L. D. Hand and J. D. Finch, Analytical Dynamics (Cambridge, 1998

We could speculate on whether the converse of Noether's Theorem holds. For example, the Laplace–Runge–Lenz vector in the Kepler problem* is a conserved quantity. Does this imply a continuous symmetry heretofore unknown to us? It does: It is called the "$O(4)$ symmetry." Like the Lenz vector example, does the existence of conservation laws necessarily imply hidden symmetries of the Lagrangian?

5.3 HAMILTONIAN DYNAMICS

Up to now, we have used a phase space in which we track the development of $q_k(t)$ and $\dot{q}_k(t)$ in time. It will be convenient to use a more general definition of phase space, one that contains the coordinates $q_k(t)$ and the *canonically conjugate momenta* $p_k(t)$, which are defined for the general case of N degrees of freedom by:

$$\boxed{\begin{gathered} p_k \equiv \frac{\partial L}{\partial \dot{q}_k} \\ (k = 1, \ldots, N). \end{gathered}} \tag{5.15}$$

Thus q_k and p_k will become the basic dynamical variables instead of q_k and $\dot{q}_k$. According to the definition (5.15), p_k is a function of $q_k, \dot{q}_k$, which are themselves functions of the time via the equations of motion.

The Lagrangian $L(q, \dot{q}, t)$ is replaced by the *Hamiltonian* $H(q, p, t)$. The Euler–Lagrange equations which determine the motion of the system are replaced by *Hamilton's equations*. Not only does this procedure lead to symmetric equations involving the dynamical variables q_k and p_k, but a whole new approach to classical mechanics is introduced, one that leads to the most powerful and sophisticated tools of theoretical physics. The concept of canonical momentum is the key concept in Hamilton's theory. You must be warned that momentum can lose its familiar definition: $\vec{p} = m\vec{v}$. It will turn out that this is still true in many simple cases, but it is often not true when generalized coordinates are used for convenience in solving a problem.

Although we defined canonical momentum in Chapter 1, let us begin again from the beginning. Our goal is to find a quantity $p(q, \dot{q})$ that is dynamically independent of the generalized coordinate q. We will explain more precisely what is meant by "dynamical independence" later. First we have to do some preliminary mathematical spadework to understand how to eliminate $\dot{q}$ and replace it with the canonical momentum p for Hamilton's theory.

5.4 THE LEGENDRE TRANSFORMATION

We begin with a purely mathematical exercise. The Legendre transformation is a recipe for starting with a function of a variable and generating a new function of a new

* The Laplace–Runge–Lenz vector is defined in Problem 2 at the end of this chapter.

variable. If the transformation is repeated, it restores the old function of the old variable. Legendre transformations are used in mathematical treatments of partial differential equations and also are used very extensively in thermodynamics to change from one set of variables to another. To focus on the mathematical content, we will use a notation that does not specifically refer to mechanics. Consider the independent mathematical variables: a passive* variable x and an active variable y. Assume a function $A(x, y)$ of these variables is known explicitly. Now introduce a third variable z and define the function of these three initially independent variables $B(x, y, z) \equiv yz - A(x, y)$. (The minus sign is not essential but will be convenient.) Small changes dx, dy, dz in x, y, z cause a change dB in the function B:

$$dB = z\,dy + y\,dz - \left.\frac{\partial A}{\partial x}\right|_y dx - \left.\frac{\partial A}{\partial y}\right|_x dy. \tag{5.16}$$

Regrouping the terms in Equation (5.16) we get

$$dB = \left(z - \left.\frac{\partial A}{\partial y}\right|_x\right) dy + y\,dz - \left.\frac{\partial A}{\partial x}\right|_y dx. \tag{5.17}$$

So far, z has been an arbitrary independent variable. We now define z to be a function of x and y by the equation

$$z = z(x, y) \equiv \left.\frac{\partial A}{\partial y}\right|_x. \tag{5.18}$$

The coefficient of the term proportional to dy in Equation (5.17) vanishes. The other partial derivatives of B, which is now only a function of x, z can be computed from Equation (5.17):

$$\left.\frac{\partial B}{\partial z}\right|_x = y, \quad \left.\frac{\partial B}{\partial x}\right|_z = -\left.\frac{\partial A}{\partial x}\right|_y. \tag{5.19}$$

To compute B explicitly, we have to invert the relation for z (5.18), solving for $y = y(x, z)$ and then substitute into $B(x, y(x, z), z)$. With the Legendre transformation, $y(x, z)$ is also obtained from the partial derivative $y = y(x, z) = \frac{\partial B}{\partial z}|_x$. This means that, given $B(x, z)$, the transformation can be inverted. For a Legendre transformation, it is possible to work either with $B(x, z)$ or with $A(x, y)$ to find the "passive" partial derivative, since $\frac{\partial B}{\partial x}|_z = -\frac{\partial A}{\partial x}|_y$.

It is often said that the advantage of the Legendre transformation is that it creates a function of x, z alone. This is true. But this could also be done by substituting an arbitrary functional relationship $y = y(x, z)$ into $B(x, y(x, z), z) = B(x, z)$. However, all information about $y(x, z)$ may be lost after the substitution, since the simple relations between

* The meaning of "passive" and "active" will become clear from the context.

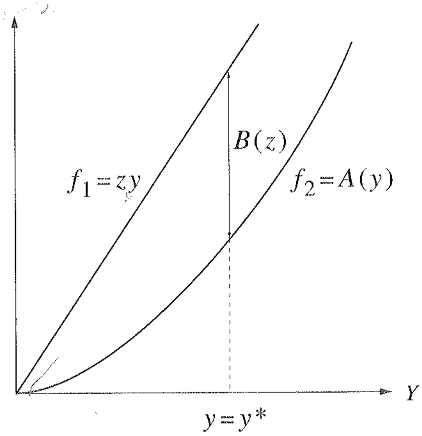

FIGURE 5.2
Graphical representation of the Legendre transformation showing the construction of z and $B(z)$ from y and $A(y)$. At the maximum separation, $\frac{d}{dy}(zy - A) = 0$, so $z = \frac{dA}{dy}|_{y=y^*}$.

partial derivatives ((5.18), (5.19)) would not be valid. A concrete example may help to make this clear.

Example
This is really an exercise. Define $A(x, y) \equiv (1 + x^2)y^2$. Prove that $B(x, z) = \frac{z^2}{4(1+x^2)}$ for the Legendre transformation. Show that you can invert the transformation using the partial derivative relations (5.19) and an "inverse" Legendre transformation to find $y(x, z)$ and $A(x, y)$ from $B(x, z)$. Now try the arbitrary substitution $y = z$ and show that the form of $y(x, z)$ cannot be recovered from knowing $B(x, z) \equiv yz - A(x, y) = -x^2z^2$.

Two different, but completely equivalent, geometric interpretations of the Legendre transformation may help the reader to visualize what the transformation means. In the first way, the distance between a line of variable slope z: $f_1(y) = zy$ and a function $f_2(y) = A(y)$ is maximized to find $y^*(z)$. (We are suppressing the passive variable x.) This shows that only convex functions $A(y)$ can be used for the Legendre transformation, since otherwise the maximum might not exist. To find the maximum distance you must solve the equation $\frac{d}{dy}(zy - A(y)) = 0$, which is the same equation as (5.18). Figure 5.2 shows this construction. The maximum distance is the function $B(z)$.

A second construction, Figure 5.3, shows the dual nature of the Legendre transformation. If z is the slope of the tangent to the curve $A(y)$ then $B(z)$ is the intercept of the line tangent to A at the point y^*. The same is true if instead we start from z and the convex function $B(z)$ and in the same way build y and $A(y)$.

QUESTION 5: *Convex versus Concave* Why does $A(y)$ have to be a convex function? What happens if it is not? (Try $A(y) = y$, for example.) Is $B(z)$ convex?

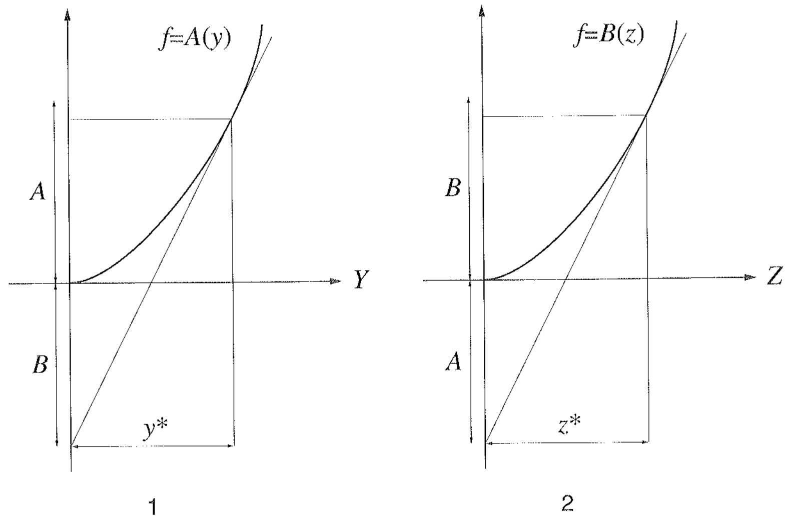

FIGURE 5.3
Dual nature of the Legendre transformation. 1) Construction of z and $B(z)$ from y and $A(y)$. Since the slope of $A(y)$ at $y = y^*$ is equal to $z(y^*)$, then $z = \frac{A(y^*)+B(z)}{y^*}$. 2) Construction of y and $A(y)$ from z and $B(z)$. Since the slope of $B(z)$ at $z = z^*$ is equal to $y(z^*)$, then $y = \frac{A(y)+B(z^*)}{z^*}$.

Why Transform?

In mechanics, start with the Lagrangian $L(q, \dot{q})$ (the possibility of explicit time dependence in the Lagrangian will be temporarily set aside, just to simplify the notation). The active variable is $\dot{q}$, and the passive variable q. By making the Legendre transformation as described above, we pass to the variable p and the Hamiltonian $H(q, p)$:

$$H \equiv p\dot{q} - L(q, \dot{q}), \qquad p \equiv \left.\frac{\partial L}{\partial \dot{q}}\right|_{\text{constant } q}. \tag{5.20}$$

The transformation is invertible as noted above.

Since the Legendre transformation can be made equally well in either direction, why do we prefer the variable p and the Hamiltonian $H(q, p)$ to the choice of $\dot{q}$ and $L(q, \dot{q})$? The key feature of using the canonical momentum p, which is the tangent to the Lagrangian, instead of $\dot{q}$, is that Hamilton's Principle holds for *independent* variations of q and p. The arbitrary variations δp and δq are truly independent at each point in time, unlike the variations δq and $\delta \dot{q}$. To see this, recall Hamilton's Principle

$$\delta S(\text{action}) = \int \delta L \, dt = 0. \tag{5.21}$$

Calculate δL in terms of the variations of q and p from (5.20):

$$\delta L = \dot{q}\delta p + p\delta\dot{q} - \delta H. \tag{5.22}$$

The chain rule for partial derivatives tells us (remember that δp, δq are arbitrary infinitesimal functions of the time) that

$$\delta H = \frac{\partial H}{\partial q}\delta q + \frac{\partial H}{\partial p}\delta p. \tag{5.23}$$

Inserting Equation (5.23) into Equation (5.22) and collecting together the coefficients of δq and δp, we have

$$\delta L = \left(\dot{q} - \frac{\partial H}{\partial p}\right)\delta p - \left(\dot{p} + \frac{\partial H}{\partial q}\right)\delta q + \frac{d}{dt}(p\delta q). \tag{5.24}$$

When Equation (5.24) is integrated with respect to time to compute the variation in the action, the total time derivative on the right-hand side of the equation will contribute nothing if $\delta q = 0$ at the end points. This has always been a requirement of Hamilton's Principle. Thus (5.24) indicates that Hamilton's Principle will only work for independent arbitrary variations of p and q if the coefficients of δp, δq vanish. (The dual transformation already shows us that the coefficient of δp vanishes automatically.)

To prove that the coefficients *do* vanish, start again from the basic defining Equation (5.20) and vary all the variables, $p, q, \dot{q}$ (this does not imply that they are independent), to obtain

$$dH = \dot{q}dp + pd\dot{q} - \left.\frac{\partial L}{\partial q}\right|_{\dot{q}} dq - \left.\frac{\partial L}{\partial \dot{q}}\right|_{q} d\dot{q}. \tag{5.25}$$

The coefficient of $d\dot{q}$ vanishes due to the definition of p. Varying q for constant p and p for constant q yields the equations

$$\dot{q} = \left.\frac{\partial H}{\partial p}\right|_{q}, \quad \left.\frac{\partial H}{\partial q}\right|_{p} = -\left.\frac{\partial L}{\partial q}\right|_{\dot{q}} = -\left.\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right|_{q} = -\dot{p}. \tag{5.26}$$

In the last step, on the right, we have made use of the passive nature of q in the Legendre transformation and the Euler–Lagrange equations of motion.

We have not only derived Hamilton's canonical equations of motion but have proved at the same time that independent infinitesimal variations in δq and δp *from the physical path in phase space* do not change the action, Equation (5.21). We can summarize our result in a single equation representing any change in H due to changes in the arguments of the function $H(q, p)$:

$$dH = \dot{q}dp - \dot{p}dq. \tag{5.27}$$

The symmetry between q and p is evident here. It is a consequence of the Legendre transformation combined with the Euler–Lagrange equation.

Before, in Chapter 2, we plotted $q(t)$ on the Y axis and time t on the X axis, making small variations from the actual graph of the physical coordinate versus time, and proving

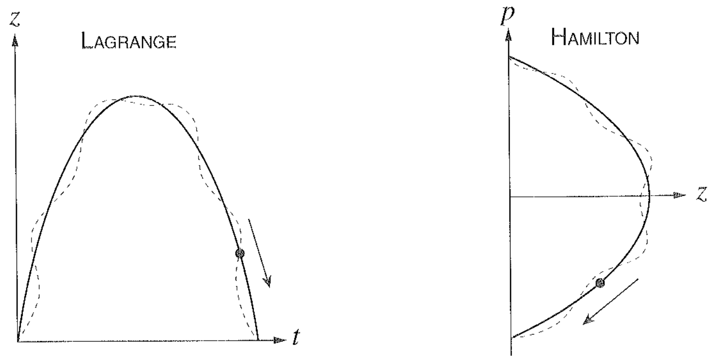

FIGURE 5.4
Two views of the dynamics of a falling body. In the Lagrange case, the height z versus the time t is plotted. In the Hamilton case, momentum p versus height z is plotted. Dashed curves are variations from the physical paths.

that the equations of motion follow from making the action integral an extremum on the physical path. Now we are plotting the phase trajectory of the moving point $q(t)$, $p(t)$ in phase space as shown in Figure 5.4. Varying this trajectory by arbitrary infinitesimal variations in q and p also leaves the action unchanged. Since we take Hamilton's Principle to be the basic law of mechanics, the trajectory in phase space with this extremum property is the solution for the motion. In this view, the time appears as a parameter that we vary in order to trace out the trajectory. There are rewards for this shift in viewpoint, which we will discuss below.

5.5 HAMILTON'S EQUATIONS OF MOTION

For N degrees of freedom, the $2N$-dimensional phase space becomes $\{q_k, p_k\}$ and the Hamiltonian H is

$$H \equiv \sum_{k=1}^{N} p_k \dot{q}_k - L,$$
$$H = H(q_1, q_2, \ldots, q_N, p_1, p_2, \ldots, p_N, t). \tag{5.28}$$

To consider the possibility that the time might appear explicitly in the Hamiltonian, add a term for this:

$$dH = \sum_{k=1}^{N} \dot{q}_k \, dp_k - \sum_{k=1}^{N} \dot{p}_k \, dq_k + \frac{\partial H}{\partial t} dt. \tag{5.29}$$

Since the time is also a passive variable in the Legendre transformation we know that

$$\left.\frac{\partial H}{\partial t}\right|_{q_1,\ldots,p_1,\ldots} = -\left.\frac{\partial L}{\partial t}\right|_{q_1,\ldots,\dot{q}_1,\ldots}. \tag{5.30}$$

The total time derivative of H can be computed:

$$\frac{dH}{dt} = \underbrace{\dot{q}\dot{p} - \dot{p}\dot{q}}_{=0} + \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{5.31}$$

If there is no explicit time dependence in L, H will be a constant of the motion. If the kinetic energy is a quadratic form in the $\dot{q}_k$s, H is also the total energy $E = T + V$.

The final result is

$$\boxed{\dot{q}_k = \frac{\partial H}{\partial p_k} \quad \dot{p}_k = -\frac{\partial H}{\partial q_k} \quad \frac{dH}{dt} = -\frac{\partial L}{\partial t}.} \tag{5.32}$$

Hamilton's equations of motion

These are the fundamental equations of Hamiltonian dynamics.

In Lagrangian dynamics N second-order differential equations must be solved. In Hamiltonian dynamics there are $2N$ first-order equations instead. This often makes very little difference in the difficulty of finding explicit solutions. The fact that q and p are treated (almost) symmetrically allows for the discovery of some important theorems: Liouville's Theorem, which we will discuss later in this chapter, and the Poincaré Recurrence Theorem, which is discussed in Appendix B. It also makes possible the development of sophisticated analytical tools such as *canonical transformations*, as we shall see in Chapter 6.

We can get (q, p) at time $t + dt$ from the knowledge of (q, p) at time t by using Hamilton's equations. Thus a step by step time integration can be performed. This is what is actually done when the equations of motion are numerically integrated on the computer.

We now summarize what you must do in order to start from a Lagrangian and convert to the use of Hamilton's dynamics:

1. Define the momentum canonically conjugate to q by the "tangent" to the Lagrangian:

$$p_k \equiv \left.\frac{\partial L}{\partial \dot{q}_k}\right|_{\text{constant } q_k}. \tag{5.33}$$

   Do this for each degree of freedom, holding the coordinates and velocities for the other degrees of freedom constant.
2. Define the Hamiltonian H as in Equation (5.28) above. This is now a mixed function of all the q_k, $\dot{q}_k$ and p_ks. This is still not the final form, since the Hamiltonian must be expressed as a function only of the q_ks and p_ks.
3. Invert the function(s) you obtained in Equation (5.33) to get $\dot{q}_k(q_1, q_2, \ldots, p_1, p_2, \ldots)$.
4. Eliminate the generalized velocities in the temporary form of the Hamiltonian from Equation (5.28). You should now have the Hamiltonian as a function of the p_ks and q_ks only. The time will appear explicitly in the Hamiltonian only if it was explicitly present (due to time-dependent constraints) in the Lagrangian.
5. Solve the $2N$ first-order Equations (5.32).