{7.8}$$

to L in (7.1), we find that

$$\delta \int_{t_1}^{t_2} [L + \sum_k \lambda_k(\dot{q}_k - r_k)]\, dt = 0, \tag{7.9}$$

where, as previously, $\delta \dot{q}_k = \frac{d}{dt}(\delta q_k)$. The variations δq_k and δr_k can now be considered to be independent, and the coefficients of each δq_k and δr_k must vanish separately. In particular, the coefficients of the δr_k determine the λ_k according to

$$\frac{\partial L}{\partial r_k} = \lambda_k. \tag{7.10}$$

Returning to (7.9) with this information, we discover that the extremum condition is equivalent to

$$\delta \int_{t_1}^{t_2} [L + \sum_k \frac{\partial L}{\partial r_k}(\dot{q}_k - r_k)]\, dt = 0, \tag{7.11}$$

or

$$\delta \int_{t_1}^{t_2} [L - \sum_k p_k \dot{q}_k + \sum_k p_k \dot{q}_k]\, dt = 0, \tag{7.12}$$

if we make use of $r_k = \dot{q}_k$ and the definition of p_k. The last relation may be written as

$$\delta \int_{t_1}^{t_2} [-H + \sum_k p_k \dot{q}_k]\, dt = 0 \tag{7.13}$$

in terms of the *Hamilton function*, or *Hamiltonian*

$$H = -L + \sum_k p_k \dot{q}_k. \tag{7.14}$$

We have met this expression previously in (1.78) of Chapter 1, there as a result of investigating the conservation properties of a system. We see from (7.13) that the principle of least action can be expressed in terms of H as well, so that H is indeed as fundamental as the Lagrange function for describing the dynamics of a system. We note that by agreement L was to be considered a function of q_k and r_k or equivalently q_k and p_k (courtesy of (7.5) in the above derivation). Consequently the same holds true for H: The Hamilton function is to be considered a function of the $2n$ *canonical variables* (q_k, p_k), i.e.

$$H = H(q, p, t) \tag{7.15}$$

using the shorthand notation $q = (q_1, q_2, \ldots, q_n)$, etc.

With this proviso in mind we can use the action principle in the form (7.13) to find the equations of motion for the variables q_k and p_k. Carrying out the indicated variations, one sees that

$$\int_{t_1}^{t_2} \sum_k [(-\frac{\partial H}{\partial q_k} - \dot{p}_k)\delta q_k + (-\frac{\partial H}{\partial p_k} + \dot{q}_k)\delta p_k] \, dt = 0, \tag{7.16}$$

after employing a partial integration to shift the time derivative of $\delta\dot{q}_k = (d/dt)\delta q_k$ onto its cofactor p_k, and the boundary conditions $\delta q_k(t_1) = \delta q_k(t_2) = 0$. Since the variations δq_k and δp_k are both arbitrary and independent, we conclude that

$$\dot{p}_k = -\frac{\partial H}{\partial q_k}; \quad \dot{q}_k = \frac{\partial H}{\partial p_k}, \quad k = 1, 2, \ldots, n. \tag{7.17}$$

This set of $2n$ *first order* differential equations constitute Hamilton's equations of motion for a system with n degrees of freedom. Note in passing that if a coordinate q_s is absent (cyclic) in L, it is also absent in H. Hence from the first equation,

$$\dot{p}_s = -\frac{\partial H}{\partial q_s} = 0, \quad \text{or} \quad p_s = \text{constant}, \tag{7.18}$$

or that each momentum that is canonical to a cyclic coordinate is conserved, as in (1.75) of Chapter 1. The Hamilton function H is itself conserved if it does not depend explicitly on t. We see this by calculating dH/dt,

$$\frac{dH}{dt} = \sum_k (\frac{\partial H}{\partial q_k}\dot{q}_k + \frac{\partial H}{\partial p_k}\dot{p}_k) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \tag{7.19}$$

and using Hamilton's equations. Hence, if

$$\frac{\partial H}{\partial t} = 0, \quad \text{then} \quad H = \text{constant}. \tag{7.20}$$

H may or may not represent the total energy, depending on the functional form of T and V in the Lagrange function, see discussion leading to (1.81) of Chapter 1: If L has the simple form $L = T - V$, with T being a homogeneous quadratic function of the $\dot{q}_k$, and V only a function of the q_k, then

$$H = T + V = E \tag{7.21}$$

is constant and equal to the total energy E of the system.

7-2 *Some Examples of Hamilton's Equations*

Experience has shown that it is usually as easy or as difficult to solve Hamilton's equations of motion as it is to solve Lagrange's equations for a given physical problem. The utility of Hamilton's equations actually lies in a different direction as will become clear presently. For the moment we content ourselves with a few examples to gain familiarity with Hamilton's method.

(i) *One-dimensional harmonic oscillator:* Calling the displacement of the particle away from equilibrium q instead of x, one has

$$L = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}m\omega_0^2 q^2 \tag{7.22}$$

if the mass m oscillates with frequency ω_0. The Hamilton function (7.14) becomes

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 q^2 \tag{7.23}$$

in this case, since

$$p = \frac{\partial L}{\partial \dot{q}} = m\dot{q} \tag{7.24}$$

can be used to eliminate $\dot{q}$ in favour of p. Hamilton's equations of motion for an oscillator are thus

$$\dot{p} = -m\omega_0^2 q; \quad \dot{q} = \frac{p}{m}. \tag{7.25}$$

The second of these just reconfirms the momentum-velocity relation (7.24). If we use this relation to eliminate p from the first equation, then

$$\ddot{q} + \omega_0^2 q = 0, \tag{7.26}$$

which is just the equation of motion for an oscillator.

(ii) *Central Motion*

Consider again the motion of a particle, mass m in a central potential field $V(r)$. In plane polar coordinates (r, θ) the Lagrange function is

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r), \tag{7.27}$$

leading to the canonical momenta

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}; \quad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}. \tag{7.28}$$

The Hamiltonian, which is also the total energy in this and the previous problem, reads

$$H = \frac{1}{2m}(p_r^2 + \frac{p_\theta^2}{r^2}) + V(r), \tag{7.29}$$

and the equations of motion are

$$\dot{p}_r = \frac{p_\theta^2}{mr^3} - \frac{\partial V}{\partial r}; \quad \dot{r} = \frac{p_r}{m} \tag{7.30}$$

and

$$\dot{p}_\theta = 0; \quad \dot{\theta} = \frac{p_\theta}{mr^2}. \tag{7.31}$$

The last set of relations confirms that the angular momentum $p_\theta = mr^2\dot{\theta}$ is conserved, while p_r and p_θ can be eliminated in the first partner of (7.30) to give the usual radial equation of motion

$$m(\ddot{r} - r\dot{\theta}^2) = F(r) \tag{7.32}$$

in a force-field $F(r) = -\partial V/\partial r$.

(iii) *Motion of a Charged Particle:*

The Lagrange function for this problem has been given before in (2.73) of Chapter 2, i.e.

$$L = \frac{1}{2}mv^2 - e\phi + e(\mathbf{v} \cdot \mathbf{A}), \tag{7.33}$$

where m is the mass of the charge e moving with velocity $\mathbf{v}$. If we locate e in cartesian coordinates at $\mathbf{r} = (x, y, z)$, then the canonical momentum $\mathbf{p} = (p_x, p_y, p_z)$ is given as

$$\mathbf{p} = m\mathbf{v} + e\mathbf{A}, \tag{7.34}$$

so that

$$H = \frac{1}{2m}(\mathbf{p} - e\mathbf{A})^2 + e\phi. \tag{7.35}$$

A straight-forward calculation, using

$$\dot{\mathbf{p}} = -\text{grad}_r H; \quad \mathbf{v} = \dot{r} = \text{grad}_p H \tag{7.36}$$

leads to the equation of motion established previously, i.e.

$$\frac{d}{dt}(m\mathbf{v}) = e\mathbf{E} + e(\mathbf{v} \times \mathbf{B}). \tag{7.37}$$

7-3 Canonical Transformations

We saw in Chapter 1 how a reformulation of Newton's second law of motion as Lagrange's equations led to a particularly efficient way of looking at dynamical problems. But the real bonus was the freedom we gained in the choice of coordinates that could be employed. This freedom was guaranteed in the derivation of the Lagrange equations from an action principle, and showed up explicitly in their invariance in form under *point transformations* of the type

$$q_k = f_k(q'_1, q'_2, \ldots, q'_n; t) \tag{7.38}$$

from one set of n coordinates q_k to another set q'_k. The particular advantage of Hamilton's formulation of the equations of motion lies in the fact that, in addition to retaining their form under point transformations like (7.38), they admit a much wider class of transformations for which this invariance also holds. This class of transformations, called *canonical* or *contact transformations* provides for the transformation of the old coordinates *and* momenta to new coordinates and momenta Q_k and P_k, via relations of the form

$$\begin{aligned} q_k &= f_k(Q_1, \ldots, Q_n; P_1, \ldots, P_n; t) \\ p_k &= g_k(Q_1, \ldots, Q_n; P_1, \ldots, P_n; t) \end{aligned} \tag{7.39}$$

such that the new canonical variables also satisfy Hamilton's equations

$$\dot{P}_k = -\frac{\partial K}{\partial Q_k}; \quad \dot{Q}_k = \frac{\partial K}{\partial P_k} \tag{7.40}$$

in terms of a *new* Hamilton function $K = K(Q, P; t)$. The construction of transformations like (7.39) that turn Hamilton's equations for the old variables into mirror relationships for the new variables seems like a tall order. Actually the problem is quite simple. The hint as to how one should view the problem comes from the variational principle in the form (7.13). There we indicated how Hamilton's equations followed from this principle. Conversely, one can show that if Hamilton's equations are valid, then (7.13) must necessarily hold. With this in mind it follows therefore that if (7.40) are to hold, K must satisfy the following variational principle

$$\delta \int_{t_1}^{t_2} [-K + \sum_k P_k \dot{Q}_k]\, dt = 0, \tag{7.41}$$

with $\delta Q(t_1) = \delta Q(t_2) = 0$. The discovery of the class of transformations that are canonical therefore boils down to the problem of introducing transformations which *automatically* guarantee that (7.41) will hold, provided that (7.13), i.e.

$$\delta \int_{t_1}^{t_2} [-H + \sum_k p_k \dot{q}_k]\, dt = 0 \tag{7.42}$$

holds in the old variables, with the familiar boundary conditions $\delta q_k(t_1) = \delta q_k(t_2) = 0$. That (7.41) should be an automatic consequence of (7.42) does not mean that their integrands are equal, however. For we can add the total time derivative dG/dt of any arbitrary function G to the integrand of either equation without changing the value of the variation. This fact follows from the nature of the δ-variation. Let $G = G(q, Q, t)$ be some function of the old and new coordinates. Then

$$\int_{t_1}^{t_2} \frac{dG}{dt}\, dt = [\delta G]_{t_1}^{t_2} = \sum_n [\frac{\partial G}{\partial q_k} \delta q_k + \frac{\partial G}{\partial Q_k}\, dQ_k]_{t_1}^{t_2} = 0, \tag{7.43}$$

due to the boundary conditions on the δq_k and δQ_k at t_1 and t_2. Consequently, (7.41) will be a consequence of (7.42) provided that

$$-H(q, p, t) + \sum_k p_k \dot{q}_k = -K(Q, P, t) + \sum_k P_k \dot{Q}_k + \frac{dG}{dt}(q, Q, t). \tag{7.44}$$

The function G is called the *generating function* of the canonical transformation. Canonical transformations are thus characterized by specifying G instead of the relations (7.39). However, these relations tell us that G can only depend on any two of the sets of variables (q_k, p_k) and (Q_k, P_k), and possibly the time t. It proves to be convenient to specify G as depending on one of the following four sets of variables that are taken to be independent: (i) (q_k, Q_k), or (ii) (q_k, P_k), or (iii) (p_k, Q_k), or (iv) (p_k, P_k), so that G has "one foot in each camp" so as to speak. The missing variables in each case are then to be calculated in terms of any of these sets from a knowledge of the generating function itself. To see how this works out in practice, let us look at Case (i) $G = G(q, Q, t)$. Calculating the time derivative in (7.44) and equating coefficients of q_k and Q_k, one finds the set of equations

$$p_k = \frac{\partial G}{\partial q_k} \tag{7.45}$$

$$P_k = -\frac{\partial G}{\partial Q_k} \tag{7.46}$$

$$K = H + \frac{\partial G}{\partial t} \tag{7.47}$$

for each value of k. Equations (7.45) and (7.46) give $p_k = g'_k(q, Q, t)$ and $P_k = f'_k(q, Q, t)$. If we can "solve" the latter equation in the form $q_k = f_k(Q, P, t)$ and use this information to eliminate the q_k from the former, $p_k = g'_k(f_k(Q, P, t), Q, t) = g_k(Q, P, t)$, we have rendered the canonical transformation $(q_k, p_k) \to (Q_k, P_k)$ in the form (7.39).

Case (ii): Since q_k and P_k are considered independent, we must eliminate the $\sum_k P_k \dot{Q}_k$ in favour of $\sum_k \dot{P}_k Q_k$ in (7.44). This can be accomplished by writing

$$G(q, Q, t) = S(q, P, t) - \sum_k P_k Q_k, \tag{7.48}$$

so that

$$\frac{dG}{dt} = \sum_k \frac{\partial S}{\partial q_k}\dot{q}_k + \sum_k \frac{\partial S}{\partial P_k}\dot{P}_k + \frac{\partial S}{\partial t} - \sum_k \dot{P}_k Q_k - \sum_k P_k \dot{Q}_k. \tag{7.49}$$

Using this information to (7.44) once more, leads to another set of relations in terms of the function S,

$$p_k = \frac{\partial S}{\partial q_k} \tag{7.50}$$

$$Q_k = \frac{\partial S}{\partial P_k} \tag{7.51}$$

$$K = H + \frac{\partial S}{\partial t}. \tag{7.52}$$

Once more, the first two equations can be solved in principle to provide the (q_k, p_k) as functions of the (Q_k, P_k).

The relation (7.48) is an example of a *Legendre transformation* that is so useful in shifting to new independent variables in thermodynamics. Cases (iii) and (iv) above can be dealt with in like manner via the Legendre tranformations $G = T(p, Q, t) + \sum_k p_k q_k$ for (iii) or $G = U(p, P, t) + \sum(p_k q_k - P_k Q_k)$ for (iv) respectively. The dependent variables are given by

$$q_k = -\frac{\partial T}{\partial p_k} \tag{7.53}$$

$$P_k = -\frac{\partial T}{\partial Q_k} \tag{7.54}$$

$$K = H + \frac{\partial T}{\partial t} \tag{7.55}$$

in case (iii), $T = T(p, Q, t)$, and

$$q_k = -\frac{\partial U}{\partial p_k} \tag{7.56}$$

$$Q_k = \frac{\partial U}{\partial P_k} \tag{7.57}$$

$$K = H + \frac{\partial U}{\partial t} \tag{7.58}$$

in case (iv), $U = U(p, P, t)$.

Some examples of familiar transformations "in canonical clothing" are useful to bear in mind. The identity transformation, $q_k = Q_k$, and $p_k = P_k$ is obviously generated by setting $G = 0$. In terms of a type (ii) or type (iii) generating function, this means that either of

$$S = \sum_k q_k P_k, \quad \text{or} \quad T = \sum_k p_k Q_k \tag{7.59}$$

generate an identity transformation, as may be verified by direct calculation. The "interchange of names" transformation $q_k = -P_k$ and $p_k = Q_k$ is obviously canonical. It is generated by either of

$$G = \sum_k q_k Q_k, \quad \text{or} \quad U = \sum_k p_k P_k. \tag{7.60}$$

These last two examples show rather clearly that which variable is termed a "coordinate" and which a "momentum" is a matter of semantics in Hamiltonian mechanics, and the term "canonical variables" for each pair (q_k, p_k) is preferable. However, it is well to bear in mind that this semantic freedom has to be tempered by the realization that the physical significance of the canonical variables will eventually have to be identified in making actual calculations.

The next transformation in ascending order of complexity, is the point transformation between coordinates,

$$q_k = f_k(Q_1, Q_2, \ldots, Q_n; t), \quad \text{or} \quad Q_k = F_k(q_1, q_2, \ldots, q_n; t). \tag{7.61}$$

The first one is generated directly by

$$T = -\sum_l f_l(Q_1, Q_2, \ldots, Q_n; t) p_l, \tag{7.62}$$

or indirectly by

$$S = \sum_l F_l(q_1, q_2, \ldots, q_n; t) P_l, \tag{7.63}$$

which generating function also gives the second form of the point transformation directly. Writing q and Q for the set of variables in f and F, one has

$$\begin{aligned} q_k &= f_k(Q, t) \\ P_k &= \sum_l \frac{\partial f_l}{\partial Q_k} p_l \end{aligned} \tag{7.64}$$

or

$$\begin{aligned} Q_k &= F_k(q, t) \\ p_k &= \sum_l \frac{\partial F_l}{\partial q_k} P_l \end{aligned} \tag{7.65}$$

upon using (7.53) -(7.55) or (7.50) -(**??**). One can show directly that each of these transformations is canonical (see Problems). In a similar fashion the choices $S = \sum_l q_l g_l(P, t)$, or $T = -\sum_l G_l(p, t) Q_l$ generate "point" transformations between momenta, $p_k = g_k(P, t)$ or $P_k = G_k(p, t)$, while $G = \sum_l q_l W_l(Q, t)$ and $U = -\sum_l p_l W_l(P, t)$ would produce "mixed" point transformations like $p_k = W_k(Q, t)$, or $q_k = W_k(P, t)$ respectively. None of these results are really surprizing when one remembers that the q's and p's are always treated on an equal footing in Hamilton theory. Instead of arbitrarily constructing more complicated canonical transformations at will, we rather ask at this stage whether there are perhaps canonical transformations that have a particular significance in dynamics. This question is examined in the following pages.

7-4 Special Canonical Transformations and Hamilton-Jacobi Theory

Our considerations thus far have ignored the fact that the Hamilton function itself is altered to

$$K = H + \frac{\partial G}{\partial t} \tag{7.66}$$

under a canonical transformation of variables. We have also seen that G may be replaced by any one of the functions S, T or U. Since K is to be a function of the new variables (Q_k, P_k), a knowledge of the associated canonical transformation in the guise of (7.39) is a prerequisite in order to eliminate the old canonical variables on the right hand side of (7.66) in favour of the new. However, there is another way of looking at (7.66): It tells us that we may "fiddle" with the functional dependence and/or value of the new Hamiltonian at will by changing the function G. Therefore, the idea presents itself of looking for those generating functions (and thus canonical transformations) *that render the new Hamilton function simpler than the old one*, when looked at from the point of view of solving Hamilton's equations. Thus (7.66) is now to be regarded *as a condition determining* G (or whatever alternate generating function is considered), rather than a prescription for calculating K.

The question as to what is the "simplest" problem to solve in terms of the new canonical variables, opens up a host of possibilities. Perhaps the simplest imaginable dynamical problem is no problem at all. $K = 0$! By (7.40) this would mean that *all* the new canonical variables are constant in time, $\dot{P}_k = 0$ and $\dot{Q}_k = 0$ for all k, or

$$P_k = \alpha_k, \quad Q_k = \beta_k, \tag{7.67}$$

where α_k and β_k are constants. The generating function that accomplishes this transformation is determined by (7.66) with K set equal to zero, i.e.

$$H(q, p, t) + \frac{\partial G}{\partial t} = 0. \tag{7.68}$$

Equation (7.68) is called the *Hamilton-Jacobi* equation. It determines the generating function that performs the canonical transformation to constant canonical variables[76]. Once G (or S, or T, or U) has been determined, one can construct from it the associated canonical transformation

$$q_k = f_k(\beta_1, \ldots, \beta_n; \alpha_1, \ldots, \alpha_n; t) \tag{7.69}$$

$$p_k = g_k(\beta_1, \ldots, \beta_n; \alpha_1, \ldots, \alpha_n; t) \tag{7.70}$$

in terms of the $2n$ constants α_k and β_k and the time t. But at $t = 0$ (or any other convenient time instant), (7.69) and (7.70) allow one to relate the $2n$ constants (α_k, β_k) to the $2n$ initial values of the canonical variables (q_k, p_k), and thus allow one to calculate the q_k and p_k at any subsequent time in terms of their initial values, i.e. to solve the equations of motion *in the original variables*. Consequently, in solving the Hamilton-Jacobi equation, one is at the same time solving the associated dynamical problem.

[76] Sometimes the name Hamilton-Jacobi equation is reserved for this equation only when $S = S(q, P, t)$ is considered as the generating function. S itself is then called *Hamilton's principal function*.

To see how this works out, we must choose a particular set of independent variables to work with. The standard choice in discussing the Hamilton-Jacobi equation is to seek a generating function of the type $S = S(q, P, t)$, although any of the remaining three will do just as well. If we choose $G \to S$ as our unknown generating function in (7.68), then we may eliminate all the p's in $H(q, p, t)$ with the help of (7.50) to find, in detail,

$$H(q_1, \ldots, q_n; \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}; t) + \frac{\partial S}{\partial t} = 0, \tag{7.71}$$

the *Hamilton-Jacobi partial differential equation* in $(n+1)$ variables for S. We have already met this equation as (1.111) of Chapter 1. The function S, designated Hamilton's principal function, depends on the n q's and the time t, as well as the n constants, $P_k = \alpha_k$, that give the constant values of the new canonical momenta. The physical significance of these constant momenta is unspecified at this stage, and we will choose them later to suit our convenience. The important point to notice, however, is that none of the constants α_k in S are additive since S and S plus an arbitrary additive constant are both solutions of (7.71) (since this equation only involved derivatives of S). Now, the complete integral of (7.71) must contain as many constants of integration as there are variables, i.e. $(n+1)$ constants. According to the above observation one of these is necessarily additive and to be discarded. We are then left with a solution $S(q_1,\dots,q_n,\alpha_1,\dots,\alpha_n;t)$ involving n non-additive constants. The second equation (7.51) now enters to relate the n constant values of the new coordinates, $Q_k = \beta_k$ to the q_k and the time,

$$\beta_k = \frac{\partial S}{\partial \alpha_k} = F_k(q_1,\dots,q_n;\alpha_1,\dots,\alpha_n;t). \tag{7.72}$$

These n relations can in principle be turned "inside out"to give the old coordinates as a function of the α_k, β_k and the time t,

$$q_k = f_k(\alpha_1,\dots,\alpha_n;\beta_1,\dots,\beta_n;t). \tag{7.73}$$

Finally, with the help of this result and the relation

$$p_k = \frac{\partial S}{\partial q_k} \tag{7.74}$$

we can solve for the p_k in terms of the constants (α_k,β_k) and t in the form

$$p_k = g_k(\alpha_1,\dots,\alpha_n;\beta_1,\dots,\beta_n;t), \tag{7.75}$$

which brings us to the point envisaged in (7.69) and (7.70), and allows one to solve for the motion.

The procedure is best appreciated via an example. For the one-dimensional harmonic oscillator, H is given in (7.23). The Hamilton-Jacobi equation (7.71) for S reads in this instance

$$\frac{1}{2m}(\frac{\partial S}{\partial q})^2 + \frac{1}{2}m\omega_0^2 q^2 + \frac{\partial S}{\partial t} = 0, \tag{7.76}$$

where $S = S(q,\alpha,t)$ contains a single non-additive constant α equal to the new canonical momentum. This partial differential equation may be solved by the method of separation of variables. Writing $S = W(q) + f(t)$ in (7.76) and using (7.52), we find that $f(t) = -Et$, where $H = E$ is the constant total energy of the oscillator. Hence,

$$S = W(q) - Et, \tag{7.77}$$

where the function $W(q)$ satisfies the ordinary differential equation

$$\frac{1}{2m}(\frac{dW}{dq})^2 + \frac{1}{2}m\omega_0^2 q^2 = E, \tag{7.78}$$

with solution

$$W = \sqrt{2m}\int\sqrt{E - \frac{1}{2}m\omega_0^2 q^2}\,dq, \tag{7.79}$$

after leaving off the trivial additive constant of integration. While this expression is simple to integrate directly, we refrain from doing so until after using (7.72) to find q as a function of time and the constant new momentum α. We note that

$$S = W(q, E) - Et \tag{7.80}$$

contains a single non-additive constant, the total energy E in this case. What then is α? The point is simply that we are now at liberty to identify the new canonical momentum (which is constant) with E, so that

$$P = \alpha = E, \tag{7.81}$$

or any function of E. Since the last term in (7.80) contains E linearly, the above identification is a natural choice. Then (7.72) reads

$$\beta = \frac{\partial S}{\partial E} = \frac{1}{\omega_0}\sin^{-1}\{\sqrt{\frac{m\omega_0^2}{2E}}q\} - t, \tag{7.82}$$

or

$$q = \sqrt{\frac{2E}{m\omega_0^2}}\sin\omega_0[(t+\beta)]. \tag{7.83}$$

The remaining relation, $p = \partial S/\partial q$ at $t = 0$, relates the canonical variables $(Q, P) = (\beta, E)$ to the initial momentum. But since we know $p = m\dot{q}$ from Hamilton's equations, it is simpler to calculate p directly from (7.83) as

$$p = \sqrt{2mE}\cos\omega_0(t+\beta). \tag{7.84}$$

Now both β and E may be determined in terms of the initial values of q and p at $t = 0$, say. Thus S given by (7.79) and (7.80) generates a canonical transformation to new canonical variables that are respectively the time $-\beta$ when $q(t)$ vanishes and the total energy E of the oscillator.

7-5 Hamilton-Jacobi Theory for Conservative Systems

For systems where the Hamilton function is conserved, $H =$ constant, it is always possible to effect the separation of the time variable as in (7.77). We only consider such systems from now on where H, when conserved, also represents the total energy, $H = E$. Then, in view of (7.80) we may write generally

$$S = W(q_1, \ldots, q_n; E, \alpha_2, \ldots, \alpha_n) - Et \tag{7.85}$$

for a system with n degrees of freedom. Here we have identified the "first" new canonical momentum α_1 with E, as we are at liberty to do. Then W satisfies the following partial differential equation if S is to satisfy the Hamilton-Jacobi equation,

$$H(q_1, \ldots, q_n; \frac{\partial W}{\partial q_1}, \ldots, \frac{\partial W}{\partial q_n}) = E. \tag{7.86}$$

The function W, which depends on the n coordinates q_k, and n constants $(E, \alpha_2, \ldots, \alpha_n)$ is called *Hamilton's characteristic function* to distinguish it from S. We also met (7.86) in Chapter 1 as (1.112), with W there called S_0 for uniformity of notation. Once (7.86) has been solved for W, we return to (7.72) to learn that $\beta_1 = \partial S/\partial E = \partial W/\partial E - t$, or that

$$t + \beta_1 = \frac{\partial W}{\partial E} \tag{7.87}$$

and

$$\beta_k = \frac{\partial S}{\partial \alpha_k} = \frac{\partial W}{\partial \alpha_k}, \quad k = 2, 3, \ldots, n. \tag{7.88}$$

The second set of relations is particularly interesting since the time t appears nowhere explicitly. Thinking in terms of a system with two degrees of freedom for simplicity, (7.88) reads in detail

$$\beta = \frac{\partial W}{\partial \alpha} = F(q_1, q_2; E, \alpha) \tag{7.89}$$

if we call $\beta_2 = \beta$ and $\alpha_2 = \alpha$. This relation determines the relation between q_1 and q_2 at all times for given values of α, β and E, i.e. it determines the *orbit*. By contrast, (7.87) determines the transit time of the system from $-\beta_1$ to t, as we already observed in (1.109) of Chapter 1. A combination of all n relations (two in number for two degrees of freedom) give both the "orbit" and the q_k as a function of time for specified initial conditions.

(i) <u>Central motion in two dimensions</u>

Again, the procedure is best appreciated via an example. We return to the case of particle motion in a central field $V(r)$ and write down (7.86) in plane polar coordinates, $q_1 = r$ and $q_2 = \theta$. Then, in view of (7.29), one has

$$\frac{1}{2m}(\frac{\partial W}{\partial r})^2 + \frac{1}{2mr^2}(\frac{\partial W}{\partial \theta})^2 + V(r) = E. \tag{7.90}$$

However, we know that, in addition to E, the angular momentum $p_\theta = mr^2\dot{\theta}$ canonical to θ is also conserved. Being constant, p_θ, which is one of the old canonical coordinates, can serve just as well as one of the new canonical momenta, $\alpha_2 = p_\theta$, say. We give expression to this fact by separating W into a sum of a function $f(r)$ of r and an identity generator in θ, thus

$$W = f(r) + \theta\alpha_2. \tag{7.91}$$

It follows at once that the new momentum $P_\theta = \alpha_2$ is identical with p_θ since $p_\theta = \partial W/\partial\theta = \alpha_2$. The function $f(r)$ is determined by

$$\frac{1}{2m}(\frac{df}{dr})^2 = E - U(r), \tag{7.92}$$

where $U = V(r) + p_\theta^2/2mr^2$ is the effective potential introduced previously in Chapter 2. Hence,

$$W = \theta p_\theta + \sqrt{2m}\int\sqrt{E-U}\,dr, \tag{7.93}$$

if we take the positive square root in evaluating f. Application of (7.88) with $\alpha_2 = p_\theta$ and $\beta_2 = \theta_0$ gives back (2.22) of Chapter 2 again,

$$\theta_0 = \frac{\partial W}{\partial p_\theta} = \theta - \int_{r_{\min}}^{r}\frac{p_\theta}{\sqrt{2m(E-U)}}\frac{dr}{r^2}, \tag{7.94}$$

if we identify θ_0 as the angle where $r = r_{\min}$, the closest approach distance. The time dependence of r is provided by the companion equation, (7.87). Calling $t_0 = -\beta_1$, the time when $r = r_{\min}$, one has

$$t - t_0 = \frac{\partial W}{\partial E} = \int_{r_{\min}}^{r}\frac{dr}{\sqrt{\frac{2}{m}(E-U)}}, \tag{7.95}$$

which is just (2.23). We have thus extracted the two relevant equations for describing central orbits from Hamilton's principal function in a rather systematic way. However, there is a price for this convenience: We must be able to solve (7.86) for W. The trick that was used in solving the central field problem by writing W as a *sum* of functions, with the coordinates distributed each to a function, is called "separating" variables.[77] The procedure is certainly successful whenever all the old canonical momenta bar one are constant. Examples of this kind are central-field motion in two or three dimensions, or the motion of a top under gravity. But this proviso is too restrictive; however we do not go into further detail and refer the interested reader to further literature research.

(ii) A non-separable problem

Before leaving the practical applications of Hamilton-Jacobi theory, it is well to point out that all is not necessarily lost if the Hamilton-Jacobi equation fails to separate. Often the resulting non-separable partial differential equation can still be solved by appealing to the vast literature on the theory of partial differential equations of the Hamilton-Jacobi type for guidance. An example of this nature is provided by the problem of the motion of a charge e in crossed electric and magnetic fields **F** and **B**. We use cartesian coordinates (x, y) to locate the charge and point $\mathbf{F} = F\hat{\mathbf{y}}$ along y and $\mathbf{B} = B\hat{\mathbf{z}}$ along z so that the motion is confined to the x, y plane. This electromagnetic field is described by scalar and vector potentials

$$\phi = -Fy, \quad \mathbf{A} = \frac{1}{2}B[-y\hat{\mathbf{x}} + x\hat{\mathbf{y}}] \tag{7.96}$$

[77] It is useful to note that the separation of variables for the corresponding quantum mechanical problem involves choosing a *product*, rather than a sum to represent the unknown function. The reason for this difference becomes clear if we recall the connection between the solutions of the wave equation in the quasi-classical limit $\hbar \to 0$ and Hamilton's characteristic function W ($\hbar$ = Planck's constant divided by 2π). Writing $\psi \sim \exp(iW/\hbar)$ for the wave function, one finds that the Schroedinger energy operator gives

$$[-\frac{\hbar^2}{2m}\nabla^2 + V]\psi \simeq [\frac{1}{2m}(\nabla W)^2 + V] \times \exp(\frac{i}{\hbar}W) = E\exp(\frac{i}{\hbar}W)$$

in the limit $\hbar \to 0$. Thus W is a solution of (7.86) for the corresponding classical problem. Now, the quantum mechanical solution $\psi(t,\theta)$ for the central field problem in two dimensions separates as $\psi(r,\theta) = F(r)\exp[\frac{ip_\theta}{\hbar}\theta] \sim \exp[\frac{i}{\hbar}f(r)]\exp[\frac{ip_\theta}{\hbar}\theta] = \exp[\frac{i}{\hbar}(f(r) + \theta p_\theta)]$ showing the genesis of $W = f + \theta p_\theta$ quite clearly.

so that (7.86) reads

$$\frac{1}{2m}[(\frac{\partial W}{\partial x}+\frac{1}{2}eBy)^2+(\frac{\partial W}{\partial y}-\frac{1}{2}eBx)^2]-eFy=E, \tag{7.97}$$

with the help of the expression (7.35) for H. This equation is not separable. However, by using the method of Charpit, one can show that

$$\frac{\partial W}{\partial x}=\frac{1}{2}eBy+\alpha, \tag{7.98}$$

where α is a constant of integration [78]. Combining this result with the original partial differential equation, we can solve for $\partial W/\partial y$ and thus construct the differential $dW=(\partial W/\partial x)dx+(\partial W/\partial y)dy$. Hence,

[78] A. R. Forsyth, *ibid.*, pp420.

$$W=\int\sqrt{2m(E+eFy)-(eBy+\alpha)^2}\,dy+\alpha x+\frac{1}{2}eBxy. \tag{7.99}$$

We take E and α as the new canonical momenta. Then (7.87) and (7.88) read

$$\beta_1+t=\int_{y_0}^{y(t)}\frac{dy}{\sqrt{\frac{2}{m}(E+eFy)-(\frac{eB}{m}y+\frac{\alpha}{m})^2}} \tag{7.100}$$

and

$$\beta_2=x-\int_{y_0}^{y(x)}\frac{(\frac{eB}{m}y+\frac{\alpha}{m})}{\sqrt{\frac{2}{m}(E+eFy)-(\frac{eB}{m}y+\frac{\alpha}{m})^2}}\,dy, \tag{7.101}$$

where $y=y_0$ at $t=-\beta_1$, or $x=\beta_2$. Integration of these expressions is elementary and gives the path of the charge as a *trochoid* in general. In the special case that the charge starts out from rest at the origin at $t=0$, one has both $E=0$ and $\alpha=0$ so that

$$\begin{aligned} y(t) &= \frac{eF}{\omega B}(1-\cos\omega t),\\ x(t) &= = \frac{eF}{\omega B}(\omega t-\sin\omega t), \quad \omega=\frac{eB}{m}, \end{aligned} \tag{7.102}$$

which are the parametric equations of a cycloid. Note that ω is just twice the Larmor frequency $\Omega=eB/2m$.

(iii) Motion in a non-inertial frame

As a final example of the utility of canonical transformations, we rederive the result given in (2.100) of Chapter 2 for the equation of motion of a particle relative to a rotating frame of reference. Let Σ' refer to a set of axes $Ox'y'z'$ that are rotating with angular velocity $\Omega=\Omega(t)$ about an axis passing through the common origin of $Ox'y'z'$ and a set of axes $Oxyz$ that constitute an inertial frame Σ. The motion of a particle in Σ is described by the Hamiltonian

$$H=\frac{p^2}{2m}+V(r), \tag{7.103}$$

leading to the equation of motion $\dot{p} = -\text{grad}V = \mathbf{F}$. The coordinates $\mathbf{r}$ and $\mathbf{r}'$ of the particle in Σ and Σ' are related by the point transformation

$$\mathbf{r}' = A(t)\mathbf{r}, \tag{7.104}$$

representing the rotation of Σ' relative to Σ. A is the operator that performs this rotation, and is represented by an orthogonal matrix $AA^T = A^TA = I$ as we already discussed in Chapter 3. The elements $A_{ij} = A_{ij}(t)$ are time dependent. This has been indicated by the blanket notation $A(t)$. The generator of the point transformation $\mathbf{r}' = A(t)\mathbf{r}$ is

$$S = \mathbf{p}' \cdot (A(t)\mathbf{r}) = (A^T(t)\mathbf{p}') \cdot \mathbf{r}, \tag{7.105}$$

as may readily be proven from matrix algebra. Here $\mathbf{p}'$ is the new canonical momentum that goes along with $\mathbf{r}'$. We have

$$\mathbf{p} = \text{grad}_{\mathbf{r}} S = A^T(t)\mathbf{p}'. \tag{7.106}$$

The new Hamiltonian determining the motion in Σ' is given by (**??**), or

$$K = \frac{p'^2}{2m} + U(\mathbf{r}') + \mathbf{p}' \cdot (\frac{\partial A}{\partial t} A^T(t))\mathbf{r}', \tag{7.107}$$

after using (7.106), and the inverse transformation $\mathbf{r} = A^T(t)\mathbf{r}'$ to eliminate the old canonical variables, and calling $V(\mathbf{r}) = V(A^T\mathbf{r}') = U(\mathbf{r}')$. The effect of the operation $(\partial A/\partial t)A^T(t)$ follows on examining the effect of $A(t)$ on the position vector $\mathbf{R}_p(0)$ at $t = 0$ of a point P rigidly attached to Σ'. At time t this point is moved to $R_p(t) = A(t)R_p(0)$ by $A(t)$. The change in position of P in time dt is therefore

$$\mathbf{R}_p(t + dt) - \mathbf{R}_0(t) \simeq dt\frac{\partial A}{\partial t}\mathbf{R}_p(0) = dt\frac{\partial A}{\partial t}A^T(t)\mathbf{R}_p(t). \tag{7.108}$$

But this shift also equals $-dt(\Omega \times \mathbf{R}_p(t))$, where Ω is the angular velocity of Σ' at time t. Notice the *minus* sign. This is necessary because as we have seen in Chapter 3, $A(t)$ rotates the vector $\mathbf{R}_p(0)$ *in the opposite sense* to the rotation of axes envisaged previously. Thus, since

$$\frac{\partial A}{\partial t}A^T(t)\mathbf{R}_p(t) = -\Omega \times \mathbf{R}_p(t) \tag{7.109}$$

holds for any $\mathbf{R}_p$ at any instant of time, we conclude that $(\partial A/\partial t)A^T(t)$ is equivalent to the "operation" $-\Omega\times$, i.e.

$$\frac{\partial A}{\partial t}A^T \to -\Omega \times . \tag{7.110}$$

Employing this result, we finally have

$$K = \frac{p'^2}{2m} + U(\mathbf{r}') - \mathbf{p}' \cdot (\Omega \times \mathbf{r}'). \tag{7.111}$$

The Hamilton equations in the new variables $\mathbf{p}'$ and $\mathbf{r}'$ become

$$\dot{\mathbf{p}}' = -\text{grad}_{\mathbf{r}'} K = \mathbf{F} + \mathbf{p}' \times \Omega \tag{7.112}$$

and

$$\dot{\mathbf{r}}' = \mathbf{v}' = \text{grad}_{\mathbf{p}'} K = \frac{1}{m}\mathbf{p}' + \mathbf{r}' \times \Omega. \tag{7.113}$$

We solve for the canonical momentum $\mathbf{p}'$ from the above equation, $\mathbf{p}' = m\mathbf{v}' + m(\Omega \times \mathbf{r}')$, and find

$$\frac{d}{dt}(m\mathbf{v}') = \mathbf{F} + 2(m\mathbf{v}' \times \Omega) + \Omega \times (m\mathbf{r}' \times \Omega) + (m\mathbf{r}' \times \dot{\Omega}), \tag{7.114}$$

as in (2.100) of Chapter 2.

7-6 Periodic Systems and Action-Angle Variables

Periodic motion has, as we have seen, a rather special place in mechanics. We have devoted considerable space to discussing central orbits that are periodic, and a whole chapter to small vibrations. In preparation for a discussion of periodic systems in Hamilton-Jacobi theory, let us first point out that Hamilton's characteristic function W, which was originally introduced as the time-independent piece of S in (7.85), induces a canonical transformation of its own that is quite different from that introduced by S. Consider a conservative system, $H = E$. Let us enquire into the properties of the generating function, which we provisionally call $W = W(q_1, \ldots, q_n, P_1, \ldots, P_n)$, that transforms H into a new Hamilton function K which *only* depends on one half of the new canonical variables. Thus, all the Q_k (say) are cyclic, all the P_k are constants, α_k. From (7.50)- (**??**) one has that

$$p_k = \frac{\partial W}{\partial q_k} \tag{7.115}$$

$$Q_k = \frac{\partial W}{\partial \alpha_k} \tag{7.116}$$

$$H = H(\alpha_1, \ldots, \alpha_n). \tag{7.117}$$

For a conservative system, $H = E$, so that W is determined by (7.86) again,

$$H(q_1, \ldots, q_n; \frac{\partial W}{\partial q_1}, \ldots, \frac{\partial W}{\partial q_n}) = E. \tag{7.118}$$

The solution for W will contain the constant $\alpha_1 = E$, and $n-1$ additional constant of integration $\alpha_2, \ldots, \alpha_n$, none of which are additive. Thus, as previously,

$$W = W(q_1, \ldots, q_n; \alpha_1, \ldots, \alpha_n), \tag{7.119}$$

where the α_k are identified with the new constant momenta P_k.

We now return to systems performing a finite motion that are in addition described by a separable Hamilton-Jacobi equation. The form of the generating function (7.119) is then

$$W = \sum_{i=1}^{n} W_i(q_i, \alpha_1, \ldots, \alpha_n), \tag{7.120}$$

where each W_i only depends on a single coordinate q_i. The conjugate momenta, p_k,

$$p_k = \frac{\partial W_k}{\partial q_k} = p_k(q_k, \alpha_1, \ldots, \alpha_n) \tag{7.121}$$

are functions of the single canonical coordinate q_k and the set of constants $\alpha_1, \alpha_2, \ldots, \alpha_n$. This circumstance, which only occurs in separable systems, allows one to classify the motion of each (q_k, p_k) pair into one of two types. If we plot p_k as a function of q_k as given by (7.121), the result is a curve in (q, p)-space, or the *phase-space* of the (q_k, p_k) variables. Then (i) if this curve is a closed one, q_k must oscillate back and forth between two turning points as the point (q_k, p_k) moves around the curve. Thus q_k returns to its original value if (q_k, p_k) moves once around the closed curve, and q_k is said to have completed its full *cycle*. This motion in q_k is called a *vibration* (or *libration*) with q_k returning to its original value after each cycle. (ii) If on the other hand q_k advances by ϕ_k every time the system returns to its original state, the motion is termed *rotational*, with a full cycle in q_k being represented by its changing by ϕ_k.

Systems performing a vibrational or rotational motion in the sense described above have special properties that are best exhibited in terms of a special set of canonical variables called *action angle* variables. The action variables J_k are a new set of constants that replace the constants α_k in (7.115) - (7.121). They are defined as (the 2π is arbitrary, but convenient)

$$2\pi J_k = \oint p_k \, dq_k, \tag{7.122}$$

where the integral $\oint dq_k \ldots$ means that q_k is taken over its full cycle. Because of the separability displayed in (7.120), the J_k are indeed constants. Using the expression for p_k in (7.121),

$$J_k = \frac{1}{2\pi} \oint \frac{\partial W_k}{\partial q_k} \, dq_k = J_k(\alpha_1, \ldots, \alpha_n), \tag{7.123}$$

since the single q_k appearing in W_k integrates away. By inverting this equation we can elimiate the α_k in favour of the J_k in (7.115) - (7.121). Then the generating function becomes

$$W = \sum_{i=1}^{n} W_i(q_i, J_a, \ldots, J_n), \tag{7.124}$$

giving rise to the following equations for the associated canonical transformation it induces:

$$p_k = \frac{\partial W_k}{\partial q_k} = p_k(q_k; J_1, \ldots, J_n) \tag{7.125}$$

$$\psi_k = \frac{\partial W}{\partial J_k} = \psi_k(q_1, \ldots, q_m; J_1, \ldots, J_n) \tag{7.126}$$

$$K = H = E(J_1, \ldots, J_n). \tag{7.127}$$

We have designated by ψ_k the variable conjugate to J_k. It is called an *angle* variable for reasons which will become clear presently.

The equations of motion and their solutions, for the ψ_k and J_k are, respectively,

$$\dot{\psi}_k = \frac{\partial K}{\partial J_k} = \frac{\partial E}{\partial J_k} = \omega_k(J_1, \ldots, J_n) \tag{7.128}$$

with solutions

$$\psi_k(t) = \omega_k t + \psi_k(0), \quad k = 1, 2, \ldots, n \tag{7.129}$$

and

$$\dot{J}_k = -\frac{\partial K}{\partial \psi_k} = 0, \tag{7.130}$$

with solutions

$$J_k = \text{constant}, \quad k = 1, 2, \ldots, n. \tag{7.131}$$

The constants $\omega_k = \omega_k(J_1, \ldots, J_n)$ are called the *fundamental frequencies* of the separable system. We will also have more to say about them in a moment.

The canonical pairs (ϕ_k, J_k) are known as *action-angle* variables. They are specific to separable systems performing a periodic motion (in order to define a q_k cycle) and, as we have seen, have the property that each ψ_k increases linearly with time, while the J_k are constants. Geometrically, $2\pi J_k$ represents the "area" in phase space, enclosed by the orbit described by the point (q_k, p_k) in a vibrational motion, or the area under this curve for one cycle of q_k, if the motion is of the rotational type. Consequently,

$$J_k = J_k(q_1, \ldots, q_n; p_1, \ldots, p_n) \tag{7.132}$$

is a single-valued function of the old coordinates so that the J_k provide for n single-valued constants of motion of a separable system. In contrast with the J_k, the ψ_k increase by 2π for every cycle in q_k, the remaining q_l's $(q_l \neq q_k)$ and all the J's being held fixed. We see this by calculating the change $\Delta_l \psi_k$ due to a full cycle in q_l: From (7.126),

$$\begin{aligned} \Delta_l \psi_k &= \oint \frac{\partial^2 W}{\partial q_l \partial J_k}\, dq_l = \frac{\partial}{\partial J_k} \oint \frac{\partial W_l}{\partial q_l}\, dq_l \\ &= 2\pi \frac{\partial J_l}{\partial J_k} = 2\pi \delta_{kl}. \end{aligned} \tag{7.133}$$

Notice that (7.133) does *not* mean ψ_k changes by 2π during q_k's cycle in the actual motion, because other q's are also changing in this case. However, it does mean that each q_k must be a *periodic* function in the ψ_k, with a fundamental period of 2π. Thus, if we solve (7.125) or (7.126) for the q_k in terms of the action-angle variables,

$$q_k = q_k(\psi_1, \psi_2, \ldots; J_1, J_2, \ldots), \tag{7.134}$$

these solutions must have the property that (the n's are integers)

$$q_k(\psi_1 + 2\pi n_1, \psi_2 + 2\pi n_2, \ldots; J_1, J_2, \ldots) = q_k(\psi_1, \psi_2, \ldots; J_1, J_2, \ldots). \tag{7.135}$$

Hence any *single-valued* q_k can be written as a multiple *Fourier series*

$$q_k = \sum_{s_1=-\infty}^{\infty} \cdots \sum_{s_n=\infty}^{\infty} A^k_{s_1 s_2 \ldots s_n} \exp\left[i \sum_{l=1}^{n} s_l \psi_l\right], \tag{7.136}$$

where the s_l are integers running over the indicated ranges. Introducing the explicit form of the ψ_k from (7.129), one obtains the time-dependence

$$\exp\left[it \sum_{l=1}^{n} s_l \omega_l\right] \tag{7.137}$$

for each term of given $(s_1 s_2 \ldots s_n)$ in (7.136). Thus, each term of that sum oscillates periodically with a frequency $\sum_{l=1}^{n} s_l \omega_l$, but the sum of such terms is not periodic in general. One then speaks of a *conditionally* periodic motion of the $q_k = q_k(t)$ as a function of time. However, it often happens that two or more fundamental frequencies are commensurable, i.e. are connected by relations of the form

$$\frac{\omega_k}{n_k} = \frac{\omega_l}{n_l} = \cdots = \omega, \tag{7.138}$$

where the n_k and n_l are integers. Such frequencies are said to be *degenerate*. If all n frequencies are degenerate, i.e. if there are n entries on the left hand side of (7.138), the degeneracy is termed *complete*. Completely degenerate systems display a simple periodic motion in all their coordinates with a common period $T = 2\pi/\omega$. We see this immediately upon replacing the ω_k in (7.137) by the degenerate frequencies $n_k\omega$. Then,

$$\exp[i\omega t \sum_l s_l n_l] \tag{7.139}$$

will return to its initial value after a time $T = 2\pi/\omega$ for any set of s_l. Consequently, the q_k in (7.136) are all periodic, $q_k(t+T) = q_k(t)$ with a common period T.

The occurrence of degeneracies is more the rule than the exception for many dynamical systems of interest. One important example is the degeneracy in particle motion in a spherically symmetric potential $V(r)$.

This degeneracy, which is due to the central nature of $V(r)$ shows up in the dependence of the particle energy

$$E = E(J_\phi + J_\theta, J_r) \tag{7.140}$$

on the sum of the first two action variables, where

$$2\pi J_\phi = \oint p_\phi \, d\phi, \quad 2\pi J_\theta = \oint p_\theta \, d\theta, \quad 2\pi J_r = \oint p_r \, dr \tag{7.141}$$

in spherical polar coordinates. Hence ω_ϕ and ω_θ are degenerate, since

$$\omega_\phi = \frac{\partial E}{\partial J_\phi} = \frac{\partial E}{\partial J_\theta} = \omega_\theta. \tag{7.142}$$

Particularizing $V(r)$ still further to be the inverse field $V(r) = -\alpha/r$ introduces an additional degeneracy,

$$E = E(J_\phi + J_\theta + J_r) = -\frac{1}{2}\frac{m\alpha^2}{(J_\phi + J_\theta + J_r)^2} \tag{7.143}$$

that is peculiar to this potential. Then all three frequencies ω_ϕ, ω_θ and ω_r are degenerate, and equal to the common frequency of motion

$$\omega = \frac{\partial E}{\partial J_\phi} = \frac{\partial E}{\partial J_\theta} = \frac{\partial E}{\partial J_r} = \frac{2}{\alpha}\sqrt{\frac{2|E|^3}{m}} \tag{7.144}$$

for all three variables. The motion is truly periodic, the orbit being a closed one as we already know from Chapter 2.

We have thus found the orbiting frequency from a knowledge of the function $E = E(J_r, J_\theta, J_\phi)$ *without* first solving the equations of motion. This result shows the power of using action-angle variables, and is typical of the general situation: If the energy $E = E(J_1, \ldots, J_n)$ is known as a function of the action variables, then so are the fundamental frequencies ω_k. Often the function $E = E(J_1, \ldots, J_n)$ can be constructed without solving the equations of motion first. However, whether or not the resulting ω_k represent actual frequencies of motion has to be decided on the basis of what degeneracies are present, as we have already discussed.

The advent of such degeneracies has a further consequence. The number of (single-valued) constants of motion then exceeds the number n represented by the n action variables $J_k = J_k(q, p)$ found in (7.67) for a finite, separable motion. Referring by way of an example to central motion in three dimensions one has the energy E, the total angular momentum l, and its projection l_z along an arbitrary axis as the three single-valued integrals of motion. If, however, $V(r) = -\alpha/r$, leading to the additional degeneracy recorded above there is an additional integral of motion,

$$(\mathbf{v} \times \mathbf{l}) - \frac{\alpha}{r}\mathbf{r} \tag{7.145}$$

by construction. We see from (7.137) how such additional single-valued integrals of motion can arise. For, apart from the n constants J_k, the $n-1$ constructs

$$\psi_k \omega_l - \psi_l \omega_k \tag{7.146}$$

are also constant in time. However, they are not single-valued. But when $\omega_k/n_k = \omega_l/n_l = \omega$, the difference

$$\psi_k n_k - \psi_l n_l \tag{7.147}$$

only changes by multiples of 2π when the ψ_k change by 2π as per (7.133). Any trigonometric function of $\psi_k n_k - \psi_l n_l$ will therefore be an additional single-valued constant of motion.

The existence of more single-valued constants of motion than degrees of freedom in separable degenerate systems[79] also means that the choice of coordinates in which to represent the n action variables cannot be unique, i.e. that the associated Hamilton-Jacobi equation must be separable in more than one set of coordinates. For example, the Hamilton-Jacobi equation with $V(r) = -\alpha/r$ separates in both spherical polar and parabolic coordinates. The occurrence of degeneracies in $E(J_1, \ldots, J_n)$, as exemplified by (7.143) also suggest a means for their removal. In the case of any central motion we can dispose of one action variable (J_ϕ say, with its attendant angle variable ψ_ϕ) by using our prior knowledge that the orbit lies in a plane. Using polar coordinates (r, θ), in this plane, the energy will in general be a function of J_θ and J_r separately, $E = E(J_\theta, J_r)$. However, if $V(r) = -\alpha/r$, one has a further degeneracy, since

$$E = -\frac{1}{2}\frac{m\alpha^2}{(J_\theta + J_r)^2}. \tag{7.148}$$

If we now consider E in (7.148) to be the "old" Hamiltonian and introduce a second canonical transformation

$$G = (\psi_\theta - \psi_r)J' + \psi_r J \tag{7.149}$$

to new action-angle variables (ψ', J') and (ψ, J) where

$$J_\theta = \frac{\partial G}{\partial \psi_\theta} = J', \quad J_r = \frac{\partial G}{\partial \psi_r} = -J' + J \tag{7.150}$$

and

$$\psi' = \frac{\partial G}{\partial J'} = \psi_\theta - \psi_r, \quad \psi = \frac{\partial G}{\partial J} = \psi_r, \tag{7.151}$$

then the "new" Hamiltonian

$$J = E(J) = -\frac{m\alpha^2}{2J^2} \tag{7.152}$$

is only a function of J. As a result, the equations of motion for the new angle variables are

$$\dot{\psi}' = 0, \quad \dot{\psi} = \frac{\partial E}{\partial J} = \omega, \tag{7.153}$$

[79] The reader will recognise that much of this commentary has a counterpart in the properties of the associated Schroedinger wave equation, especially in view of what was said a few pages back about the connection between the solutions of the wave equation and the Hamilton-Jacobi equation. L.P. Eisenhart, Phys. Rev. **74**, 87 (1948), has identified potentials for which the Schroedinger equation is separable. The same considerations obviously apply for the separation of the Hamilton-Jacobi equation.

showing that $\psi' =$ constant, while ψ increases linearly with time as before,

$$\psi(t) = \omega t + \psi(0). \tag{7.154}$$

Referring back to (7.93), it is interesting to construct the canonical transformation from the old to the new variables explicitly. We have for ψ in particular,

$$\psi = \psi_r{}' = \frac{\partial W}{\partial J_r} = \omega \frac{\partial W}{\partial E} = \xi - e \sin \xi, \tag{7.155}$$

either by direct integration, or by comparison with (2.34) of Chapter 2, if we set $\psi = 0$ at the perihelion. The geometric constants of this orbit can also be expressed in terms of J and J'. Noting that $J' = J_\theta = p_\theta$, the constant angular momentum of the motion, one finds

$$a = \frac{J^2}{m\alpha}, \quad e = \sqrt{1 - \frac{J'^2}{J^2}} \tag{7.156}$$

for the semi-major axis and the eccentricity of the elliptic orbit (refer to (2.32) of Chapter 2).

7-7 Perturbations

The action-angle variables introduced in the previous section turn out to be very suitable variables in which to study the effect of perturbations of the Hamiltonian K in (7.127). We do not intend to present a systematic development of such *canonical perturbation theory*, on account of its complexity in relation to the sort of problem we want to discuss, but rather wish to illustrate some salient features by way of an example. For this purpose consider the motion of a particle in the central field $V(r) = -\alpha/r$, upon which a constant force field $\mathbf{F}$ has been superposed. The potential energy for the perturbed problem is thus

$$V(r) = -\frac{\alpha}{r} - (\mathbf{r} \cdot \mathbf{F}). \tag{7.157}$$

We wish to investigate how the motion in the inverse field, i.e. the unperturbed motion, is affected by the additional interaction $(\mathbf{r} \cdot \mathbf{F})$, assuming this to be "small". The philosophy of perturbation theory runs as follows: We know the geometric constants of the unperturbed elliptic orbit of the particle in terms of J and J'. The presence of a small perturbation will therefore have the effect of introducing a slow time dependence (also relative to the period of elliptic motion) into J and J'. Thus, the geometric constants (and orientation) of the unperturbed ellipse are expected to change slowly with time. The problem is therefore to find out how the J's vary with time. But since the variation in the shape and orientation of the ellipse is a slow one, we can as a first approximation replace $(\mathbf{r} \cdot \mathbf{F})$ by its time-average $\langle \mathbf{r} \cdot \mathbf{F} \rangle$ over one period of unperturbed motion. Now, the time-average $\langle \mathbf{r} \rangle$ of the position vector $\mathbf{r}$ of the particle relative to the center of force obviously lies along the major axis of the ellipse by symmetry. Its value is

$$\langle x \rangle = \frac{1}{r}\int_0^r r\cos\theta\, dt = \frac{\alpha}{2\pi}\int_0^{2\pi}(\cos\xi - e)(1 - e\cos\xi)\, d\xi = -\frac{3}{2}ae, \quad (7.158)$$

after appealing to (2.34) and (2.43) of Chapter 2. Hence, the *average* perturbing interaction is

$$H' = -\langle \mathbf{r}\cdot\mathbf{F}\rangle = \frac{3}{2}aeF\cos\psi', \quad (7.159)$$

where ψ' is the angle between $\mathbf{F}$ and the semi-major axis of the ellipse. If we now assume that $\mathbf{F}$ lies in the plane of the orbit, then ψ' is just the constant angle variable conjugate to J' that gives the orientation of the unperturbed ellipse relative to some arbitrary direction that we now choose to be along $\mathbf{F}$. We see this by noting that $\psi' = \partial W/\partial J'$, leads to (7.94) again, with $\psi' = \theta_0$. The perturbed Hamiltonian $K' = K + H'$ is thus

$$K' = E(J) + \frac{3}{2}F(\frac{J^2}{m\alpha})\sqrt{1 - \frac{J'^2}{J^2}}\cos\psi', \quad (7.160)$$

when expressed in the canonical variables J, J', and ψ'. K' determines the evolution in time of these variables to first order in F. We have

$$\dot{J} = -\frac{\partial K'}{\partial \psi} = 0 \quad (7.161)$$

$$\dot{J}' = -\frac{\partial K'}{\partial \psi'} = \frac{3aF}{2J}\sqrt{J^2 - J'^2}\sin\psi' \quad (7.162)$$

$$\dot{\psi}' = \frac{\partial K'}{\partial J'} = -\frac{3aF}{2J}\frac{J'}{\sqrt{J^2 - J'^2}}\cos\psi'. \quad (7.163)$$

The first equation shows that J, and therefore the semi-major axis a, is not changed by H'. We have therefore re-introduced a into (7.162) and (7.163), since it is constant in time. Noting that

$$\frac{d}{dt}(\sqrt{1 - J'^2/J^2}\cos\psi') = 0, \quad (7.164)$$

in view of (7.162) and (7.163), or that

$$(\sqrt{1 - J'^2/J^2}\cos\psi') = \text{constant} = e_0, \quad (7.165)$$

(e_0 is the eccentricity of the unperturbed motion), we can eliminate ψ' between (7.162) and (7.163) to find

$$\ddot{J}' + (\frac{3aF}{2J})^2 J' = 0. \quad (7.166)$$

Hence,

$$J'(t) = p_\theta \cos(\frac{3aF}{2J}t), \quad (7.167)$$

where p_θ is the unperturbed value of J'. The eccentricity $e(t)$ and orientation $\psi'(t)$ of the major axis oscillate with the same frequency

$$\omega' = \frac{3aF}{2J} \tag{7.168}$$

as $J'(t)$. From (7.156) and (7.165),

$$\begin{aligned} e(t) &= \sqrt{1-(1-e_0^2)\cos^2\omega' t} \\ \cos\psi'(t) &= \frac{e_0}{e(t)}. \end{aligned} \tag{7.169}$$

These results show that the effect of H' in (7.159) is to cause the semi-major axis of the ellipse to oscillate with frequency ω' about the direction of **F** while maintaining its original length. The eccentricity $e(t)$ waxes and wanes with the same frequency. It is interesting to note that $J'(t)$ changes sign periodically, indicating that the particle motion is reversed periodically (assuming that the perturbation theory holds for sufficiently long times for this to happen). The frequency of oscillation ω' is expected to be much slower than the frequency ω of elliptic motion. From (7.156),

$$\frac{1}{\omega'}\frac{\partial E}{\partial J} \simeq \frac{2m\alpha^2}{3aFJ^2} = \frac{4}{3}\frac{|E|}{aF}, \tag{7.170}$$

a ratio that is large if the size of the perturbing potential (as measured by aF) is small relative to the total energy in the elliptic motion.

The above calculations apply in particular to the case of a hydrogen atom placed in an external electric field (classical theory of the Stark effect). Actually, the Hamilton-Jacobi equation with $V(r)$ given by (7.157) is separable in parabolic coordinates, so that the problem is exactly solvable for arbitrary external electric fields. One can thus investigate the range of validity of our approximate solutions, as well as questions like whether the bounded elliptic motion can become unbounded if the external field is strong enough (see Problems).

7-8 Further Aspects of Canonical Transformations

We return to (7.52) and consider the case that the generating function $S = S(q, P)$ is (a) independent of time, and (b) differs infinitesimally from the identity transformation $\sum_k q_k P_k$ thus:

$$S = \sum_k q_k P_k + \eta f(q, P). \tag{7.171}$$

Here f is an arbitrary function, and η a small parameter characterizing the difference between S and the identity transformation. The equations of transformation read

$$p_k = \frac{\partial S}{\partial q_k} = P_k + \eta \frac{\partial f}{\partial q_k}$$
$$Q_k = \frac{\partial S}{\partial P_k} = q_k + \eta \frac{\partial f}{\partial P_k}. \qquad (7.172)$$

If η is small however, we obtain the differences

$$\delta p_k = P_k - p_k \simeq -\eta \frac{\partial}{\partial q_k} f(q,p)$$
$$\delta q_k = Q_k - q_k \simeq \eta \frac{\partial}{\partial p_k} f(q,p), \qquad (7.173)$$

correct to first order in η, upon replacing P_k by p_k everywhere in the last terms on the right of (7.172). The generator S is then said to generate the *infinitesimal canonical transformation* given in (7.173).

Consider now any function $F = F(q,p)$ of the canonical variables but not an explicit function of the time. If q_k and p_k are shifted by the amounts given by (7.173) in an infinitesimal canonical transformation, the change in F is

$$\delta F = F(q+\delta q, p+\delta p) = \eta \sum_k (\frac{\partial F}{\partial q_k}\frac{\partial f}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial f}{\partial q_k}). \qquad (7.174)$$

The combination of derivatives under the sum appears often enough to merit a name and a special symbol. The expression

$$\sum_k (\frac{\partial F}{\partial q_k}\frac{\partial f}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial f}{\partial q_k}) = [F,f]_{q,p} \qquad (7.175)$$

is called a *Poisson bracket*, and is denoted by the square bracket symbol on the right. Thus the change in F induced by the infinitesimal generator ηf appears as

$$\delta F = \eta [F,f]_{q,p}. \qquad (7.176)$$

We will discover later that the subscripts (q,p) are not necessary. Poisson brackets are invariant under canonical transformations and may thus be evaluated in any convenient set of canonical variables. Let us investigate some special cases of (7.176). Since we are allowed to choose both f and the infinitesimal parameter η, take $f = H(q,p)$ a time-independent Hamilton function of some system, and $\eta = \delta t$, an infinitesimal element of time. Then,

$$\delta F = \delta t [F,H], \qquad (7.177)$$

or, writing $\delta F = \dot{F}\delta t$ as $\delta t = 0$,

$$\dot{F} = [F,H]. \qquad (7.178)$$

The time rate of change of F equals its Poisson bracket with the Hamiltonian responsible for changing q_k and p_k by δq_k and δp_k in time δt. In fact, setting $F = p_k$ and q_k respectively one finds

$$\delta p_k = \dot{p}_k \delta t = \delta t[p_k, H] = -\delta t \frac{\partial H}{\partial q_k}$$
$$\delta q_k = \dot{q}_k \delta t = \delta t[q_k, H] = \delta t \frac{\partial H}{\partial p_k}. \quad (7.179)$$

In a very real sense then, the Hamilton function is the generator of the infinitesimal transformation that shifts q_k and p_k from their values at time t, to their values that obtain at time $t + \delta t$ during the actual motion in the time δt.

Two further views of (7.178) and its special case are important. Firstly, if $F = F(q, p)$ happens to be a constant of the motion, $\dot{F} = 0$, then

$$[F, H] = 0, \quad (7.180)$$

i.e. all constants of motion have a vanishing Poisson bracket with the Hamiltonian[80]. Secondly if we go back to (7.176) and calculate the change in H itself under the infinitesimal transformation ηf, then

$$\delta H = \eta[H, f]. \quad (7.181)$$

Comparing this with the preceeding equation, we may say: The generators f of those infinitesimal canonical transformations that leave the Hamiltonian invariant, $\delta H = 0$; are constants of the motion.

The symmetry properties of H, and the constants of motion of the dynamical system it describes, are brought to the fore once again. For example if H is translationally invariant, the total momentum is conserved. To see this in the context of the present discussion, consider an N-particle system with a translationally invariant H. Then $\delta H = 0$ upon shifting each particle position vector $\mathbf{r}_i$ by a small common amount $\delta \mathbf{a} = \delta \mathbf{r}_i$. The generator for this infinitesimal transformation is

$$\eta f = \sum_{i=1}^{N} \delta \mathbf{a} \cdot \mathbf{p}_i = \delta \mathbf{a} \cdot \mathbf{P}, \quad (7.182)$$

where $\mathbf{P}$ is the total momentum. If $\delta H = 0$ then by (7.181) the quantity $\delta \mathbf{a} \cdot \mathbf{P}$ is a constant of motion. Thus, $\mathbf{P}$ itself must be constant since $\delta \mathbf{a}$ is an arbitrary vector.

Likewise, a rotationally invariant Hamiltonian conserves the total angular momentum of the system it describes. A common rotation of all particle coordinates through an angle $\delta\theta$ about any axis $\hat{\mathbf{n}}$ shifts the coordinates and associated momenta by

$$\delta \mathbf{r}_i = \delta\theta(\hat{\mathbf{n}} \times \mathbf{r}_i)$$
$$\delta \mathbf{p}_i = \delta\theta(\hat{\mathbf{n}} \times \mathbf{p}_i).$$

[80] An identical statement holds in quantum mechanics if the Poisson bracket is replaced by the commutator bracket.

The generator for this rotation is found to be

$$\eta f = \delta\theta \sum_{i=1}^{N} \hat{\mathbf{n}} \cdot (\mathbf{r}_i \times \mathbf{p}_i) = \delta\theta(\hat{\mathbf{n}} \cdot \mathbf{L}), \tag{7.183}$$

where $\mathbf{L} = \sum_i (\mathbf{r}_i \times \mathbf{p}_i)$ is the total angular momentum. If H is left invariant by this rotation, then $\delta\theta(\hat{\mathbf{n}} \cdot \mathbf{L})$, and therefore $\mathbf{L}$ itself is a constant of motion.

7-9 *Liouville's Equation*

There is another way of looking at Hamilton's equations of motion,

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}, \quad k = 1, 2, \ldots, n, \tag{7.184}$$

that is associated with the idea of the *phase space* of the system. For, if instead of thinking of the $2n$ variables (q_k, p_k) and how each varies with time, we follow the motion of the single representative point with "co-ordinates" $(q_1, \ldots, q_n; p_1, p_2, \ldots, p_n)$ in the $2n$-dimensional space with axes labelled by the q's and p's, then Hamilton's equations give us the "velocity" components of this point directly. The $2n$-dimensional space spanned by the $2n$ canonical variables is called the *phase space* of the system, and the point $(q_1, \ldots, q_n; p_1, p_2, \ldots, p_n)$ occupied by the system at time t its representative point, or *system point* in phase space. As the system moves according to the dictates of its Hamilton function and initial conditions, its system point moves along some curve, called a phase path, in phase space. Notice that the initial conditions (specification of the q's and q's at some initial time) tells us where the system point starts out in phase space, while Hamilton's equations determine how it proceeds away from this initial configuration.

(i) Liouville's Theorem

One important property of phase space is that the volume element

$$d\Omega = dq_1 \ldots dq_n dp_1 \ldots dp_n, \tag{7.185}$$

and therefore also any finite volume $\int d\Omega$ of phase space is invariant under a canonical transformation of variables. For consider by way of illustration a system with one degree of freedom, and introduce new canonical variables (Q, P) in place of (q, p). Then

$$d\Omega = dqdp = JdQdP = Jd\Omega', \tag{7.186}$$

where

$$J = \frac{\partial(q, p)}{\partial(Q, P)} = [q, p]_{Q,P} \tag{7.187}$$

is the Jacobian for the transformation $[q, p] \to [Q, P]$. The anticipated invariance $d\Omega = dq\,dp = d\Omega'$ therefore amounts to showing that the

Jacobian of all canonical transformations is unity. This brings in the idea of the fundamental Poisson brackets $[q, p]$, $[q, q]$ and $[p, p]$ and raises the question of how Poisson brackets transform under canonical transformations. Consider any Poisson bracket $[F, G]_{q,p}$, where F and G are functions of the canonical variables q and p. We assert that

$$[F, G]_{q,p} = [F, G]_{Q,P}, \tag{7.188}$$

where (Q, P) are new canonical variables, i.e. that the Poisson bracket is invariant under canonical transformations, and that both sets of subscripts are unnecessary in (7.188). The proof of (7.188) involves straightforward but tedious algebra using the properties of canonical transformations (see Problems). A direct argument for proving the invariance has been suggested by Landau and Lifshitz[81]. If we suppose G is the Hamiltonian of some fictitious system then, in view of (7.178), the Poisson bracket $[F, G]_{q,p}$ gives the time rate of change of F. But this must be given equally by $[F, G]_{Q,P}$, since the time rate of change can only depend on the properties of the system, not the choice of coordinates. Therefore, (7.188) must be true. Returning to our problem, we compute the invariant values

[81] L. Landau and E. Lifshitz, *Mechanics*, (Addison-Wesley Inc., Cambridge Mass., 1960) p. 145.

$$[q, p]_{Q,P} = [q, p]_{q,p} = 1 \tag{7.189}$$

$$[q, q]_{Q,P} = [q, q]_{q,p} = 0 \tag{7.190}$$

$$[p, p]_{Q,P} = [p, p]_{q,p} = 0 \tag{7.191}$$

of the fundamental Poisson brackets. Consequently $J = 1$ in (7.186), confirming the invariance $dq\,dp = dQ\,dP$. The proof for the $2n$-dimensional volume element in (7.185) is similar (see Problems). Therefore, it also follows that

$$\int dq_1 \ldots dq_n dp_1 \ldots dp_n = \int dQ_1 \ldots dQ_n dP_1 \ldots dP_n, \tag{7.192}$$

where the integral is taken over some volume of phase space.

The invariance of the volume in phase space under a canonical transformation leads to the following conclusion, known as *Liouville's Theorem*. We consider a collection of system points in a given volume of phase space. Such a collection might arise for example from considering a large number of identical mechanical systems with differing starting conditions. As time progresses each system point moves in accordance with Hamilton's equations, so that the boundaries of the phase space under consideration moves as well. However, as we saw in (7.179), the change in each q_k and p_k during the actual motion is made up of a series of infinitesimal canonical transformations generated by the Hamiltonian.

Therefore the volume occupied by the system points under consideration cannot change, or

$$\int d\Omega = \text{constant}, \tag{7.193}$$

a result known as Liouville's theorem.

(ii) Liouville's Equation

We can follow up one consequence of Liouville's theorem a little further by introducing the concept of density $\rho(q_1, \ldots, q_n; p_1, \ldots, p_n; t)$ of system points in phase space. Then

$$\rho d\Omega \tag{7.194}$$

gives the number of system points in the volume element $d\Omega = dq_1 \ldots dq_n \times dp_1 \ldots dp_n$ at time t. As time marches on, the system points move in accordance with Hamilton's equations, to occupy a new volume $d\Omega'$ at time $t + \delta t$. However, the *number* of system points in $d\Omega$ at time t must be the same as the number in $d\Omega'$ at time $t + \delta t$ since no system points can leave or enter the volume under consideration. Now, each representative system point changes position and momentum by $\dot{q}_k \delta t$ and $\dot{p}_k \delta t$ in time δt. Therefore the conservation of the number of system points means that

$$\rho(q_1 + \dot{q}_t \delta t, \ldots; p_1 + \dot{p}_1 \delta t, \ldots; t + \delta t) d\Omega' = \rho(q_1, \ldots; p_q \ldots; t) d\Omega', \tag{7.195}$$

or, since $d\Omega = d\Omega'$ by Liouville's theorem, that

$$\sum_k (\dot{q}_k \frac{\partial \rho}{\partial q_k} + \dot{p}_k \frac{\partial \rho}{\partial p_k}) + \frac{\partial \rho}{\partial t} = 0. \tag{7.196}$$

We insert Hamilton's equation for q_k and p_k at this point to find Liouville's equation,

$$[\rho, H] + \frac{\partial \rho}{\partial t} = 0 \tag{7.197}$$

for the density of system points in phase space. Obviously, the left hand side of either equation gives the total time derivative of ρ, so that Liouville's equation may also be written simply as $\dot{\rho} = 0$.

Further developments along these lines lead into statistical mechanics, an aspect that lies beyond the purpose of this chapter. Nevertheless, we cannot resist presenting a problem, due to Max Born[82], that bears on the question of in what sense classical mechanics may be regarded as deterministic. It is obvious from the mathematical problem posed by Hamilton's equations that the values of the canonical variables q_k and p_k can be calculated *exactly* at any later time in terms of their initial values $q_k(0)$ and $p_k(0)$ at time $t = 0$ say. That is to say, Hamilton's equations are fully deterministic. However, this statement tacitly presupposes that such initial values, or starting conditions, are known with arbitrary precision. We now pose, with Max Born, the following question: Suppose the

[82] M. Born, Physikalische Blaetter **15**, 342 (1959).

starting conditions are known only imprecisely. What then can be said about q_k and p_k at later times? Consider the motion of a free particle in one dimension by way of an example. Its position and momentum at any time t are given by

$$q(t) = q_0 + \frac{p_0}{m}t, \quad p = p_0 \tag{7.198}$$

in terms of their initial values q_0 and p_0 at $t = 0$. If, however, these initial values are imprecise, we can still assign a *probability*

$$f(q_0, p_0)d\Omega_0 \tag{7.199}$$

to the occurrence of a given q_0 and p_0 in the phase volume $d\Omega_0 = dq_0 dp_0$ at $t = 0$. This is equivalent to considering N identical particles that start off with different values of (q_0, p_0) such that the density of system points at $t = 0$ is

$$\rho(q_0, p_0, 0) = Nf(q_0, p_0). \tag{7.200}$$

By Liouville's equation, this must also equal the density at all later times,

$$\rho(q, p, t) = \rho(q_0, p_0, 0), \tag{7.201}$$

or

$$f(q, p, t) = f(q_0, p_0). \tag{7.202}$$

Born considers the example where q_0 and p_0 have independent Gaussian distributions about their respective means, so that

$$f(q_0, p_0) = \frac{1}{2\pi\sigma\tau} \exp\{(\frac{q_0 - \bar{q}}{\sigma})^2 + (\frac{p_0 - \bar{p}}{\tau})^2\}, \tag{7.203}$$

where σ and τ measure the scatter in q_0 and p_0 about their means $\bar{q}$ and $\bar{p}$. At all later time, therefore

$$f(q, p, t) = f(q - \frac{p}{m}t, p) = \frac{1}{2\pi\sigma\tau} \exp\{(\frac{q - \bar{q} - pt/m}{\sigma})^2 + (\frac{p - \bar{p}}{\tau})^2\} \tag{7.204}$$

in view of (7.202) and (7.198). Notice that the q and p distributions interlock at later times, $f(q, p, t)$ no longer being a product of a function of q times a function of p. Curves of equal probability are therefore ellipses

$$(\frac{q - \bar{q} - pt/m}{\sigma})^2 + (\frac{p - \bar{p}}{\tau})^2 = C \tag{7.205}$$

that enclose an area in phase space within which $f(q, p, t) >$ a constant determined by C. We examine the representative case $C = 1$. Setting

$$q - \bar{q} - \frac{\bar{p}}{m}t = \sigma\xi, \quad p - \bar{p} = \tau\eta, \tag{7.206}$$

we obtain

$$(\xi - \alpha t\eta)^2 + \eta^2 = 1. \tag{7.207}$$

This is an ellipse in the variables ξ and η, referred to its center. Its semi-major axis a makes an angle

$$\chi = \frac{1}{2}\cot^{-1}(\frac{1}{2}\alpha t) \simeq \frac{1}{\alpha t}, \quad \alpha t \gg 1, \tag{7.208}$$

with the positive ξ-axis and has the value

$$a = \cot\chi \simeq \alpha t. \tag{7.209}$$

At $t = 0$, this ellipse reduces to a unit circle in the variables (ξ, η), the center of which lies at $(\bar{q}, \bar{p})$. The area, π, of this circle must by Liouville's theorem equal the area of the ellipse (7.207) at all later times. Therefore, the axes a and b of the ellipse are reciprocal,

$$b = \frac{1}{a} = \tan\chi \simeq \frac{1}{\alpha t}. \tag{7.210}$$

So, as time progresses, two things happen: From (7.207) we see that the center of the circle moves with the mean speed $\bar{v} = \bar{p}/m$ of the distribution, while the circle itself deforms into an ellipse that becomes more and more "emaciated" as its major axis dips down towards the ξ-axis, see Fig. 7.1.

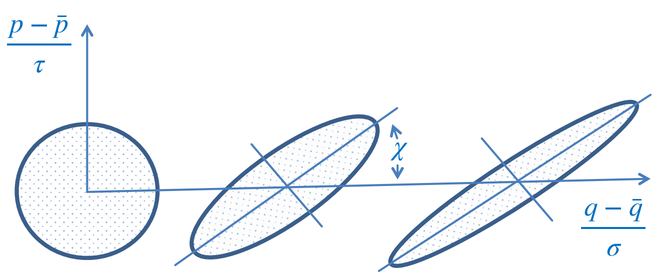

Figure 7.1: Area-preserving ellipses in the phase space of a free particle, shown at three equal intervals of time. System points inside the circular boundary at $t = 0$ remain inside the corresponding elliptic boundary as time increases.

Thus, the definition of the one canonical variable, p gets sharper at the expense of the other, q, a result we can understand by noting again that $f(q, p, t)$ is *not* a simple product of q and p distributions for $t > 0$,

The sharpening up of one canonical variable at the expense of the other has an interesting consequence for periodic systems. We have seen that such systems can always be described in terms of action-angle variables. Thus, if we introduce action-angle variables (ψ, J) for a one-dimensional system with Hamiltonian $K = E(J)$, then

$$\begin{aligned} \psi(t) &= \psi_0 + \omega(J)t \\ J(t) &= J_0 \\ \omega(J) &= \frac{\partial E}{\partial J}. \end{aligned} \tag{7.211}$$

We suppose that ψ_0 and J_0 also have a Gaussian distribution as in (7.203). If we further assume that the J-distribution is sharply peaked at $\bar{J}$ so that the first two terms in

$$\omega(J) = \omega(\bar{J}) + \frac{\partial^2 E}{\partial \bar{J}^2}(J - \bar{J}) + \dots \tag{7.212}$$

give an adequate representation of $\omega(J)$, we regain (7.207) once more, with

$$\psi - \bar{\psi} - \omega(\bar{J})t = \sigma\xi, \quad J - \bar{J} = \tau\eta \tag{7.213}$$

and

$$\alpha = \frac{\mu\sigma}{\tau}, \quad \mu = \frac{\partial^2 E}{\partial \bar{J}^2}. \tag{7.214}$$

Consequently, the action variable becomes better defined at the expense of the angle variable as time passes. In fact, since ψ increases by 2π during each cycle of the system, we will have lost all information of ψ after a time t, where the semi-major axis of the ellipse, $a \simeq \alpha t$; exceeds $2\pi/\sigma$, i.e. when

$$t \gg t_c = \frac{2\pi}{\mu\tau}. \tag{7.215}$$

Thus, t_c represents a critical time for the periodic system. After this time we have lost all information on the angle variable. Notice that the time t_c is inversely proportional to the spread τ of the action variable, provided that $\mu \neq 0$. The latter circumstance is a special feature of a linear harmonic oscillator for which $E(J) = \omega J$ and $\mu = 0$. Thus, $t_c \to \infty$ in this case, a result that is immediately understood by noting that the ξ and η (or ψ and J) distributions remain uncoupled for all times if $\mu = 0$.

Similar considerations follow for a system with n degrees of freedom, described by a Hamiltonian $K = E(J_1, \dots, J_n)$. Equations (7.211) and (7.212) are replaced by

$$\begin{aligned} \psi_k(t) &= \psi_{0,k} + \omega_k t \\ J_k(t) &= J_{0,k} \\ \omega_k &= \frac{\partial E}{\partial J_k} \end{aligned} \tag{7.216}$$

and

$$\begin{aligned} \omega_k &= \bar{\omega}_k + \sum_l \frac{\partial^2 E}{\partial \bar{J}_k \partial \bar{J}_l}(J_l - \bar{J}_l) + \dots \\ \bar{\omega}_k &= \omega_k(\bar{J}_1, \dots, \bar{J}_n) \end{aligned} \tag{7.217}$$

for each value of k. Assuming Gaussian distributions in $\psi_{0,k}$ and $J_{0,k}$ with widths σ_k and τ_k,

$$\sum_{k=1}^{n} \{(\xi_k - t\sum_{l=1}^{n} \alpha_{kl}\eta_l)^2 + \eta_k^2\} = 1 \tag{7.218}$$

replaces (7.207), if we set

$$\psi_k - \bar{\psi}_k - \bar{\omega}_k t = \sigma_k\xi_k, \quad J_k - \bar{J}_k = \tau_k\eta_k, \tag{7.219}$$

and call

$$\alpha_{kl} = \frac{\mu_{kl}\tau_l}{\sigma_k}, \quad \mu_{kl} = \frac{\partial^2 E}{\partial \bar{J}_k \partial J_l}. \tag{7.220}$$

We take $\sigma_k\tau_k = \sigma_l\tau_l \neq 0$ for simplicity and exclude systems (harmonic oscillators) for which $\mu_{kl} = 0$. Then α_{kl} is a real, symmetric matrix which can be diagonalized by a real orthogonal matrix U. Call its eigenvalues α_i, and transform ξ and η to new variables $U\xi'$ and $U\eta'$. Then (7.218) is replaced by

$$\sum_{i=1}^{n}\{(\xi_i' - \alpha_i t\eta_i')^2 + \eta_i'^2\} = 1. \tag{7.221}$$

Being a sum of terms like those in (7.207), we can analyze this equation for a longest critical time (α is the smallest eigenvalue of α_{kl})

$$t_c = \frac{2\pi}{\alpha}[\sum_{k=1}^{n}\frac{1}{\sigma_k^2}]^{\frac{1}{2}}, \tag{7.222}$$

after which time we lack all information about one half of the action-angle variables: The remarkable feature about this result is that it holds for any system for which action-angle variables can be defined. The Hamiltonian $K = E(J_1, \ldots, J_n)$ itself has remained unspecified and incidental to the whole discussion.

Problems

7-1. Show that the generating functions $T(p, Q, t)$ and $U(p, P, t)$ lead to the canonical transformations given by (7.53) - (7.58).

7-2. Prove directly that the point transformation $q_k = f_k(Q_1, \ldots, Q_n; t)$ where $k = 1, 2, \ldots, n$ is canonical.

7-3. Set up and solve the Hamilton-Jacobi equation for the motion of a projectile near the surface of the earth. Find both the equation for the path, as well as the coordinates as a function of the time.

7-4. Set up and solve the Hamilton-Jacobi equations for the motion of a top spinning about a fixed point. Take the analysis as far as you can.

7-5. Confirm the results given in paragraph 7-5(ii) of the text concerning the motion of a charged particle in crossed electric and magnetic fields.

7-6. Obtain the result

$$E = -\frac{1}{2}\frac{m\alpha^2}{(J_\phi + J_\theta + J_r)^2} \tag{7.223}$$

for the energy of a particle, mass m, moving in the central potential $V(r) = -\alpha/r$. (Remark: The action variable $2\pi J_r = \oint p_r\, dr$ can be evaluated in a straightforward way by solving the energy equation for p_r. The resulting integral over r can then be done either pedantically, using integration tables, or elegantly using contour integration.)

7-7. Using parabolic coordinates, set up and solve the Hamilton-Jacobi equation for a hydrogen atom in an external electric field E_0. Introduce action-angle variables and check the conclusions reached in paragraph 7-7 of the text. Address the question of whether there is a critical electric field strength for which the orbit is rendered unbound. Estimate a typical value for ea_0E_0 (e = electron charge, a_0 = Bohr radius) and decide whether or not perturbation theory is applicable to hydrogen for the sorts of electric fields usually available in the laboratory.

7-8. A hydrogen atom is placed in a constant magnetic field **B**, pointing perpendicular to the plane of the orbiting electron. Assuming **B** is "small", carry out a similar analysis to that in paragraph 7-7, to study how the electron orbit is changed by the external field.

7-9. It is known from electrodynamics that an accelerating charge e radiates energy at the rate

$$P = \frac{2}{3}\frac{e^2}{c^2}|\dot{\mathbf{v}}|^2, \tag{7.224}$$

where $\dot{\mathbf{v}}$ is the acceleration and c the velocity of light, and the charge e is measured in esu. Now consider the electron in the lowest state of a hydrogen atom. Classically its orbit is a circle of radius $e_0 = \hbar^2/mc^2 = 0.53 \times 10^{-8}$ cm (the Bohr radius). Assuming that this orbit remains approximately circular while the electron is syphoning off energy at the rate P, show that the radius of the orbit will contract to

$$r(t) = a_0[1 - \frac{8}{3}\frac{r_0^3}{a_0^3}\frac{t}{\tau}]^{\frac{1}{3}}. \tag{7.225}$$

Here the constants are the classical electron radius $r_0 = e^2/mc^2$ (= 2.8×10^{-13} cm) and τ, where $c\tau = (2/3)r_0$ is the characteristic time for light to travel across the classical electron. Assuming the above formula to be valid at all distances, calculate the time for a classical hydrogen atom to contract to a point. How many orbits does the electron complete in this time? Why do real hydrogen atoms not behave in this way? (Don't just say "quantum mechanics". Try to understand the reason in detail).

7-10. The orientation in space and geometric constants of an orbit of a planet are sometimes described in terms of its *Delaunay elements*. Look up what these are and how they relate to the action-angle variables. Then study the problem posed by (7.157) again, that is particle motion in the field $V(r) = -\alpha/r - (\mathbf{r} \cdot \mathbf{F})$, but where the force $\mathbf{F}$ now has an *arbitrary* orientation relative to the plane of the elliptic orbit. Generalize problem 7-8 in the same way.

7-11. Prove that the Jacobian

$$J = \frac{\partial(q_1, \ldots, q_n, p_1, \ldots, p_n)}{\partial(Q_1, \ldots, Q_n, P_1, \ldots, P_n)} = 1 \tag{7.226}$$

if the transformation from the (q_k, p_k) to the (Q_k, P_k) is a canonical one.

7-12. Show that the fundamental Poisson brackets of $2n$ canonical variables q_k and p_k are

$$[q_k, p_k] = \delta_{kl}, \quad [q_k, q_l] = 0, \quad [p_k, p_l] = 0. \tag{7.227}$$

We saw in the text that Poisson brackets are canonical invariants. Prove the converse of this statement for the fundamental Poisson brackets, that is to say, transform to new (not necessarily canoncial) variables Q_k and P_k, and show that these variables also satisfy the canonical equations of motion if q_k and p_k do, provided that

$$[Q_k, P_k]_{q,p} = \delta_{kl}, \quad [Q_k, Q_l]_{q,p} = 0, \quad [P_k, P_l]_{q,p} = 0. \tag{7.228}$$

Assume the transformation $(q, p) \to (Q, P)$ to be time-independent for simplicity. Thus the canonical invariance of the fundamental Poisson brackets serves as both a necessary and a sufficient condition for a transformation to be canonical.

7-13. Prove directly that the Poisson bracket $[F, G]_{q,p}$ of any two functions $F(q, p)$ and $G(q, p)$ is a canonical invariant. Hint: Start by proving the statement for the fundamental brackets and then use these as building blocks for the general proof.

7-14. Show that the action variables J_k defined in (7.123) also have the property of being *adiabatic invariants*. This means that if the J_k are constructed for a Hamiltonian $H(q, p, \varepsilon)$ that depends on some time-dependent external parameter $\varepsilon(t)$, then the J_k remain constant for slow changes in $\varepsilon(t)$ (slow means $\dot{\varepsilon} \ll \varepsilon/T$, where T is a typical period of motion of the system if ε is constant). Start off your proof by observing that the generating function W inducing the transformation $(q_k, p_k) \to (\psi_k, J_k)$ will depend on $\varepsilon(t)$ in this case, and thus become time-dependent. Therefore, the new Hamiltonian will be given by

$$K = E(J_1, \ldots, J_n) + \dot{\varepsilon}\frac{\partial W}{\partial \varepsilon}, \tag{7.229}$$

where $W = W(q, J, \varepsilon) = \sum_{i=1}^{n} W_i(q_i, J, \varepsilon)$. Now write down the equation of motion for J_k and average over a time large relative to T but small relative to time over which $\varepsilon(t)$ changes appreciably. Complete the proof by showing that $\partial^2 W/\partial\psi_k \partial\varepsilon$ is a one-valued function of the q_k and thus expressable as a periodic function in the ψ_k. The time-average of such a function vanishes.