2

Review of Classical Mechanics

In this chapter we will develop the Lagrangian and Hamiltonian formulations of mechanics starting from Newton's laws. These subsequent reformulations of mechanics bring with them a great deal of elegance and computational ease. But our principal interest in them stems from the fact that they are the ideal springboards from which to make the leap to quantum mechanics. The passage from the Lagrangian formulation to quantum mechanics was carried out by Feynman in his path integral formalism. A more common route to quantum mechanics, which we will follow for the most part, has as its starting point the Hamiltonian formulation, and it was discovered mainly by Schrödinger, Heisenberg, Dirac, and Born.

It should be emphasized, and it will soon become apparent, that all three formulations of mechanics are essentially the same theory, in that their domains of validity and predictions are identical. Nonetheless, in a given context, one or the other may be more inviting for conceptual, computational, or simply aesthetic reasons.

2.1. The Principle of Least Action and Lagrangian Mechanics

Let us take as our prototype of the Newtonian scheme a point particle of mass m moving along the x axis under a potential $V(x)$. According to Newton's Second Law,

$$m\frac{d^2x}{dt^2}=-\frac{dV}{dx} \tag{2.1.1}$$

If we are given the initial state variables, the position $x(t_i)$ and velocity $\dot{x}(t_i)$, we can calculate the classical trajectory $x_{cl}(t)$ as follows. Using the initial velocity and acceleration [obtained from Eq. (2.1.1)] we compute the position and velocity at a time $t_i+\Delta t$. For example,

$$x_{\text{cl}}(t_i+\Delta t)=x(t_i)+\dot{x}(t_i)\Delta t$$

Having updated the state variables to the time $t_i+\Delta t$, we can repeat the process again to inch forward to $t_i+2\Delta t$ and so on.

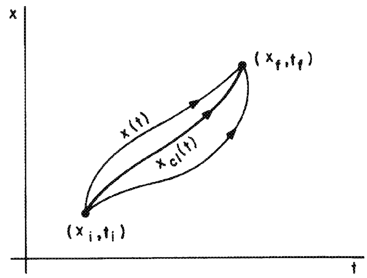


Figure 2.1. The Lagrangian formalism asks what distinguishes the actual path $x_{cl}(t)$ taken by the particle from all possible paths connecting the end points (x_i, t_i) and (x_f, t_f).

The equation of motion being second order in time, two pieces of data, $x(t_i)$ and $\dot{x}(t_i)$, are needed to specify a unique $x_{cl}(t)$. An equivalent way to do the same, and one that we will have occasion to employ, is to specify two space-time points (x_i, t_i) and (x_f, t_f) on the trajectory.

The above scheme readily generalizes to more than one particle and more than one dimension. If we use n Cartesian coordinates $(x_1, x_2, \ldots, x_n)$ to specify the positions of the particles, the spatial configuration of the system may be visualized as a point in an n-dimensional *configuration space*. (The term "configuration space" is used even if the n coordinates are not Cartesian.) The motion of the representative point is given by

$$m_j \frac{d^2 x_j}{dt^2} = -\frac{\partial V}{\partial x_j} \tag{2.1.2}$$

where m_j stands for the mass of the particle whose coordinate is x_j. These equations can be integrated step by step, just as before, to determine the trajectory.

In the Lagrangian formalism, the problem of a single particle in a potential $V(x)$ is posed in a different way: given that the particle is at x_i and x_f at times t_i and t_f, respectively, what is it that distinguishes the actual trajectory $x_{cl}(t)$ from all other trajectories or paths that connect these points? (See Fig. 2.1.)

The Lagrangian approach is thus global, in that it tries to determine at one stroke the entire trajectory $x_{cl}(t)$, in contrast to the local approach of the Newtonian scheme, which concerns itself with what the particle is going to do in the next infinitesimal time interval.

The answer to the question posed above comes in three parts:

(1) Define a function $\mathscr{L}$, called the *Lagrangian*, given by $\mathscr{L} = T - V$, T and V being the kinetic and potential energies of the particle. Thus $\mathscr{L} = \mathscr{L}(x, \dot{x}, t)$. The explicit t dependence may arise if the particle is in an external time-dependent field. We will, however, assume the absence of this t dependence.

(2) For each path $x(t)$ connecting (x_i, t_i) and (x_f, t_f), calculate the *action* $S[x(t)]$ defined by

$$S[x(t)] = \int_{t_i}^{t_f} \mathscr{L}(x, \dot{x})\, dt \tag{2.1.3}$$

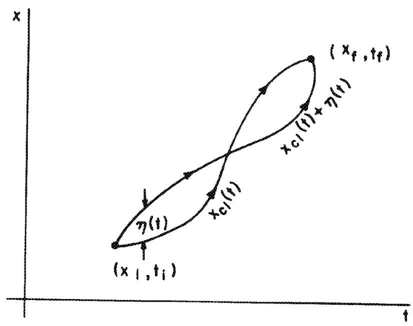


Figure 2.2. If $x_{cl}(t)$ minimizes S, then $\delta S^{(1)}=0$ if we go to any nearby path $x_{cl}(t)+\eta(t)$.

We use square brackets to enclose the argument of S to remind us that the function S depends on an entire path or function $x(t)$, and not just the value of x at some time t. One calls S a *functional* to signify that it is a function of a function.

(3) The classical path is one on which S is a minimum. (Actually we will only require that it be an extremum. It is, however, customary to refer to this condition as the *principle of least action.*)

We will now verify that this principle reproduces Newton's Second Law.

The first step is to realize that a functional $S[x(t)]$ is just a function of n variables as $n\to\infty$. In other words, the function $x(t)$ simply specifies an infinite number of values $x(t_i), \ldots, x(t), \ldots, x(t_f)$, one for each instant in time t in the interval $t_i \le t \le t_f$, and S is a function of these variables. To find its minimum we simply generalize the procedure for the finite n case. Let us recall that if $f=f(x_1, \ldots, x_n)=f(\mathbf{x})$; the minimum $\mathbf{x}^0$ is characterized by the fact that if we move away from it by a small amount $\boldsymbol{\eta}$ in any direction, the first-order change $\delta f^{(1)}$ in f vanishes. That is, if we make a Taylor expansion:

$$f(\mathbf{x}^0+\boldsymbol{\eta})=f(\mathbf{x}^0)+\sum_{i=1}^{n}\left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{x}^0}\eta_i+\text{higher-order terms in }\eta \tag{2.1.4}$$

then

$$\delta f^{(1)}\equiv\sum_{i=1}^{n}\left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{x}^0}\eta_i=0 \tag{2.1.5}$$

From this condition we can deduce an equivalent and perhaps more familiar expression of the minimum condition: every first-order partial derivative vanishes at $\mathbf{x}^0$. To prove this, for say, $\partial f/\partial x_i$, we simply choose $\boldsymbol{\eta}$ to be along the ith direction. Thus

$$\left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{x}^0}=0, \qquad i=1, \ldots, n \tag{2.1.6}$$

Let us now mimic this procedure for the action S. Let $x_{cl}(t)$ be the path of least action and $x_{cl}(t)+\eta(t)$ a "nearby" path (see Fig. 2.2). The requirement that all paths coincide at t_i and t_f means

$$\eta(t_i)=\eta(t_f)=0 \tag{2.1.7}$$

Now

$$
\begin{aligned}
S[x_{\rm cl}(t)+\eta(t)] &= \int_{t_i}^{t_f} \mathscr{L}(x_{\rm cl}(t)+\eta(t);\dot{x}_{\rm cl}(t)+\dot{\eta}(t))\,dt \\
&= \int_{t_i}^{t_f}\left[\mathscr{L}(x_{\rm cl}(t),\dot{x}_{\rm cl}(t))+\left.\frac{\partial\mathscr{L}}{\partial x(t)}\right|_{x_{\rm cl}}\cdot\eta(t)\right. \\
&\qquad \left. +\left.\frac{\partial\mathscr{L}}{\partial\dot{x}(t)}\right|_{x_{\rm cl}}\cdot\dot{\eta}(t)+\cdots\right]dt \\
&= S[x_{\rm cl}(t)]+\delta S^{(1)}+\text{higher-order terms}
\end{aligned}
$$

We set $\delta S^{(1)}=0$ in analogy with the finite variable case:

$$
0=\delta S^{(1)}=\int_{t_i}^{t_f}\left[\left.\frac{\partial\mathscr{L}}{\partial x(t)}\right|_{x_{\rm cl}}\cdot\eta(t)+\left.\frac{\partial\mathscr{L}}{\partial\dot{x}(t)}\right|_{x_{\rm cl}}\cdot\dot{\eta}(t)\right]dt
$$

If we integrate the second term by parts, it turns into

$$
\left.\frac{\partial\mathscr{L}}{\partial\dot{x}(t)}\right|_{x_{\rm cl}}\cdot\left.\eta(t)\right|_{t_i}^{t_f}-\int_{t_i}^{t_f}\left[\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\dot{x}(t)}\right]_{x_{\rm cl}}\cdot\eta(t)\,dt
$$

The first of these terms vanishes due to Eq. (2.1.7). So that

$$
0=\delta S^{(1)}=\int_{t_i}^{t_f}\left[\frac{\partial\mathscr{L}}{\partial x(t)}-\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\dot{x}(t)}\right]_{x_{\rm cl}}\cdot\eta(t)\,dt \tag{2.1.8}
$$

Note that the condition $\delta S^{(1)}=0$ implies that S is extremized and not necessarily minimized. We shall, however, continue the tradition of referring to this extremum as the minimum. This equation is the analog of Eq. (2.1.5): the discrete variable η_i is replaced by $\eta(t)$; the sum over i is replaced by an integral over t, and $\partial f/\partial x_i$ is replaced by

$$
\frac{\partial\mathscr{L}}{\partial x(t)}-\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\dot{x}(t)}
$$

There are two terms here playing the role of $\partial f/\partial x_i$ since $\mathscr{L}$ (or equivalently S) has both explicit and implicit (through the $\dot{x}$ terms) dependence on $x(t)$. Since $\eta(t)$ is arbitrary, we may extract the analog of Eq. (2.1.6):

$$
\left\{\frac{\partial\mathscr{L}}{\partial x(t)}-\frac{d}{dt}\left[\frac{\partial\mathscr{L}}{\partial\dot{x}(t)}\right]\right\}_{x_{\rm cl}(t)}=0 \quad \text{for } t_i\le t\le t_f \tag{2.1.9}
$$

To deduce this result for some specific time t_0, we simply choose an $\eta(t)$ that vanishes everywhere except in an infinitesimal region around t_0.

Equation (2.1.9) is the celebrated *Euler–Lagrange equation.* If we feed into it $\mathscr{L}=T-V$, $T=\frac{1}{2}m\dot{x}^2$, $V=V(x)$, we get

$$\frac{\partial \mathscr{L}}{\partial \dot{x}}=\frac{\partial T}{\partial \dot{x}}=m\dot{x}$$

and

$$\frac{\partial \mathscr{L}}{\partial x}=-\frac{\partial V}{\partial x}$$

so that the Euler-Lagrange equation becomes just

$$\frac{d}{dt}(m\dot{x})=-\frac{\partial V}{\partial x}$$

which is just Newton's Second Law, Eq. (2.1.1).

If we consider a system described by n Cartesian coordinates, the same procedure yields

$$\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{x}_i}\right)=\frac{\partial \mathscr{L}}{\partial x_i} \qquad (i=1,\ldots,n) \tag{2.1.10}$$

Now

$$T=\tfrac{1}{2}\sum_{i=1}^{n} m_i(\dot{x}_i)^2$$

and

$$V=V(x_1,\ldots,x_n)$$

so that Eq. (2.1.10) becomes

$$\frac{d}{dt}(m_i\dot{x}_i)=-\frac{\partial V}{\partial x_i}$$

which is identical to Eq. (2.1.2). Thus the minimum (action) principle indeed reproduces Newtonian mechanics if we choose $\mathscr{L}=T-V$.

Notice that we have assumed that V is velocity-independent in the above proof. An important force, that of a magnetic field $\mathbf{B}$ on a moving charge is excluded by this restriction, since $\mathbf{F}_B=q\mathbf{v}\times\mathbf{B}$, q being the charge of the particle and $\mathbf{v}=\dot{\mathbf{r}}$ its velocity. We will show shortly that this force too may be accommodated in the Lagrangian formalism, in the sense that we can find an $\mathscr{L}$ that yields the correct force law when Eq. (2.1.10) is employed. But this $\mathscr{L}$ no longer has the form $T-V$. One therefore frees oneself from the notion that $\mathscr{L}=T-V$; and views $\mathscr{L}$ as some function $\mathscr{L}(x_i, \dot{x}_i)$ which yields the correct Newtonian dynamics when fed into the Euler–Lagrange equations. To the reader who wonders why one bothers to even deal with a Lagrangian when all it does is yield Newtonian force laws in the end, I present a few of its main attractions besides its closeness to quantum mechanics. These will then be illustrated by means of an example.

(1) In the Lagrangian scheme one has merely to construct a single *scalar* $\mathscr{L}$ and all the equations of motion follow by simple differentiation. This must be contrasted with the Newtonian scheme, which deals with vectors and is thus more complicated.

(2) The Euler–Lagrange equations (2.1.10) have the same *form* if we use, instead of the n Cartesian coordinates $x_1, \ldots, x_n$, *any* general set of n independent coordinates $q_1, q_2, \ldots, q_n$. To remind us of this fact we will rewrite Eq. (2.1.10) as

$$\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{q}_i}\right)=\frac{\partial \mathscr{L}}{\partial q_i} \tag{2.1.11}$$

One can either verify this by brute force, making a change of variables in Eq. (2.1.10) and seeing that an identical equation with x_i replaced by q_i follows, or one can simply go through our derivation of the minimum action condition and see that nowhere were the coordinates assumed to be Cartesian. Of course, at the next stage, in showing that the Euler–Lagrange equations were equivalent to Newton's, Cartesian coordinates *were* used, for in these coordinates the kinetic energy T and the Newtonian equations have simple forms. But once the principle of least action is seen to generate the correct dynamics, we can forget all about Newton's laws and use Eq. (2.1.11) as the equations of motion. What is being emphasized is that these equations, which express the condition for least action, are form invariant under an arbitrary change of coordinates. This form invariance must be contrasted with the Newtonian equation (2.1.2), which presumes that the x_i are Cartesian. If one trades the x_i for another non-Cartesian set of q_i, Eq. (2.1.2) will have a different form (see Example 2.1.1 at the end of this section).

Equation (2.1.11) can be made to resemble Newton's Second Law if one defines a quantity

$$p_i=\frac{\partial \mathscr{L}}{\partial \dot{q}_i} \tag{2.1.12}$$

called the *canonical momentum conjugate to* q_i and the quantity

$$F_i=\frac{\partial \mathscr{L}}{\partial q_i} \tag{2.1.13}$$

called the *generalized force conjugate to* q_i. Although the rate of change of the canonical momentum equals the generalized force, one must remember that neither is p_i always a linear momentum (mass times velocity or "mv" momentum), nor is F_i always a force (with dimensions of mass times acceleration). For example, if q_i is an angle θ, p_i will be an angular momentum and F_i a torque.

(3) Conservation laws are easily obtained in this formalism. Suppose the Lagrangian depends on a certain velocity $\dot{q}_i$ but not on the corresponding coordinate q_i. The latter is then called a *cyclic coordinate.* It follows that the corresponding p_i is conserved:

$$\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{q}_i}\right)=\frac{dp_i}{dt}=\frac{\partial \mathscr{L}}{\partial q_i}=0 \tag{2.1.14}$$

Although Newton's Second Law, Eq. (2.1.2), also tells us that if a Cartesian coordinate x_i is cyclic, the corresponding momentum $m_i\dot{x}_i$ is conserved, Eq. (2.1.14) is more general. Consider, for example, a potential $V(x, y)$ in two dimensions that depends only upon $\rho=(x^2+y)^{1/2}$, and not on the polar angle ϕ, so that $V(\rho, \phi)=V(\rho)$. It follows that ϕ is a cyclic coordinate, as T depends only on $\dot{\phi}$ (see Example 2.1.1 below). Consequently $\partial\mathscr{L}/\partial\dot{\phi}=p_\phi$ is conserved. In contrast, no obvious conservation law arises from the Cartesian Eqs. (2.1.2) since neither x nor y is cyclic. If one rewrites Newton's laws in polar coordinates to exploit $\partial V/\partial\phi=0$, the corresponding equations get complicated due to centrifugal and Coriolis terms. It is the Lagrangian formalism that allows us to choose coordinates that best reflect the symmetry of the potential, without altering the simple form of the equations.

Example 2.1.1. We now illustrate the above points through an example. Consider a particle moving in a plane. The Lagrangian, in Cartesian coordinates, is

$$\begin{aligned}\mathscr{L}&=\tfrac{1}{2}m(\dot{x}^2+\dot{y}^2)-V(x, y)\\ &=\tfrac{1}{2}m\mathbf{v}\cdot\mathbf{v}-V(x, y)\end{aligned} \tag{2.1.15}$$

where $\mathbf{v}$ is the velocity of the particle, with $\mathbf{v}=\dot{\mathbf{r}}$, $\mathbf{r}$ being its position vector. The corresponding equations of motion are

$$m\ddot{x}=-\frac{\partial V}{\partial x} \tag{2.1.16}$$

$$m\ddot{y}=-\frac{\partial V}{\partial y} \tag{2.1.17}$$

which are identical to Newton's laws. If one wants to get the same Newton's laws in terms of polar coordinates ρ and ϕ, some careful vector analysis is needed to unearth the centrifugal and Coriolis terms:

$$m\ddot{\rho}=-\frac{\partial V}{\partial\rho}+m\rho(\dot{\phi})^2 \tag{2.1.18}$$

$$m\ddot{\phi}=-\frac{1}{\rho^2}\frac{\partial V}{\partial\phi}-\frac{2m\dot{\rho}\dot{\phi}}{\rho} \tag{2.1.19}$$

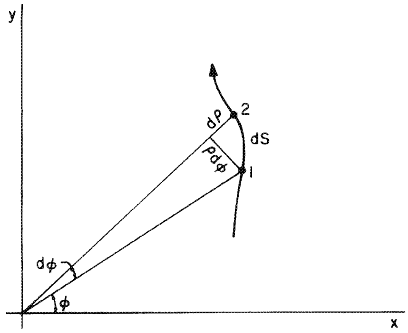


Figure 2.3. Points (1) and (2) are positions of the particle at times differing by Δt.

Notice the difference in form between Eqs. (2.1.16) and (2.1.17) on the one hand and Eqs. (2.1.18) and (2.1.19) on the other.

In the Lagrangian scheme one has only to recompute $\mathscr{L}$ in polar coordinates. From Fig. 2.3 it is clear that the distance traveled by the particle in time Δt is

$$dS = [(d\rho)^2 + (\rho\, d\phi)^2]^{1/2}$$

so that the magnitude of velocity is

$$v = \frac{dS}{dt} = [(\dot{\rho})^2 + \rho^2(\dot{\phi})^2]^{1/2}$$

and

$$\mathscr{L} = \tfrac{1}{2}m(\dot{\rho}^2 + \rho^2\dot{\phi}^2) - V(\rho, \phi) \tag{2.1.20}$$

(Notice that in these coordinates T involves not just the velocities $\dot{\rho}$ and $\dot{\phi}$ but also the coordinate ρ. This does not happen in Cartesian coordinates.) The equations of motion generated by this $\mathscr{L}$ are

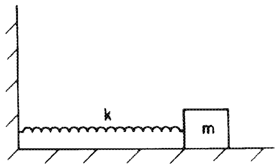
$$\frac{d}{dt}(m\dot{\rho}) = -\frac{\partial V}{\partial \rho} + m\rho\dot{\phi}^2 \tag{2.1.21}$$

$$\frac{d}{dt}(m\rho^2\dot{\phi}) = -\frac{\partial V}{\partial \phi} \tag{2.1.22}$$

which are the same as Eqs. (2.1.18) and (2.1.19). In Eq. (2.1.22) the canonical momentum $p_\phi = m\rho^2\dot{\phi}$ is the angular momentum and the generalized force $-\partial V/\partial\phi$ is the torque, both along the z axis. Notice how easily the centrifugal and Coriolis forces came out.

Finally, if $V(\rho, \phi) = V(\rho)$, the conservation of p_ϕ is obvious in Eq. (2.1.22). The conservation of p_ϕ follows from Eq. (2.1.19) only after some manipulations and is practically invisible in Eqs. (2.1.16) and (2.1.17). Both the conserved quantity and its conservation law arise naturally in the Lagrangian scheme. □

*Exercise 2.1.1.** Consider the following system, called a *harmonic oscillator*. The block has a mass m and lies on a frictionless surface. The spring has a force constant k.



Write the Lagrangian and get the equations of motion.

*Exercise 2.1.2.** Do the same for the coupled-mass problem discussed at the end of Section 1.8. Compare the equations of motion with Eqs. (1.8.24) and (1.8.25).

*Exercise 2.1.3.** A particle of mass m moves in three dimensions under a potential $V(r, \theta, \phi) = V(r)$. Write its $\mathscr{L}$ and find the equations of motion.

2.2. The Electromagnetic Lagrangian‡

Recall that the force on a charge q due to an electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ is given by

$$\mathbf{F} = q\left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right) \tag{2.2.1}$$

where $\mathbf{v} = \dot{\mathbf{r}}$ is the velocity of the particle. Since the force is velocity-dependent, we must analyze the problem afresh, not relying on the preceding discussion, which was restricted to velocity-independent forces.

Now it turns out that if we use

$$\mathscr{L}_{e \cdot m} = \tfrac{1}{2}m\mathbf{v}\cdot\mathbf{v} - q\phi + \frac{q}{c}\mathbf{v}\cdot\mathbf{A} \tag{2.2.2}$$

we get the correct electromagnetic force laws. In Eq. (2.2.2) c is the velocity of light, while ϕ and $\mathbf{A}$ are the scalar and vector potentials related to $\mathbf{E}$ and $\mathbf{B}$ via

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \tag{2.2.3}$$

and

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{2.2.4}$$

‡ See Section 18.4 for a review of classical electromagnetism.

The Euler–Lagrange equations corresponding to $\mathscr{L}_{e\cdot m}$ are

$$\frac{d}{dt}\left(m\dot{x}_i+\frac{q}{c}A_i\right)=-q\frac{\partial\phi}{\partial x_i}+\frac{q}{c}\frac{\partial(\mathbf{v}\cdot\mathbf{A})}{\partial x_i}, \qquad i=1,2,3 \tag{2.2.5}$$

Combining the three equations above into a single vector equation we get

$$\frac{d}{dt}\left(m\mathbf{v}+\frac{q\mathbf{A}}{c}\right)=-q\nabla\phi+\frac{q}{c}\nabla(\mathbf{v}\cdot\mathbf{A}) \tag{2.2.6}$$

The canonical momentum is

$$\mathbf{p}=m\mathbf{v}+\frac{q\mathbf{A}}{c} \tag{2.2.7}$$

Rewriting Eq. (2.2.6), we get

$$\frac{d}{dt}(m\mathbf{v})=-q\nabla\phi+\frac{q}{c}\left[-\frac{d\mathbf{A}}{dt}+\nabla(\mathbf{v}\cdot\mathbf{A})\right] \tag{2.2.8}$$

Now, the total derivative $d\mathbf{A}/dt$ has two parts: an explicit time dependence $\partial A/\partial t$, plus an implicit one $(\mathbf{v}\cdot\nabla)\mathbf{A}$ which represents the fact that a spatial variation in $\mathbf{A}$ will appear as a temporal variation to the moving particle.Now Eq. (2.2.8) becomes

$$\frac{d}{dt}(m\mathbf{v})=-q\nabla\phi-\frac{q}{c}\frac{\partial\mathbf{A}}{\partial t}+\frac{q}{c}[\nabla(\mathbf{v}\cdot\mathbf{A})-(\mathbf{v}\cdot\nabla)\mathbf{A}] \tag{2.2.9}$$

which is identical to Eq. (2.2.1) by virtue of the identity

$$\mathbf{v}\times(\nabla\times\mathbf{A})=\nabla(\mathbf{v}\cdot\mathbf{A})-(\mathbf{v}\cdot\nabla)\mathbf{A}$$

Notice that $\mathscr{L}_{e\cdot m}$ is not of the form $T-V$, for the quantity $U=q\phi-(q/c)\mathbf{v}\cdot\mathbf{A}$ (sometimes called the *generalized potential*) cannot be interpreted as the potential energy of the charged particle. First of all, the force due to a time-dependent electromagnetic field is not generally conservative and does not admit a path-independent work function to play the role of a potential. Even in the special cases when the force is conservative, only $q\phi$ can be interpreted as the electrical potential energy. The $[-q(\mathbf{v}\cdot\mathbf{A})/c]$ term is not a magnetic potential energy, since the magnetic force $\mathbf{F}_B=q(\mathbf{v}\times\mathbf{B})/c$ never does any work, being always perpendicular to the velocity. To accommodate forces such as the electro-magnetic, we must, therefore, redefine $\mathscr{L}$ to be that function $\mathscr{L}(q,\dot{q},t)$ which, when fed into the Euler–Lagrange equations, reproduces the correct dynamics. The rule $\mathscr{L}=T-V$ becomes just a useful mnemonic for the case of conservative forces.

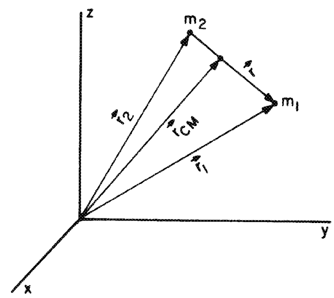


Figure 2.4. The relation between $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{r}_{CM}$, $\mathbf{r}$.

2.3. The Two-Body Problem

We discuss here a class of problems that plays a central role in classical physics: that of two masses m_1 and m_2 exerting equal and opposite forces on each other. Since the particles are responding to each other and nothing external, it follows that the potential between them depends only on the *relative coordinate* $\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2$ and not the individual positions $\mathbf{r}_1$ and $\mathbf{r}_2$. But $V(\mathbf{r}_1, \mathbf{r}_2)=V(\mathbf{r}_1-\mathbf{r}_2)$ means in turn that there are three cyclic coordinates, for V depends on only three variables rather than the possible six. (In Cartesian coordinates, since T is a function only of velocities, a coordinate missing in V is also cyclic.) The corresponding conserved momenta will of course be the three components of the total momentum, which are conserved in the absence of external forces. To bring out these features, it is better to trade $\mathbf{r}_1$ and $\mathbf{r}_2$ in favor of

$$\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2 \tag{2.3.1}$$

and

$$\mathbf{r}_{CM}=\frac{m_1\mathbf{r}_1+m_2\mathbf{r}_2}{m_1+m_2} \tag{2.3.2}$$

where $\mathbf{r}_{CM}$ is called the *center-of-mass (CM) coordinate*. One can invert Eqs. (2.3.1) and (2.3.2) to get (see Fig. 2.4)

$$\mathbf{r}_1=\mathbf{r}_{CM}+\frac{m_2\mathbf{r}}{m_1+m_2} \tag{2.3.3}$$

$$\mathbf{r}_2=\mathbf{r}_{CM}-\frac{m_1\mathbf{r}}{m_1+m_2} \tag{2.3.4}$$

If one rewrites the Lagrangian

$$\mathscr{L}=\tfrac{1}{2}m_1|\dot{\mathbf{r}}_1|^2+\tfrac{1}{2}m_2|\dot{\mathbf{r}}_2|^2-V(\mathbf{r}_1-\mathbf{r}_2) \tag{2.3.5}$$

in terms of $\mathbf{r}_{CM}$ and $\mathbf{r}$, one gets

$$\mathscr{L}=\frac{1}{2}(m_1+m_2)|\dot{\mathbf{r}}_{CM}|^2+\frac{1}{2}\frac{m_1m_2}{m_1+m_2}|\dot{\mathbf{r}}|^2-V(\mathbf{r}) \tag{2.3.6}$$

The main features of Eq. (2.3.6) are the following.

(1) The problem of two mutually interacting particles has been transformed to that of two fictitious particles that do not interact with each other. In other words, the equations of motion for $\mathbf{r}$ do not involve $\mathbf{r}_{\text{CM}}$ and vice versa, because $\mathscr{L}(\mathbf{r}, \dot{\mathbf{r}}; \mathbf{r}_{\text{CM}}, \dot{\mathbf{r}}_{\text{CM}}) = \mathscr{L}(\mathbf{r}, \dot{\mathbf{r}}) + \mathscr{L}(\mathbf{r}_{\text{CM}}, \dot{\mathbf{r}}_{\text{CM}})$.

(2) The first fictitious particle is the CM, of mass $M = m_1 + m_2$. Since $\mathbf{r}_{\text{CM}}$ is a cyclic variable, the momentum $\mathbf{p}_{\text{CM}} = M\dot{\mathbf{r}}_{\text{CM}}$ (which is just the total momentum) is conserved as expected. Since the motion of the CM is uninteresting one usually ignores it. One clear way to do this is to go to the CM frame in which $\dot{\mathbf{r}}_{\text{CM}} = 0$, so that the CM is completely eliminated in the Lagrangian.

(3) The second fictitious particle has mass $\mu = m_1 m_2/(m_1 + m_2)$ (called the *reduced mass*), momentum $\mathbf{p} = \mu\dot{\mathbf{r}}$ and moves under a potential $V(\mathbf{r})$. One has just to solve this one-body problem. If one chooses, one may easily return to the coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$ at the end, using Eqs. (2.3.1) and (2.3.2).

*Exercise 2.3.1.** Derive Eq. (2.3.6) from (2.3.5) by changing variables.

2.4. How Smart Is a Particle?

The Lagrangian formalism seems to ascribe to a particle a tremendous amount of foresight: a particle at (x_i, t_i) destined for (x_f, t_f) manages to calculate ahead of time the action for every possible path linking these points, and takes the one with the least action. But this, of course, is an illusion. The particle need not know its entire trajectory ahead of time, it needs only to obey the Euler–Lagrange equations at each instant in time to minimize the action. This in turn means just following Newton's law, which is to say, the particle has to sample the potential in its immediate vicinity and accelerate in the direction of greatest change.

Our esteem for the particle will sink further when we learn quantum mechanics. We will discover that far from following any kind of strategy, the particle, in a sense, goes from (x_i, t_i) to (x_f, t_f) along all possible paths, giving equal weight to each! How it is that despite this, classical particles do seem to follow $x_{\text{cl}}(t)$ is an interesting question that will be answered when we come to the path integral formalism of quantum mechanics.

2.5. The Hamiltonian Formalism

In the Lagrangian formalism, the independent variables are the coordinates q_i and velocities $\dot{q}_i$. The momenta are derived quantities defined by

$$p_i = \frac{\partial \mathscr{L}}{\partial \dot{q}_i} \tag{2.5.1}$$

In the Hamiltonian formalism one exchanges the roles of $\dot{q}$ and p: one replaces the Lagrangian $\mathscr{L}(q, \dot{q})$‡ by a Hamiltonian $\mathscr{H}(q, p)$ which generates the equations of motion, and $\dot{q}$ becomes a derived quantity,

$$\dot{q}_i = \frac{\partial \mathscr{H}}{\partial p_i} \tag{2.5.2}$$

thereby completing the role reversal of the $\dot{q}$'s and the p's.

There exists a standard procedure for effecting such a change, called a *Legendre transformation*, which is illustrated by the following simple example. Suppose we have a function $f(x)$ with

$$u(x) = \frac{df}{dx} \tag{2.5.3}$$

Let it be possible to invert $u(x)$ to get $x(u)$. [For example if $u(x) = x^3$, $x(u) = u^{1/3}$, etc.] If we define a function

$$g(u) = x(u)u - f(x(u)) \tag{2.5.4}$$

then

$$\frac{dg}{du} = \frac{dx}{du} \cdot u + x(u) - \frac{df}{dx} \cdot \frac{dx}{du} = x(u) \tag{2.5.5}$$

That is to say, in going from f to g (or vice versa) we exchange the roles of x and u. One calls Eq. (2.5.4) a *Legendre transformation* and f and g *Legendre transforms* of each other.

More generally, if $f = f(x_1, x_2, \ldots, x_n)$, one can eliminate a subset $\{x_i, i = 1 \text{ to } j\}$ in favor of the partial derivatives $u_i = \partial f/\partial x_i$ by the transformation

$$g(u_1, \ldots, u_j, x_{j+1}, \ldots, x_n) = \sum_{i=1}^{j} u_i x_i - f(x_1, \ldots, x_n) \tag{2.5.6}$$

It is understood in the right-hand side of Eq. (2.5.6) that all the x_i's to be eliminated have been rewritten as functions of the allowed variables in g. It can be easily verified that

$$\frac{\partial g}{\partial u_i} = x_i \tag{2.5.7}$$

where in taking the above partial derivative, one keeps all the other variables in g constant.

‡ We will often refer to $q_1, \ldots, q_n$ as q and $p_1, \ldots, p_n$ as p.

Table 2.1. Comparison of the Lagrangian and Hamiltonian Formalisms

Lagrangian formalism	Hamiltonian formalism
(1) The state of a system with n degrees of freedom is described by n coordinates $(q_1, \ldots, q_n)$ and n velocities $(\dot{q}_1, \ldots, \dot{q}_n)$, or in a more compact notation by $(q, \dot{q})$.	(1) The state of a system with n degrees of freedom is described by n coordinates and n momenta $(q_1, \ldots, q_n; p_1, \ldots, p_n)$ or, more succinctly, by (q, p).
(2) The state of the system may be represented by a point moving with a definite velocity in an *n-dimensional configuration space*.	(2) The state of the system may be represented by a point in a *$2n$-dimensional phase space*, with coordinates $(q_1, \ldots, q_n; p_1, \ldots, p_n)$.
(3) The n coordinates evolve according to n second-order equations.	(3) The $2n$ coordinates and momenta obey $2n$ first-order equations.
(4) For a given $\mathscr{L}$, several trajectories may pass through a given point in configuration space depending on $\dot{q}$.	(4) For a given $\mathscr{H}$ only one trajectory passes through a given point in phase space.

Applying these methods to the problem in question, we define

$$\mathscr{H}(q, p) = \sum_{i=1}^{n} p_i \dot{q}_i - \mathscr{L}(q, \dot{q}) \tag{2.5.8}$$

where the $\dot{q}$'s are to be written as functions of q's and p's. This inversion is generally easy since $\mathscr{L}$ is a polynomial of rank 2 in $\dot{q}$, and $p_i = \partial\mathscr{L}/\partial\dot{q}_i$ is a polynomial of rank 1 in the $\dot{q}$'s, e.g., Eq. (2.2.7). Consider now

$$\frac{\partial\mathscr{H}}{\partial p_i} = \frac{\partial}{\partial p_i}\left(\sum_j p_j \dot{q}_j - \mathscr{L}\right) \tag{2.5.9}$$

$$= \dot{q}_i + \sum_j p_j \frac{\partial \dot{q}_j}{\partial p_i} - \sum_j \frac{\partial\mathscr{L}}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial p_i}$$

$$= \dot{q}_i \quad \left(\text{since } p_j = \frac{\partial\mathscr{L}}{\partial \dot{q}_j}\right) \tag{2.5.10}$$

[There are no $(\partial\mathscr{L}/\partial q_j)(\partial q_j/\partial p_i)$ terms since q is held constant in $\partial\mathscr{H}/\partial p_i$; that is, q and p are independent variables.] Similarly,

$$\frac{\partial\mathscr{H}}{\partial q_i} = \sum_j p_j \frac{\partial \dot{q}_j}{\partial q_i} - \frac{\partial\mathscr{L}}{\partial q_i} - \sum_j \frac{\partial\mathscr{L}}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial q_i} = -\frac{\partial\mathscr{L}}{\partial q_i} \tag{2.5.11}$$

We now feed in the dynamics by replacing $(\partial\mathscr{L}/\partial q_i)$ by $\dot{p}_i$, and obtain *Hamilton's canonical equations*:

$$\frac{\partial\mathscr{H}}{\partial p_i} = \dot{q}_i, \qquad -\frac{\partial\mathscr{H}}{\partial q_i} = \dot{p}_i \tag{2.5.12}$$

Note that we have altogether $2n$ first-order equations (in time) for a system with n degrees of freedom. Given the initial-value data, $(q_i(0), p_i(0))$, $i = 1, \ldots, n$, we can integrate the equations to get $(q_i(t), p_i(t))$.

Table 2.1 provides a comparison of the Lagrangian and Hamiltonian formalisms.

Now, just as $\mathscr{L}$ may be interpreted as $T-V$ if the force is conservative, so there exists a simple interpretation for $\mathscr{H}$ in this case. Consider the sum $\sum_i p_i \dot{q}_i$. Let us use Cartesian coordinates, in terms of which

$$T=\sum_{i=1}^{n} \tfrac{1}{2} m_i \dot{x}_i^2$$

$$p_i=\frac{\partial \mathscr{L}}{\partial \dot{x}_i}=\frac{\partial T}{\partial \dot{x}_i}=m_i \dot{x}_i$$

and

$$\sum_{i=1}^{n} p_i \dot{x}_i=\sum_{i=1}^{n} m_i \dot{x}_i^2=2T \tag{2.5.13}$$

so that

$$\mathscr{H}=\sum_i p_i \dot{x}_i-\mathscr{L}=T+V \tag{2.5.14}$$

the total energy. Notice that although we used Cartesian coordinates along the way, the resulting equation (2.5.14) is a relation among scalars and thus coordinate independent.

Exercise 2.5.1. Show that if $T=\sum_i \sum_j T_{ij}(q)\dot{q}_i \dot{q}_j$, where $\dot{q}$'s are generalized velocities, $\sum_i p_i \dot{q}_i=2T$.

The Hamiltonian method is illustrated by the simple example of a harmonic oscillator, for which

$$\mathscr{L}=\tfrac{1}{2} m \dot{x}^2-\tfrac{1}{2} k x^2$$

The canonical momentum is

$$p=\frac{\partial \mathscr{L}}{\partial \dot{x}}=m \dot{x}$$

It is easy to invert this relation to obtain $\dot{x}$ as a function of p:

$$\dot{x}=p/m$$

and obtain

$$\mathscr{H}(x, p) = T + V = \tfrac{1}{2}m[\dot{x}(p)]^2 + \tfrac{1}{2}kx^2$$
$$= \frac{p^2}{2m} + \frac{1}{2}kx^2 \tag{2.5.15}$$

The equations of motion are

$$\frac{\partial \mathscr{H}}{\partial p} = \dot{q} \rightarrow \frac{p}{m} = \dot{x} \tag{2.5.16}$$

$$-\frac{\partial \mathscr{H}}{\partial q} = \dot{p} \rightarrow -kx = \dot{p} \tag{2.5.17}$$

These equations can be integrated in time, given the initial q and p. If, however, we want the familiar second-order equation, we differentiate Eq. (2.5.16) with respect to time, and feed it into Eq. (2.5.17) to get

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

Exercise 2.5.2. Using the conservation of energy, show that the trajectories in phase space for the oscillator are ellipses of the form $(x/a)^2 + (p/b)^2 = 1$, where $a^2 = 2E/k$ and $b^2 = 2mE$.

Exercise 2.5.3. Solve Exercise 2.1.2 using the Hamiltonian formalism.

*Exercise 2.5.4.** Show that $\mathscr{H}$ corresponding to $\mathscr{L}$ in Eq. (2.3.6) is $\mathscr{H} = |\mathbf{p}_{\text{CM}}|^2/2M + |\mathbf{p}|^2/2\mu + V(\mathbf{r})$, where M is the total mass, μ is the reduced mass, $\mathbf{p}_{\text{CM}}$ and $\mathbf{p}$ are the momenta conjugate to $\mathbf{r}_{\text{CM}}$ and $\mathbf{r}$, respectively.

2.6. The Electromagnetic Force in the Hamiltonian Scheme

The passage from $\mathscr{L}_{e\cdot m}$ to its Legendre transform $\mathscr{H}_{e\cdot m}$ is not sensitive in any way to the velocity-dependent nature of the force. If $\mathscr{L}_{e\cdot m}$ generated the correct force laws, so will $\mathscr{H}_{e\cdot m}$, the dynamical content of the schemes being identical. In contrast, the velocity independence of the force was assumed in showing that the numerical value of $\mathscr{H}$ is $T + V$, the total energy. Let us therefore repeat the analysis for the electromagnetic case. As

$$\mathscr{L}_{e\cdot m} = \tfrac{1}{2}m\mathbf{v}\cdot\mathbf{v} - q\phi + \frac{q}{c}\mathbf{v}\cdot\mathbf{A}$$

and‡

‡ Note that in this discussion, q is the charge and not the coordinate. The (Cartesian) coordinate $\mathbf{r}$ is hidden in the functions $\mathbf{A}(\mathbf{r}, t)$ and $\phi(\mathbf{r}, t)$.

$$\mathbf{p} = m\mathbf{v} + \frac{q\mathbf{A}}{c}$$

we have

$$\begin{aligned}\mathscr{H}_{e\cdot m} &= \mathbf{p}\cdot\mathbf{v} - \mathscr{L}_{e\cdot m} \\ &= m\mathbf{v}\cdot\mathbf{v} + q\frac{\mathbf{v}\cdot\mathbf{A}}{c} - \frac{1}{2}m\mathbf{v}\cdot\mathbf{v} + q\phi - \frac{q\mathbf{v}\cdot\mathbf{A}}{c} \\ &= \tfrac{1}{2}m\mathbf{v}\cdot\mathbf{v} + q\phi = T + q\phi \end{aligned} \tag{2.6.1}$$

Now, there is something very disturbing about Eq. (2.6.1): the vector potential **A** seems to have dropped out along the way. How is $\mathscr{H}_{e\cdot m}$ to generate the correct dynamics without knowing what **A** is? The answer is, of course, the $\mathscr{H}$ is more than just $T+q\phi$; it is $T+q\phi$ written in terms of the correct variables, in particular, in terms of **p** and not **v**. Making the change of variables, we get

$$\mathscr{H}_{e\cdot m} = \frac{|(\mathbf{p} - q\mathbf{A}/c)|^2}{2m} + q\phi \tag{2.6.2}$$

with the vector potential very much in the picture.

2.7. Cyclic Coordinates, Poisson Brackets, and Canonical Transformations

Cyclic coordinates are defined here just as in the Lagrangian case and have the same significance: if a coordinate q_i is missing in $\mathscr{H}$, then

$$\dot{p}_i = -\frac{\partial\mathscr{H}}{\partial q_i} = 0 \tag{2.7.1}$$

Now, there will be other quantities, such as the energy, that may be conserved in addition to the canonical momenta.§ There exists a nice method of characterizing these in the Hamiltonian formalism. Let $\omega(p, q)$ be some function of the state variables, *with no explicit dependence on t*. Its time variation is given by

$$\begin{aligned}\frac{d\omega}{dt} &= \sum_i \left(\frac{\partial\omega}{\partial q_i}\dot{q}_i + \frac{\partial\omega}{\partial p_i}\dot{p}_i\right) \\ &= \sum_i \left(\frac{\partial\omega}{\partial q_i}\frac{\partial\mathscr{H}}{\partial p_i} - \frac{\partial\omega}{\partial p_i}\frac{\partial\mathscr{H}}{\partial q_i}\right) \\ &\equiv \{\omega, \mathscr{H}\} \end{aligned} \tag{2.7.2}$$

§ Another example is the conservation of $l_z = xp_y - yp_x$ when $V(x, y) = V(x^2+y^2)$. There are no cyclic coordinates here. Of course, if we work in polar coordinates, $V(\rho, \phi) = V(\rho)$, and $p_\phi = m\rho^2\dot{\phi} = l_z$ is conserved because it is the momentum conjugate to the cyclic coordinate ϕ.

where we have defined the Poisson bracket (PB) between two variables $\omega(p, q)$ and $\lambda(p, q)$ to be

$$\{\omega, \lambda\} \equiv \sum_i \left(\frac{\partial\omega}{\partial q_i}\frac{\partial\lambda}{\partial p_i} - \frac{\partial\omega}{\partial p_i}\frac{\partial\lambda}{\partial q_i}\right) \tag{2.7.3}$$

It follows from Eq. (2.7.2) that any variable whose PB with $\mathscr{H}$ vanishes is constant in time, i.e., conserved. In particular $\mathscr{H}$ itself is a constant of motion (identified as the total energy) if it has no explicit t dependence.

*Exercise 2.7.1.** Show that

$$\{\omega, \lambda\} = -\{\lambda, \omega\}$$

$$\{\omega, \lambda+\sigma\} = \{\omega, \lambda\} + \{\omega, \sigma\}$$

$$\{\omega, \lambda\sigma\} = \{\omega, \lambda\}\sigma + \lambda\{\omega, \sigma\}$$

Note the similarity between the above and Eqs. (1.5.10) and (1.5.11) for commutators.

Of fundamental importance are the PB between the q's and the p's. Observe that

$$\{q_i, q_j\} = \{p_i, p_j\} = 0 \tag{2.7.4a}$$

$$\{q_i, p_j\} = \delta_{ij} \tag{2.7.4b}$$

since $(q_i, \ldots, p_n)$ are independent variables ($\partial q_i/\partial q_j = \delta_{ij}$, $\partial q_i/\partial p_k = 0$, etc.). Hamilton's equations may be written in terms of PB as

$$\dot{q}_i = \{q_i, \mathscr{H}\} \tag{2.7.5a}$$

$$\dot{p}_i = \{p_i, \mathscr{H}\} \tag{2.7.5b}$$

by setting $\omega = q_i$ or p_i in Eq. (2.7.2).

*Exercise 2.7.2.** (i) Verify Eqs. (2.7.4) and (2.7.5). (ii) Consider a problem in two dimensions given by $\mathscr{H} = p_x^2 + p_y^2 + ax^2 + by^2$. Argue that if $a = b$, $\{l_z, \mathscr{H}\}$ must vanish. Verify by explicit computation.

Canonical Transformations

We have seen that the Euler–Lagrange equations are form invariant under an arbitrary‡ change of coordinates in configuration space

$$q_i \to \bar{q}_i(q_1, \ldots, q_n), \qquad i = 1, \ldots, n \tag{2.7.6a}$$

‡ We assume the transformation is invertible, so we may write q in terms of $\bar{q}$: $q = q(\bar{q})$. The transformation may also depend on time explicitly [$\bar{q} = q(q, t)$], but we do not consider such cases.

or more succinctly

$$q \to \bar{q}(q) \tag{2.7.6b}$$

The response of the velocities to this transformation follows from Eq. (2.7.6a):

$$\dot{\bar{q}}_i = \bar{\dot{q}}_i = \frac{d\bar{q}_i}{dt} = \sum_j \left(\frac{\partial \bar{q}_i}{\partial q_j}\right) \dot{q}_j \tag{2.7.7}$$

The response of the canonical momenta may be found by rewriting $\mathscr{L}$ in terms of $(\bar{q}, \dot{\bar{q}})$ and taking the derivative with respect to $\dot{\bar{q}}$:

$$\bar{p}_i = \frac{\partial \mathscr{L}(\bar{q}, \dot{\bar{q}})}{\partial \dot{\bar{q}}_i} \tag{2.7.8}$$

The result is (Exercise 2.7.8):

$$\bar{p}_i = \sum_j \left(\frac{\partial q_j}{\partial \bar{q}_i}\right) p_j \tag{2.7.9}$$

Notice that although $\mathscr{L}$ enters Eq. (2.7.8), it drops out in Eq. (2.7.9), which connects $\bar{p}$ to the old variables. This is as it should be, for we expect that the response of the momenta to a coordinate transformation (say, a rotation) is a purely kinematical question.

A word of explanation about $\mathscr{L}(\bar{q}, \dot{\bar{q}})$. By $\mathscr{L}(\bar{q}, \dot{\bar{q}})$ we mean the Lagrangian (say $T-V$, for definiteness) written in terms of $\bar{q}$ and $\dot{\bar{q}}$. Thus the numerical value of the Lagrangian is unchanged under $(q, \dot{q}) \to (\bar{q}, \dot{\bar{q}})$; for $(q, \dot{q})$ and $(\bar{q}, \dot{\bar{q}})$ refer to the *same physical state*. The functional form of the Lagrangian, however, *does* change and so we should really be using two different symbols $\mathscr{L}(q, \dot{q})$ and $\bar{\mathscr{L}}(\bar{q}, \dot{\bar{q}})$. Nonetheless we follow the convention of denoting a given dynamical variable, such as the Lagrangian, by a fixed symbol in all coordinate systems.

The invariance of the Euler–Lagrange equations under $(q, \dot{q}) \to (\bar{q}, \dot{\bar{q}})$ implies the invariance of Hamilton's equation under $(q, p) \to (\bar{q}, \bar{p})$, i.e., $(\bar{q}, \bar{p})$ obey

$$\dot{\bar{q}}_i = \partial\mathscr{H}/\partial\bar{p}_i, \qquad \dot{\bar{p}}_i = -(\partial\mathscr{H}/\partial\bar{q}_i) \tag{2.7.10}$$

where $\mathscr{H} = \mathscr{H}(\bar{q}, \bar{p})$ is the Hamiltonian written in terms of $\bar{q}$ and $\bar{p}$. The proof is simple: we start with $\mathscr{L}(\bar{q}, \dot{\bar{q}})$, perform a Legendre transform, and use the fact that $\bar{q}$ obeys Euler–Lagrange equations.

The transformation

$$q_i \to \bar{q}_i(q_1, \ldots, q_n), \qquad \bar{p}_i = \sum_j \left(\frac{\partial q_j}{\partial \bar{q}_i}\right) p_j \tag{2.7.11}$$

is called a *point transformation*. If we view the Hamiltonian formalism as something derived from the Lagrangian scheme, which is formulated in n-dimensional configuration space, this is the most general (time-independent) transformation which preserves the form of Hamilton's equations (that we can think of). On the other hand, if we view the Hamiltonian formalism in its own right, the backdrop is the $2n$-dimensional phase space. In this space, the point transformation is unnecessarily restrictive. One can contemplate a more general transformation of phase space coordinates:

$$\begin{aligned} q &\to \bar{q}(q, p) \\ p &\to \bar{p}(q, p) \end{aligned} \tag{2.7.12}$$

Although all sets of $2n$ independent coordinates $(\bar{q}, \bar{p})$ are formally adequate for describing the state of the system, not all of them will preserve the canonical form of Hamilton's equations. (This is like saying that although Newton's laws may be written in terms of any complete set of coordinates, the simple form $m\ddot{q}_i = -\partial V/\partial q_i$ is valid only if the q_i are Cartesian). If, however, $(\bar{q}, \bar{p})$ obey the canonical equations (2.7.10), we say that they are *canonical coordinates* and that Eq. (2.7.12) defines a *canonical transformation*. Any set of coordinates $(q_1, \ldots, q_n)$, and the corresponding momenta generated in the Lagrangian formalism $(p_i = \partial\mathscr{L}/\partial \dot{q}_i)$, are canonical coordinates. Given one set, (q, p), we can get another, $(\bar{q}, \bar{p})$, by the point transformation, which is a special case of the canonical transformation. This does not, however, exhaust the possibilities. Let us now ask the following question. *Given a new set of coordinates* $(\bar{q}(q, p), \bar{p}(q, p))$, *how can we tell if they are canonical* [*assuming* (q, p) *are*]*?* Now it is true for any $\omega(q, p)$ that

$$\dot{\omega} = \{\omega, \mathscr{H}\} = \sum_i \left(\frac{\partial \omega}{\partial q_i} \frac{\partial \mathscr{H}}{\partial p_i} - \frac{\partial \omega}{\partial p_i} \frac{\partial \mathscr{H}}{\partial q_i} \right) \tag{2.7.13}$$

Applying this to $\bar{q}_j(q, p)$ we find

$$\dot{\bar{q}}_j = \sum_i \left(\frac{\partial \bar{q}_j}{\partial q_i} \frac{\partial \mathscr{H}}{\partial p_i} - \frac{\partial \bar{q}_j}{\partial p_i} \frac{\partial \mathscr{H}}{\partial q_i} \right) \tag{2.7.14}$$

If we view $\mathscr{H}$ as a function of $(\bar{q}, \bar{p})$ and use the chain rule, we get

$$\frac{\partial \mathscr{H}(q, p)}{\partial p_i} = \frac{\partial \mathscr{H}(\bar{q}, \bar{p})}{\partial p_i} = \sum_k \left(\frac{\partial \mathscr{H}}{\partial \bar{q}_k} \frac{\partial \bar{q}_k}{\partial p_i} + \frac{\partial \mathscr{H}}{\partial \bar{p}_k} \frac{\partial \bar{p}_k}{\partial p_i} \right) \tag{2.7.15a}$$

and

$$\frac{\partial \mathscr{H}(q, p)}{\partial q_i} = \frac{\partial \mathscr{H}(\bar{q}, \bar{p})}{\partial q_i} = \sum_k \left(\frac{\partial \mathscr{H}}{\partial \bar{q}_k} \frac{\partial \bar{q}_k}{\partial q_i} + \frac{\partial \mathscr{H}}{\partial \bar{p}_k} \frac{\partial \bar{p}_k}{\partial q_i} \right) \tag{2.7.15b}$$

Feeding all this into Eq. (2.7.14) we find, upon regrouping terms,

$$\dot{\bar{q}}_j = \sum_k \left(\frac{\partial \mathscr{H}}{\partial \bar{q}_k} \{\bar{q}_j, \bar{q}_k\} + \frac{\partial \mathscr{H}}{\partial \bar{p}_k} \{\bar{q}_j, \bar{p}_k\} \right) \tag{2.7.16}$$

It can similarly be established that

$$\dot{\bar{p}}_j = \sum_k \left(\frac{\partial \mathscr{H}}{\partial \bar{q}_k} \{\bar{p}_j, \bar{q}_k\} + \frac{\partial \mathscr{H}}{\partial \bar{p}_k} \{\bar{p}_j, \bar{p}_k\} \right) \tag{2.7.17}$$

If Eqs. (2.7.16) and (2.7.17) are to reduce to the canonical equations (2.7.10) for any $\mathscr{H}(q, p)$, we must have

$$\begin{aligned} \{\bar{q}_j, \bar{q}_k\} &= 0 = \{\bar{p}_j, \bar{p}_k\} \\ \{\bar{q}_j, \bar{p}_k\} &= \delta_{jk} \end{aligned} \tag{2.7.18}$$

These then are the conditions to be satisfied by the new variables if they are to be canonical. Notice that these constraints make no reference to the specific functional form of $\mathscr{H}$: the equations defining canonical variables are purely kinematical and true for any $\mathscr{H}(q, p)$.

Exercise 2.7.3. Fill in the missing steps leading to Eq. (2.7.18) starting from Eq. (2.7.14).

Exercise 2.7.4. Verify that the change to a rotated frame

$$\bar{x} = x \cos\theta - y \sin\theta$$

$$\bar{y} = x \sin\theta + y \cos\theta$$

$$\bar{p}_x = p_x \cos\theta - p_y \sin\theta$$

$$\bar{p}_y = p_x \sin\theta + p_y \cos\theta$$

is a canonical transformation.

Exercise 2.7.5. Show that the polar variables $\rho = (x^2 + y^2)^{1/2}$, $\phi = \tan^{-1}(y/x)$,

$$p_\rho = \hat{e}_\rho \cdot \mathbf{p} = \frac{x p_x + y p_y}{(x^2 + y^2)^{1/2}}, \qquad p_\phi = x p_y - y p_x (= l_z)$$

are canonical. ($\hat{e}_\rho$ is the unit vector in the radial direction.)

*Exercise 2.7.6.** Verify that the change from the variables $\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2$ to $\mathbf{r}_{CM}, \mathbf{p}_{CM}, \mathbf{r}$, and $\mathbf{p}$ is a canonical transformation. (See Exercise 2.5.4).

Exercise 2.7.7. Verify that

$$\bar{q} = \ln(q^{-1} \sin p)$$

$$\bar{p} = q \cot p$$

is a canonical transformation.

Exercise 2.7.8. We would like to derive here Eq. (2.7.9), which gives the transformation of the momenta under a coordinate transformation in configuration space:

$$q_i \to \bar{q}_i(q_1, \ldots, q_n)$$

(1) Argue that if we invert the above equation to get $q = q(\bar{q})$, we can derive the following counterpart of Eq. (2.7.7):

$$\dot{q}_i = \sum_j \frac{\partial q_i}{\partial \bar{q}_j} \dot{\bar{q}}_j$$

(2) Show from the above that

$$\left(\frac{\partial \dot{q}_i}{\partial \dot{\bar{q}}_j}\right)_{\bar{q}} = \frac{\partial q_i}{\partial \bar{q}_j}$$

(3) Now calculate

$$\bar{p}_i = \left[\frac{\partial \bar{\mathscr{L}}(\bar{q}, \dot{\bar{q}})}{\partial \dot{\bar{q}}_i}\right]_{\bar{q}} = \left[\frac{\partial \mathscr{L}(q, \dot{q})}{\partial \dot{\bar{q}}_i}\right]_{\bar{q}}$$

Use the chain rule and the fact that $q = q(\bar{q})$ and not $q(\bar{q}, \dot{\bar{q}})$ to derive Eq. (2.7.9).

(4) Verify, by calculating the PB in Eq. (2.7.18), that the point transformation is canonical.

If (q, p) and $(\bar{q}, \bar{p})$ are both canonical, we must give them both the same status, for Hamilton's equations have the same appearance when expressed in terms of either set. Now, we have defined the PB of two variables ω and σ in terms of (q, p) as

$$\{\omega, \sigma\} = \sum_i \left(\frac{\partial \omega}{\partial q_i}\frac{\partial \sigma}{\partial p_i} - \frac{\partial \omega}{\partial p_i}\frac{\partial \sigma}{\partial q_i}\right) \equiv \{\omega, \sigma\}_{q,p}$$

Should we not also define a PB, $\{\omega, \sigma\}_{\bar{q},\bar{p}}$ for every canonical pair $(\bar{q}, \bar{p})$? Fortunately it turns out that *the PB are invariant under canonical transformations*:

$$\{\omega, \sigma\}_{q,p} = \{\omega, \sigma\}_{\bar{q},\bar{p}} \qquad (2.7.19)$$

(It is understood that ω and σ are written as functions of $\bar{q}$ and $\bar{p}$ on the right-hand side.)

Exercise 2.7.9. Verify Eq. (2.7.19) by direct computation. Use the chain rule to go from q, p derivatives to $\bar{q}, \bar{p}$ derivatives. Collect terms that represent PB of the latter.

Besides the proof by direct computation (as per Exercise 2.7.9 above) there is an alternate way to establish Eq. (2.7.19).

Consider first $\sigma = \mathscr{H}$. We know that since (q, p) obey canonical equations,

$$\dot{\omega} = \{\omega, \mathscr{H}\}_{q,p}$$

But then $(\bar{q}, \bar{p})$ also obey canonical equations, so

$$\dot{\omega} = \{\omega, \mathscr{H}\}_{\bar{q},\bar{p}}$$

Now ω is some physical quantity such as the kinetic energy or the component of angular momentum in some fixed direction, so its rate of change is independent of the phase space coordinates used, i.e., $\dot{\omega}$ is $\dot{\omega}$, whether $\omega = \omega(q, p)$ or $\omega(\bar{q}, \bar{p})$. So

$$\{\omega, \mathscr{H}\}_{q,p} = \{\omega, \mathscr{H}\}_{\bar{q},\bar{p}} \tag{2.7.20}$$

Having proved the result for what seems to be the special case $\sigma = \mathscr{H}$, we now pull the following trick. Note that nowhere in the derivation did we have to assume that $\mathscr{H}$ was any particular function of q and p. In fact, Hamiltonian dynamics, as a consistent mathematical scheme, places no restriction on $\mathscr{H}$. It is the physical requirement that the time evolution generated by $\mathscr{H}$ coincide with what is *actually* observed, that restricts $\mathscr{H}$ to be $T+V$. Thus $\mathscr{H}$ could have been any function at all in the preceding argument and in the result Eq. (2.7.20) (which is just a relation among partial derivatives.) If we understand that $\mathscr{H}$ is not $T+V$ *in this argument* but an arbitrary function, call it σ, we get the desired result.

Active Transformations

So far, we have viewed the transformation

$$\bar{q} = \bar{q}(q, p)$$

$$\bar{p} = \bar{p}(q, p)$$

as passive: both (q, p) and $(\bar{q}, \bar{p})$ refer to the same point in phase space described in two different coordinate systems. Under the transformation $(q, p) \to (\bar{q}, \bar{p})$, the numerical values of all dynamical variables are unchanged (for we are talking about the same physical state), but their functional form is changed. For instance, under a change from Cartesian to spherical coordinates, $\omega(x, y, z) = x^2 + y^2 + z^2 \to \omega(r, \theta, \phi) = r^2$. As mentioned earlier, we use the same symbol for a given variable even if its functional dependence on the coordinates changes when we change coordinates.

Consider now a restricted class of transformations, called *regular transformations*, which preserve the range of the variables: (q, p) and $(\bar{q}, \bar{p})$ have the same range. A change from one Cartesian coordinate to a translated or rotated one is regular (each variable goes from $-\infty$ to $+\infty$ before and after), whereas a change to spherical coordinates (where some coordinates are nonnegative, some are bounded by 2π, etc.) is not.

A regular transformation $(q, p) \to (\bar{q}, \bar{p})$ permits an alternate interpretation: instead of viewing $(\bar{q}, \bar{p})$ as the same phase space point in a new coordinate system, we may view it as a new point in the same coordinate system. This corresponds to an active transformation which changes the state of the system. Under this change, the numerical value of any dynamical variable $\omega(q, p)$ will generally change: $\omega(q, p) \neq \omega(\bar{q}, \bar{p})$, though its functional dependence will not: $\omega(\bar{q}, \bar{p})$ is the same function $\omega(q, p)$ evaluated at the new point $(q = \bar{q}, p = \bar{p})$.

We say that ω is *invariant* under the regular transformation $(q, p) \to (\bar{q}, \bar{p})$ if

$$\omega(q, p) = \omega(\bar{q}, \bar{p}) \tag{2.7.21}$$

(This equation has content only if we are talking about the active transformations, for it is true for any ω under a passive transformation.)

Whether we view the transformation $(q, p) \to (\bar{q}, \bar{p})$ as active or passive, it is called canonical if $(\bar{q}, \bar{p})$ obey Eq. (2.7.18). As we shall see, only regular *canonical* transformations are physically interesting.

2.8. Symmetries and Their Consequences

Let us begin our discussion by examining what the word "symmetry" means in daily usage. We say that a sphere is a very symmetric object because it looks the same when seen from many directions. Or, equivalently, a sphere looks the same before and after it is subjected to a rotation around *any* axis passing through its center. A cylinder has symmetry too, but not as much: the rotation must be performed around its axis. Generally then, the symmetry of an object implies its invariance under some transformations, which in our examples are rotations.

A symmetry can be discrete or continuous, as illustrated by the example of a hexagon and a circle. While the rotation angles that leave a hexagon unchanged form a discrete set, namely, multiples of 60°, the corresponding set for a circle is a continuum. We may characterize the continuous symmetry of the circle in another way. Consider the *identity transformation*, which does nothing, i.e., rotates by 0° in our example. This leaves both the circle and the hexagon invariant. Consider next an *infinitesimal transformation*, which is infinitesimally "close" to the identity; in our example this is a rotation by an infinitesimal angle ε. The infinitesimal rotation leaves the circle invariant but not the hexagon. The circle is thus characterized by its invariance under infinitesimal rotations. Given this property, its invariance under finite rotations follows, for any finite rotation may be viewed as a sequence of infinitesimal rotations (each of which leaves it invariant).

It is also possible to think of functions of some variables as being symmetric in the sense that if one changes the values of the variables in a certain way, the value of the function is invariant. Consider for example

$$f(x, y) = x^2 + y^2$$

If we make the following change

$$\begin{aligned} x \to \bar{x} &= x \cos\theta - y \sin\theta \\ y \to \bar{y} &= x \sin\theta + y \cos\theta \end{aligned} \tag{2.8.1}$$

in the arguments, we find that f is invariant. We say that f is symmetric under the above transformation. In the terminology introduced earlier, the transformation in question is continuous: its infinitesimal version is

$$\begin{aligned} x \to \bar{x} &= x \cos\varepsilon - y \sin\varepsilon = x - y\varepsilon \\ y \to \bar{y} &= x \sin\varepsilon + y \cos\varepsilon = x\varepsilon + y \qquad (\textit{to order } \varepsilon) \end{aligned} \tag{2.8.2}$$

Consider now the function $\mathscr{H}(q, p)$. There are two important dynamical consequences that follow from its invariance under *regular canonical* transformations.

I. If $\mathscr{H}$ is invariant under the following *infinitesimal* transformation (which you may verify is canonical, Exercise 2.8.2),

$$\begin{aligned} q_i \to \bar{q}_i &= q_i + \varepsilon \frac{\partial g}{\partial p_i} \equiv q_i + \delta q_i \\ p_i \to \bar{p}_i &= p_i - \varepsilon \frac{\partial g}{\partial q_i} \equiv p_i + \delta p_i \end{aligned} \tag{2.8.3}$$

where $g(q, p)$ is any dynamical variable, *then g is conserved*, i.e., a constant of motion. One calls g the *generator of the transformation.*

II. If $\mathscr{H}$ is invariant under the regular, canonical, but not necessarily infinitesimal, transformation $(q, p) \to (\bar{q}, \bar{p})$, and if $(q(t), p(t))$ is a solution to the equations of motion, so is the transformed (translated, rotated, etc.) trajectory, $(\bar{q}(t), \bar{p}(t))$.

Let us now analyze these two consequences.

Consequence I. Let us first verify that g is indeed conserved if $\mathscr{H}$ is invariant under the transformation it generates. Working to first order in ε, if we equate the change in $\mathscr{H}$ under the change of its arguments to zero, we get

$$\delta\mathscr{H} = \sum_i \frac{\partial \mathscr{H}}{\partial q_i}\left(\varepsilon \frac{\partial g}{\partial p_i}\right) + \frac{\partial \mathscr{H}}{\partial p_i}\left(-\varepsilon \frac{\partial g}{\partial q_i}\right) = \varepsilon\{\mathscr{H}, g\} = 0 \tag{2.8.4}$$

But according to Eq. (2.7.2),

$$\{g, \mathscr{H}\} = 0 \to g \text{ is conserved} \tag{2.8.5}$$

(More generally, the response of any variable ω to the transformation is

$$\delta\omega = \varepsilon\{\omega, g\} \tag{2.8.6}$$

Note that δp and δq in Eq. (2.8.3) may also be written as PBs.) Consider as an example, a particle in one dimension and the case $g=p$. From Eq. (2.8.3),

$$\begin{aligned}\delta x &= \varepsilon \frac{\partial p}{\partial p} = \varepsilon \\ \delta p &= -\varepsilon \frac{\partial p}{\partial x} = 0\end{aligned} \tag{2.8.7}$$

which we recognize to be an infinitesimal translation. Thus the linear momentum p is the generator of spatial translations and is conserved in a translationally invariant problem. The physics behind this result is clear. Since p is unchanged in a translation, so is $T=p^2/2m$. Consequently $V(x+\varepsilon)=V(x)$. But if the potential doesn't vary from point to point, there is no force and p is conserved.

Next consider an example from two dimensions with $g=l_z=xp_y-yp_x$. Here,

$$\begin{aligned}\delta x &= -y\varepsilon \left(=\varepsilon \frac{\partial l_z}{\partial p_x}\right) \\ \delta y &= x\varepsilon \left(=\varepsilon \frac{\partial l_z}{\partial p_y}\right) \\ \delta p_x &= -p_y\varepsilon \left(=-\varepsilon \frac{\partial l_z}{\partial x}\right) \\ \delta p_y &= p_x\varepsilon \left(=-\varepsilon \frac{\partial l_z}{\partial y}\right)\end{aligned} \tag{2.8.8}$$

which we recognize to be an infinitesimal rotation around the z axis, [Eq. (2.8.2)]. Thus the angular momentum around the z axis is the generator of rotations around that axis, and is conserved if $\mathscr{H}$ is invariant under rotations of the state around that axis. The relation between the symmetry and the conservation law may be understood in the following familiar terms. Under the rotation of the coordinates and the momenta, $|\mathbf{p}|$ doesn't change and so neither does $T=|\mathbf{p}|^2/2m$. Consequently, V is a constant as we go along any circle centered at the origin. This in turn means that there is no force in the tangential direction and so no torque around the z axis. The conservation of l_z then follows.

Exercise 2.8.1. Show that $p=p_1+p_2$, the total momentum, is the generator of infintesimal translations for a two-particle system.

*Exercise 2.8.2.** Verify that the infinitesimal transformation generated by any dynamical variable g is a canonical transformation. (Hint: Work, as usual, to first order in ε.)

Exercise 2.8.3. Consider

$$\mathscr{H}=\frac{p_x^2+p_y^2}{2m}+\frac{1}{2}m\omega^2(x^2+y^2)$$

whose invariance under the rotation of the coordinates *and* momenta leads to the conservation of l_z. But $\mathscr{H}$ is also invariant under the rotation of *just the coordinates*. Verify that this is a *noncanonical* transformation. Convince yourself that in this case it is not possible to write $\delta\mathscr{H}$ as $\varepsilon\{\mathscr{H}, g\}$ for any g, i.e., that no conservation law follows.

*Exercise 2.8.4.** Consider $\mathscr{H} = \frac{1}{2}p^2 + \frac{1}{2}x^2$, which is invariant under infinitesimal rotations in *phase space* (the x–p plane). Find the generator of this transformation (after verifying that it is canonical). (You could have guessed the answer based on Exercise 2.5.2.).

The preceding analysis yields, as a by-product, a way to generate infinitesimal canonical transformations. We take any function $g(q, p)$ and obtain the transformation given by Eq. (2.8.6). (Recall that although we defined a canonical transformation earlier, until now we had no means of generating one.) Given an infinitesimal canonical transformation, we can get a finite one by "integrating" it. The following examples should convince you that this is possible. Consider the transformation generated by $g = \mathscr{H}$. We have

$$\begin{aligned} \delta q_i &= \varepsilon\{q_i, \mathscr{H}\} \\ \delta p_i &= \varepsilon\{p_i, \mathscr{H}\} \end{aligned} \tag{2.8.9}$$

But we know from the equations of motion that $\dot{q}_i = \{q_i, \mathscr{H}\}$ etc. So

$$\begin{aligned} \delta q_i &= \varepsilon \dot{q}_i \\ \delta p_i &= \varepsilon \dot{p}_i \end{aligned} \tag{2.8.10}$$

Thus the new point in phase space $(\bar{q}, \bar{p}) = (q + \delta q, p + \delta p)$ obtained by this canonical transformation of (q, p) is just the point to which (q, p) would move in an infinitesimal time interval ε. In other words, the motion of points in phase space under the time evolution generated by $\mathscr{H}$ is an active canonical transformation. Now, you know that by integrating the equations of motion, we can find $(\bar{q}, \bar{p})$ at any future time, i.e., get the finite canonical transformation. Consider now a general case of $g \neq \mathscr{H}$. We still have

$$\begin{aligned} \delta q_i &= \varepsilon\{q_i, g\} \\ \delta p_i &= \varepsilon\{p_i, g\} \end{aligned} \tag{2.8.11}$$

Mathematically, these equations are identical to Eq. (2.8.9), with g playing the role of the Hamiltonian. Clearly there should be no problem integrating these equations for the evolution of the phase space points under the "fake" Hamiltonian g, and fake "time" ε. Let us consider for instance the case $g = l_z$ which has units erg sec and the corresponding fake time $\varepsilon = \delta\theta$, an angle. The transformation of the coordinates is

$$\begin{aligned} \delta x &= \varepsilon\{x, l_z\} = -\varepsilon y \equiv (-\delta\theta)y \\ \delta y &= (\delta\theta)x \end{aligned} \tag{2.8.12}$$

The fake equations of motion are

$$\frac{dx}{d\theta}=-y, \qquad \frac{dy}{d\theta}=x \tag{2.8.13}$$

Differentiating first with respect to θ, and using the second, we get

$$\frac{d^2x}{d\theta^2}+x=0$$

and likewise,

$$\frac{d^2y}{d\theta^2}+y=0$$

So

$$x=A\cos\theta+B\sin\theta$$

$$y=C\sin\theta+D\cos\theta$$

We find the constants from the "initial" ($\theta=0$) coordinates and "velocities": $A=x_0$, $D=y_0$, $B=(\partial x/\partial\theta)_0=-y_0$, $C=(\partial y/\partial\theta)_0=x_0$. Reverting to the standard notation in which (x, y), rather than (x_0, y_0), labels the initial point and $(\bar{x}, \bar{y})$, rather than (x, y), denotes the transformed one, we may write the finite canonical transformation (a finite rotation) as

$$\begin{aligned}\bar{x}&=x\cos\theta-y\sin\theta\\ \bar{y}&=x\sin\theta+y\cos\theta\end{aligned} \tag{2.8.14}$$

Similar equations may be derived for $\bar{p}_x$ and $\bar{p}_y$ in terms of p_x and p_y.

Although a wide class of canonical transformations is now open to us, there are many that aren't. For instance, $(q, p)\to(-q, -p)$ is a discrete canonical transformation that has no infinitesimal version. There are also the transformations that are not regular, such as the change from Cartesian to spherical coordinates, which have neither infintesimal forms, nor an active interpretation. We do not consider ways of generating these.‡

Consequence II. Let us understand the content of this result through an example before turning to the proof. Consider a two-particle system whose Hamiltonian is invariant under the translation of the entire system, i.e., both particles. Let an observer S_A prepare, at $t=0$, a state $(x_1^0, x_2^0; p_1^0, p_2^0)$ which evolves as $(x_1(t), x_2(t); p_1(t), p_2(t))$ for some time and ends up in the state $(x_1^T, x_2^T; p_1^T, p_2^T)$ at time T. Let

‡ For an excellent and lucid treatment of this question and many other topics in advanced classical mechanics, see H. Goldstein, *Classical Mechanics*, Addison-Wesley, Reading, Massachusetts (1950); E. C. G. Sudharshan and N. Mukunda, *Classical Dynamics: A Modern Perspective*, Wiley, New York (1974).

us call the final state the outcome of the experiment conducted by S_A. We are told that as a result of the translational invariance of $\mathscr{H}$, any other trajectory that is related to this by an arbitrary translation a is also a solution to the equations of motion. In this case, the initial state, for example, is $(x_1^0+a, x_2^0+a; p_1^0, p_2^0)$. *The final state and all intermediate states are likewise displaced by the same amount.* To an observer S_B, displaced relative to S_A by an amount a, the evolution of the second system will appear to be identical to what S_A saw in the first. Assuming for the sake of this argument that S_B had in fact prepared the second system, we may say that a given experiment and its translated version will give the same result (as seen by the observers who conducted them) if $\mathscr{H}$ is translationally invariant.

The physical idea is the following. For the usual reasons, translational invariance of $\mathscr{H}$ implies the invariance of $V(x_1, x_2)$. This in turn means that $V(x_1, x_2)=V(x_1-x_2)$. Thus each particle cares only about where the other is relative to it, and not about where the system as a whole is in space. Consequently the outcome of the experiment is not affected by an overall translation.

Consequence II is just a generalization of this result to other canonical transformations that leave $\mathscr{H}$ invariant. For instance, if $\mathscr{H}$ is rotationally invariant, a given experiment and its rotated version will give the same result (according to the observers who conducted them).

Let us now turn to the proof of the general result.

Proof. Imagine a trajectory $(q(t), p(t))$ in phase space that satisfies the equations of motion. Let us associate with it an image trajectory, $(\bar{q}(t), \bar{p}(t))$, which is obtained by transforming each point (q, p) to the image point $(\bar{q}, \bar{p})$ by means of a regular canonical transformation. We ask if the image point moves according to Hamilton's equation of motion, i.e., if

$$\dot{\bar{q}}_j=\frac{\partial\mathscr{H}(\bar{q},\bar{p})}{\partial\bar{p}_j}, \qquad \dot{\bar{p}}_j=-\frac{\partial\mathscr{H}(\bar{q},\bar{p})}{\partial\bar{q}_j} \tag{2.8.15}$$

if $\mathscr{H}$ is invariant under the transformation $(q, p)\to(\bar{q}, \bar{p})$. Now $\bar{q}_j(q, p)$, like any dynamical variable $\omega(q, p)$, obeys

$$\dot{\bar{q}}_j=\{\bar{q}_j, \mathscr{H}(q, p)\}_{q,p} \tag{2.8.16}$$

If $(q, p)\to(\bar{q}, \bar{p})$ were a *passive* canonical transformation, we could write, since the PB are invariant under such a transformation,

$$\dot{\bar{q}}_j=\{\bar{q}_j, \mathscr{H}(q, p)\}_{q,p}=\{\bar{q}_j, \mathscr{H}(\bar{q}, \bar{p})\}_{\bar{q},\bar{p}}=\frac{\partial\mathscr{H}(\bar{q},\bar{p})}{\partial\bar{p}_j}$$

But it is an active transformation. However, *because of the symmetry of* $\mathscr{H}$, i.e., $\mathscr{H}(q, p)=\mathscr{H}(\bar{q}, \bar{p})$, we can go through the very same steps that led to Eq. (2.7.16) from Eq. (2.7.14) and prove the result. If you do not believe this, you may verify it by explicit computation using $\mathscr{H}(q, p)=\mathscr{H}(\bar{q}, \bar{p})$. A similar argument shows that

$$\dot{\bar{p}}_j=-\frac{\partial \mathscr{H}(\bar{q}, \bar{p})}{\partial \bar{q}_j} \tag{2.8.17}$$

So the image point moves according to Hamilton's equations. Q.E.D.

Exercise 2.8.5. Why is it that a *non*canonical transformation that leaves $\mathscr{H}$ invariant does not map a solution into another? Or, in view of the discussions on consequence II, why is it that an experiment and its transformed version do not give the same result when the transformation that leaves $\mathscr{H}$ invariant is not canonical? It is best to consider an example. Consider the potential given in Exercise 2.8.3. Suppose I release a particle at $(x=a, y=0)$ with $(p_x=b, p_y=0)$ and you release one in the transformed state in which $(x=0, y=a)$ and $(p_x=b, p_y=0)$, i.e., you rotate the coordinates but not the momenta. This is a noncanonical transformation that leaves $\mathscr{H}$ invariant. Convince yourself that at later times the states of the two particles are not related by the same transformation. Try to understand what goes wrong in the general case.

As you go on and learn quantum mechanics, you will see that the symmetries of the Hamiltonian have similar consequences for the dynamics of the system.

A Useful Relation Between S and E

We now prove a result that will be invoked in Chapter 16:

$$\frac{\partial S_{\text{cl}}(x_f, t_f; x_i, t_i)}{\partial t_f}=-\mathscr{H}(t_f)$$

where $S_{\text{cl}}(x_f, t_f; x_i, t_i)$ is the action of the classical path from x_i, t_i to x_f, t_f and $\mathscr{H}$ is the Hamiltonian at the upper end point. Since we shall be working with problems where energy is conserved we may write

$$\frac{\partial S_{\text{cl}}(x_f, t_f; x_i, t_i)}{\partial t_f}=-E \tag{2.8.18}$$

where E is the conserved energy, constant on the whole trajectory.

At first sight you may think that since

$$S_{\text{cl}}=\int_{t_i}^{t_f} \mathscr{L}\, dt$$

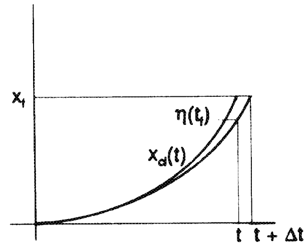


Figure 2.5. The upper trajectory takes time t while the lower takes $t+\Delta t$.

the right side must equal $\mathscr{L}$ and not $-E$. The explanation requires Fig. 2.5 wherein we have set $x_i=t_i=0$ for convenience.

The derivative we are computing is governed by the change in action of the *classical path* due to a change in travel by Δt holding the end points x_i and x_f fixed. From the figure it is clear that now the particle takes a different classical trajectory

$$x(t)=x_{\rm cl}\,(t)+\eta(t) \quad \text{with} \quad \eta(0)=0.$$

so that the total change in action comes from the difference in paths between $t=0$ and $t=t_f$ as well as the entire action due to the extra travel between t_f and $t_f+\Delta t_f$. Only the latter is given $\mathscr{L}\Delta t$. The correct answer is then

$$\begin{aligned}\delta S_{\rm cl}&=\int_0^{t_f}\left[\frac{\partial\mathscr{L}}{\partial x}\,\eta(t)+\frac{\partial\mathscr{L}}{\partial\dot{x}}\,\dot{\eta}(t)\right]dt+\mathscr{L}(t_f)\,\Delta t\\
&=\int_0^{t_f}\left(-\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial\dot{x}}+\frac{\partial\mathscr{L}}{\partial x}\right)_{x_{\rm cl}}\eta(t)\,dt+\int_0^{t_f}\frac{d}{dt}\left[\frac{\partial\mathscr{L}}{\partial\dot{x}}\,\eta(t)\right]dt+\mathscr{L}(t_f)\,\Delta t\\
&=0+\frac{\partial\mathscr{L}}{\partial\dot{x}}\,\eta(t)\bigg|_{t_f}+\mathscr{L}(t_f)\,\Delta t.\end{aligned}$$

It is clear from the figure that $\eta(t_f)=-\dot{x}(t_f)\,\Delta t$ so that

$$\delta S=\left[-\frac{\partial\mathscr{L}}{\partial\dot{x}}\,\dot{x}+\mathscr{L}\right]_{t_f}\Delta t=-\mathscr{H}(t_f)\,\Delta t$$

from which the result follows.

Exercise 2.8.6. Show that $\partial S_{\rm cl}/\partial x_f=p(t_f)$.

Exercise 2.8.7. Consider the harmonic oscillator, for which the general solution is

$$x(t)=A\cos\omega t+B\sin\omega t.$$

Express the energy in terms of A and B and note that it does not depend on time. Now choose A and B such that $x(0)=x_1$ and $x(T)=x_2$. Write down the energy in terms of x_1, x_2, and T. Show that the action for the trajectory connecting x_1 and x_2 is

$$S_{\mathrm{cl}}(x_1, x_2, T)=\frac{m\omega}{2\sin\omega T}[(x_1^2+x_2^2)\cos\omega T-2x_1x_2].$$

Verify that $\partial S_{\mathrm{cl}}/\partial T=-E$.