CHAPTER SIX

THEORETICAL MECHANICS:

FROM CANONICAL TRANSFORMATIONS TO ACTION–ANGLE VARIABLES

OVERVIEW OF CHAPTER 6

Canonical transformations are transformations from one set of canonically conjugate variables q, p to another conjugate set Q, P. A transformation is said to be canonical if, after the transformation, Hamilton's equations are still the correct dynamical equations for the time development of the new variables. The new Hamiltonian may look quite different from the old one. It may prove easier to solve the EOM in terms of the new variables Q, P. The concept of a *generating function* is introduced, which gives an "automatic" method for producing canonical transformations. There are four types of generating functions for canonical transformations. It will be explained how these different generating functions are connected by Legendre transformations.

Poisson brackets will be introduced, which are invariant under canonical transformations. If Hamilton's dynamics is formulated in terms of Poisson brackets, we have a coordinate-free way to express the equations of motion. The close resemblance of Poisson brackets used in classical mechanics to commutators of operators in quantum mechanics is not an accident, since Poisson brackets played a fundamental role in the invention of quantum mechanics.

We proceed from the general notion of a generating function to the special generating function S, which produces a canonical transformation leading to the *Hamilton–Jacobi equation*. The Hamilton–Jacobi equation leads to a geometric picture of dynamics relating the dynamics to wave motion. There is a close connection between the Hamilton–Jacobi equation in mechanics and the Schrödinger equation in quantum mechanics.

The generating function S turns out to be the time integral of the Lagrangian – the action. The Hamilton–Jacobi equation for time-independent Hamiltonians describing periodic motion leads to the concept of a special set of canonically conjugate variables – *action–angle* variables. The action variables are constants of the motion, and the angle variables increase linearly with time. Thus the time development of the dynamical system takes a simple form when cast in terms of action–angle variables. These variables are also important for further theoretical analysis of dynamical systems.

Conservative systems (systems without damping) come in two types: *integrable* and *nonintegrable*. Only the latter type can exhibit chaos. Integrable systems with N

degrees of freedom by definition have N constants of the motion. Each of these constants confines the motion to a $(2N-1)$-dimensional subspace of the $2N$-dimensional phase space. The intersection of N of these subspaces requires that all the motion takes place on an N-dimensional surface embedded in the $2N$-dimensional phase space. Trajectories on this "surface" can either be periodic or quasiperiodic.

Separable systems are a subset of integrable systems. All analytically soluble mechanics problems are of this type. The Hamilton–Jacobi equation is the most powerful technique for solving separable systems.

6.1 CANONICAL TRANSFORMATIONS

Up to now we have discussed only coordinate transformations (technically known as *point transformations*) between two different sets of space coordinates Q and q:

$$\boxed{\begin{array}{c} Q = Q(q(t), t). \\ \text{point transformation} \end{array}} \tag{6.1}$$

or, with more degrees of freedom,

$$Q_k = Q_k(q_1, q_2, \ldots q_N, t) \quad (k = 1, \ldots, N). \tag{6.2}$$

You have probably always referred to point transformations as a change of variables. Notice that this kind of transformation is in the mechanical system's configuration space. It is not the most general mathematical transformation possible in phase space. There could be more general transformations in which the coordinates and the momenta are interdependent. Transformations of this general type are called *contact transformations*:

$$\boxed{\begin{array}{c} Q = Q(q, p, t), \quad P = P(q, p, t) \\ \text{contact transformation} \end{array}} \tag{6.3}$$

The terminology originated with projective geometry.

Transformations of the type (6.3) are *equivalent* descriptions of the dynamics of a given system if there exists a new Hamiltonian, a function of Q, P and perhaps t, that gives equations of motion in terms of the new variables, which are again Hamilton's equations (5.32). *Canonical transformations will, by definition, take us from one set of coordinates q and canonically conjugate momenta p to another set Q, P, in such a way that the structure of Hamilton's equations for* all *dynamical systems is preserved by this transformation.* The canonically conjugate relation between Q and P will also be preserved. This means that Hamilton's equations will continue to describe the motion for a given specific dynamical system, but with a new Hamiltonian which is a function of the new variables Q, P. The new Hamiltonian for any particular system can be derived from the old Hamiltonian by applying a simple rule. A given canonical transformation does not depend on any specific problem or Hamiltonian; it necessarily preserves the form of Hamilton's equations for all dynamical systems to which it is applied. The main application is towards a better understanding of the theoretical structure of mechanics.

Canonical transformations may or may not be of practical use in solving problems. In some cases the equations of motion can be drastically simplified such that the main features of the motion are more clearly revealed. If the contact transformation were *not* canonical, we would sacrifice all of the theoretical advantages that flow out of Hamilton's analytical mechanics, such as Liouville's theorem. It is usually not easy to guess the form of the canonical transformation that will simplify the EOM for a specific dynamical system, but there is a definite mathematical technique that will guarantee to produce a canonical transformation, useful or not.

Recall that two different descriptions of the same physical system are equivalent if their Lagrangians differ by a total time derivative of the form $\frac{dF(q,t)}{dt}$. You may want to review a proof of this before proceeding further. (See Problem 1.6 and/or Question 2.5. Why can't F depend on $\dot{q}$?) Imagine that we have two ways of describing a physical system. Call $\bar{L}(Q, \dot{Q}, t)$ the Lagrangian of the system using the "Q" description, and $L(q, \dot{q}, t)$ the Lagrangian using the "q" description. The two descriptions refer to the same physical system if

$$\boxed{\bar{L}(Q, \dot{Q}, t) = \lambda L(q, \dot{q}, t) - \frac{dF(q, Q, t)}{dt}.} \tag{6.4}$$

Time derivatives of q and/or Q are not allowed to appear in F. (We choose the minus sign in front of F for convenience.) λ is a constant factor. However, by definition, only $\lambda = 1$ can be called a canonical transformation. $\lambda \neq 1$ is associated with a change of units, which is not considered to be a canonical transformation in the most common sense of the term.

We will use Hamilton's Principle to prove that the Euler–Lagrange equations still hold in terms of the new variables for $\bar{L}$ if they hold in terms of the old variables for L. Integrating (6.4), we obtain

$$\int_{t_1}^{t_2} \bar{L}\, dt = \int_{t_1}^{t_2} L\, dt + F(q(t_1), Q(t_1), t_1) - F(q(t_2), Q(t_2), t_2). \tag{6.5}$$

Since Hamilton's Principle holds in the old (q) system, it must also hold in the new (Q) variables. This follows immediately by taking the variation of Equation (6.5) and assuming that arbitrary variations $\delta q(t)$ imply arbitrary variations $\delta Q(t)$. It is necessary to replace our previous assumption that $\delta q = 0$ at the end points of the action integral with the new assumption $\delta F = 0$ at the end points. Then the two descriptions are equivalent, that is, *the physics is the same, independent of which coordinate system we use to describe the system.*

The function F can be used with any specific Lagrangian to "generate" a new, but equivalent, description of the particular physical system described by this Lagrangian. We think of the canonical transformation as associated with a given form for F rather than with a particular L. There are some restrictions on what you can use for $F(q, Q, t)$. A necessary and sufficient condition for an acceptable F is that $\frac{\partial^2 F}{\partial q \partial Q} \neq 0$. If the mixed second derivative vanishes, it can be shown that the transformation will not be invertible.

F is called a *generating function*. There are four possible types of generating functions, as will be discussed below. The chain rule for the time derivative of $F(q, Q, t)$ is

$$\frac{dF}{dt} = \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial Q}\dot{Q} + \frac{\partial F}{\partial t}. \tag{6.6}$$

Since from (6.4) $\dot{q}$ does not appear explicitly in $\bar{L}$,

$$\frac{\partial \bar{L}}{\partial \dot{q}} = \frac{\partial L}{\partial \dot{q}} - \frac{\partial F}{\partial q} = 0, \quad \text{i.e., } p = \frac{\partial F}{\partial q}, \tag{6.7}$$

$$P \equiv \frac{\partial \bar{L}}{\partial \dot{Q}} = -\frac{\partial F}{\partial Q}. \tag{6.8}$$

To summarize:

$$\boxed{P = -\frac{\partial F}{\partial Q}, \quad p = \frac{\partial F}{\partial q}.} \tag{6.9}$$

Equations (6.9) give two equations for the two unknowns, $P(p, q)$, $Q(p, q)$. To find an explicit form for the transformation, solve Equation (6.7) to express $Q = Q(q, p, t)$, and then insert this relation into Equation (6.8) (after taking the partial derivative) to get $P = P(q, p, t)$. In some cases this may be difficult or even impossible to carry out analytically.

➲ Example

As a rather simple example, suppose we take $F = qQ$. Then, according to Equations (6.9), $P = -\frac{\partial F}{\partial Q} = -q$ and $p = \frac{\partial F}{\partial q} = Q$. This particular generating function interchanges the role of coordinate and momentum. To anticipate the result derived below, the new Hamiltonian $\bar{H}(Q, P) = H(-P, Q)$, if $H(q, p)$ is the original Hamiltonian. The reader should check that $\frac{\partial \bar{H}}{\partial P} = \dot{Q}$ and $\frac{\partial \bar{H}}{\partial Q} = -\dot{P}$ if these transformations are made. (Notice that the minus sign in (6.9) is necessary to preserve the form of Hamilton's equations.) We emphasize that Hamilton's equations will always be preserved, since a generating function will automatically generate a canonical transformation.

What if there is more than one degree of freedom? Then F becomes a function of the q_ks and Q_ks ($k = 1, \ldots, N$), and possibly the time. Using $p_k = \frac{\partial F}{\partial q_k}$ and $P_k = -\frac{\partial F}{\partial Q_k}$ we now have the $2N$ equations like (6.9), which give us the transformation rules implicitly, since we have to solve $2N$ equations in $2N$ unknowns Q_k, P_k. With more variables, we have to keep track of the indices, but the basic rules for transformation remain the same as for one degree of freedom.

To find the new Hamiltonian $\bar{H}(Q, P)$, we need to return to the definition of how the Hamiltonian is derived from the Lagrangian by a Legendre transformation (5.20):

$$\bar{H}(Q, P, t) \equiv P\dot{Q} - \bar{L} = -\frac{\partial F}{\partial Q}\dot{Q} - L + \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial Q}\dot{Q} + \frac{\partial F}{\partial t}. \tag{6.10}$$

(We've used (6.6) for $\frac{dF}{dt}$ above.) Thus

$$\bar{H}(Q, P, t) = p\dot{q} - L + \frac{\partial F}{\partial t} \tag{6.11}$$

and so

$$\boxed{\bar{H}(Q, P, t) = H(q(Q, P), p(Q, P), t) + \frac{\partial F(q(Q, P), Q, t)}{\partial t}.} \tag{6.12}$$

Equation (6.12) says: "*to find the new Hamiltonian, just insert the inverse of the transformation equations expressing P and Q in terms of p and q into the old Hamiltonian H. If F had an explicit time dependence, then add* $\frac{\partial F}{\partial t}$ *as well.*" This procedure will preserve Hamilton's equations of motion as the new equations of motion, since we know Hamilton's Principle is obeyed for either set of variables. It is only rarely the case that we have an explicitly time-dependent F, so usually $\bar{H}(Q, P) = H(q(Q, P), p(Q, P))$.

In summary, the recipe for a canonical transformation involves these steps:

1. Specify a specific generating function $F(q, Q, t)$.
2. Equations (6.9) give a set of implicit equations for the canonical transformation.
3. Use (6.12) to find $\bar{H}(Q, P, t)$, expressing p and q in terms of P and Q.

There are methods for finding F that we will discuss in Section 6.5, but often you simply make an educated guess.

If we start with a contact transformation in the form of Equations (6.3), how do we find the F which generates it? First, express p, P as functions of q, Q, and t. Then consider Equations (6.9) as partial differential equations to be solved for $F(q, Q, t)$. This may or not be possible to solve, however. Not every possible contact transformation is a canonical transformation. The Hamilton–Jacobi equation is a special case of this "inverse" procedure which we will discuss later.

QUESTION 1: *Canonical Transformation* Follow the recipe for a canonical transformation outlined in the previous section for $F = q + Q$ (use it on your favorite

Hamiltonian). You will find that you do not obtain Hamilton's equations in terms of the new variables. Why does this happen? What is wrong with our generating function?

QUESTION 2: *Change of Scale* A scale change (change of units) clearly does not change the motion of any dynamics. Let's look at what happens to the Hamiltonian and the Lagrangian in this case. 1) Prove that if $H(p, q)$ is the original Hamiltonian, and you make the scale change $Q = \mu q$, $P = \nu p$, then $\bar{H}(Q, P) = \lambda H(\frac{Q}{\mu}, \frac{P}{\nu})$ is the new Hamiltonian in Q, P, where $\lambda = \mu\nu$ where μ, ν are constants. 2) Knowing the form of the Hamiltonian above, prove that the new Lagrangian is of the form of Equation (6.4) with $\lambda = \mu\nu$.

➲ Example: Harmonic Oscillator Solved by a Canonical Transformation

For this problem (see (5.34))

$$H(q, p) \equiv \frac{1}{2}(p^2 + \omega^2 q^2). \tag{6.13}$$

With 20/20 foresight, choose*

$$F(q, Q) = \frac{1}{2}\omega q^2 \cot 2\pi Q. \tag{6.14}$$

Carry out the canonical transformation:

$$\begin{aligned} p &= \frac{\partial F}{\partial q} = \omega q \cot 2\pi Q, \\ P &= -\frac{\partial F}{\partial Q} = \frac{\pi \omega q^2}{\sin^2 2\pi Q}. \end{aligned} \tag{6.15}$$

Solve the implicit transformation equations for the explicit (inverse) transformation equations:

$$\begin{aligned} p &= \sqrt{\frac{\omega P}{\pi}} \cos 2\pi Q, \\ q &= \sqrt{\frac{P}{\pi \omega}} \sin 2\pi Q. \end{aligned} \tag{6.16}$$

Substitute Equations (6.16) into H:

$$\bar{H} = \frac{\omega}{2\pi} P. \tag{6.17}$$

* Not the most obvious function to choose for F, we admit.