

8. Canonical Transformations

We found that certain choices of coordinates give rise to cyclic coordinates leading to their corresponding conjugate momenta being constants. We will see that it is possible to find a set of coordinate transformations that lead to all coordinates being cyclic.

In the Lagrangian formalism, it is difficult to determine these optimal transformations; this is because both the coordinates and their time derivatives appear in the Lagrangian. In the Hamiltonian formalism, we can treat the coordinates and their conjugate momenta as independent, and we thus have more liberty to seek transformations that make the coordinates cyclic. In general, we call transformations that preserve Hamilton's canonical equations, "canonical transformations". The transformed coordinates are denoted by Q_i and P_i , and the transformed Hamiltonian by K . The conditions satisfied by a canonical transformation are thus:

$$\begin{aligned} H(q_i, p_i, t) &\rightarrow K(Q_i, P_i, t) \\ \dot{Q}_i &= \frac{\partial K}{\partial P_i} \\ \dot{P}_i &= -\frac{\partial K}{\partial Q_i} \end{aligned} \tag{8.1}$$

8.1 Types of transformations

Point transformations

The simplest type of transformation, "point transformations" simply change the q_i to a new set of coordinates, Q_i . We know from our liberty in choosing a set of generalized coordinates that the Lagrangian is invariant under point transformation. Naturally, Hamilton's equations are also preserved, and point transformations of the type:

$$q_i = q_i(Q_i, t) \tag{8.2}$$

are canonical.

General transformations

Consider the general invertible canonical transformation of the form:

$$\begin{aligned} q_i &= q_i(Q_i, P_i, t) \\ p_i &= p_i(Q_i, P_i, t) \end{aligned} \tag{8.3}$$

The equations of motion will be preserved if the integrand of the action (i.e. the Lagrangian) is modified at most by a total time differential of some function, F :

$$\sum_i \dot{q}_i p_i - H = \sum_i \dot{Q}_i P_i - K + \frac{d}{dt} F(q_i, Q_i, t) \quad (8.4)$$

where we have, arbitrarily, chosen that F depends on the pair of variables q_i, Q_i . We could have just as well chosen any pair of variables that mixes the new and old coordinates, $F(p_i, P_i)$, $F(p_i, Q_i)$, or $F(q_i, P_i)$. We assume (impose) that the transformation is canonical, so that:

$$\begin{aligned} \dot{Q}_i &= \frac{\partial K}{\partial P_i} \\ \dot{P}_i &= -\frac{\partial K}{\partial Q_i} \end{aligned} \quad (8.5)$$

We will require that Equation 8.4 is true so that we can make the transformation canonical. Consider the time derivative of F :

$$\frac{dF}{dt} = \sum_i \left(\frac{\partial F}{\partial q_i} \dot{q}_i + \frac{\partial F}{\partial Q_i} \dot{Q}_i \right) + \frac{\partial F}{\partial t} \quad (8.6)$$

which we can substitute back into 8.4:

$$\begin{aligned} \sum_i \dot{q}_i p_i - H &= \sum_i \dot{Q}_i P_i - K + \sum_i \left(\frac{\partial F}{\partial q_i} \dot{q}_i + \frac{\partial F}{\partial Q_i} \dot{Q}_i \right) + \frac{\partial F}{\partial t} \\ \sum_i \left(p_i - \frac{\partial F}{\partial q_i} \right) \dot{q}_i - H &= \sum_i \left(P_i + \frac{\partial F}{\partial Q_i} \right) \dot{Q}_i - K + \frac{\partial F}{\partial t} \end{aligned} \quad (8.7)$$

We can guarantee the validity of this equation by setting the coefficients of $\dot{q}_i$ and $\dot{Q}_i$ to zero:

$$\begin{aligned} p_i &= \frac{\partial}{\partial q_i} F(q_i, Q_i, t) \\ P_i &= -\frac{\partial}{\partial Q_i} F(q_i, Q_i, t) \end{aligned}$$

and by requiring that the new Hamiltonian is then given by:

$$K(Q_i, P_i, t) = H(p_i, q_i, t) + \frac{\partial F}{\partial t} \quad (8.9)$$

We call F the “generator” of the canonical transformation, since it tells how to define the new Hamiltonian and the new coordinates. Given F and the above equations, one can always invert the transformation equations to get:

$$\begin{aligned} q_i &= q_i(Q_i, P_i, t) \\ p_i &= p_i(Q_i, P_i, t) \\ Q_i &= Q_i(q_i, p_i, t) \\ P_i &= P_i(q_i, p_i, t) \end{aligned} \tag{8.10}$$

and Hamilton’s canonical equations for $K(Q_i, P_i, t)$.

We have four possible “types” of canonical transformations depending on the variables that F depends on:

1. $F = F_1(q_i, Q_i, t)$ are transformations of the first type
2. $F = F_2(q_i, P_i, t) - \sum_i Q_i P_i$ are transformations of the second type
3. $F = F_3(p_i, Q_i, t) + \sum_i p_i q_i$ are transformations of the third type
4. $F = F_4(p_i, P_i, t) + \sum_i p_i q_i - \sum_i Q_i P_i$ are transformations of the fourth type

These relations between the generating functions are similar to Legendre transformations, although it is not always true that, for a given situation, one can use any of the coordinate transformations (this may lead to expressions that are singular, ill-defined, etc.). It should also be noted that one can use generating functions that mix the above possibilities for different indices. For example, a valid generating function could be of one type for $i = 1$ and of another for $i = 2$.

One can use the same formalism as above to determine how the various generating functions give different coordinate transformations. These are summarized in table [8.1](#).

Type	Transformation equations
$F = F_1(q_i, Q_i, t)$	$p_i = \frac{\partial F_1}{\partial q_i} \quad P_i = -\frac{\partial F_1}{\partial Q_i}$
$F = F_2(q_i, P_i, t) - \sum_i Q_i P_i$	$p_i = \frac{\partial F_2}{\partial q_i} \quad Q_i = \frac{\partial F_2}{\partial P_i}$
$F = F_3(p_i, Q_i, t) + \sum_i p_i q_i$	$q_i = -\frac{\partial F_3}{\partial p_i} \quad P_i = -\frac{\partial F_3}{\partial Q_i}$
$F = F_4(p_i, P_i, t) + \sum_i p_i q_i - \sum_i Q_i P_i$	$q_i = -\frac{\partial F_4}{\partial p_i} \quad Q_i = \frac{\partial F_4}{\partial P_i}$

Table 8.1 Summary of the four types of canonical transformations

It is interesting to note that one can choose an arbitrary function $F(q_i, Q_i, t)$ and generate a canonical transformation, giving us a new set of arbitrary coordinates and Hamiltonian that will automatically satisfy the equations of motion. Being able to re-write the equations of motion in an arbitrary coordinate system is a powerful tool.

EXAMPLE 8-1

Find the transformation equations $Q_i(q_i, p_i, t)$ and $P_i(q_i, p_i, t)$ for the generating function given by $F = \sum_i Q_i q_i$

For a canonical transformation of the first type, we have

$$\begin{aligned}
 p_i &= \frac{\partial F}{\partial q_i} = Q_i \\
 P_i &= -\frac{\partial F}{\partial Q_i} = -q_i \\
 K &= H
 \end{aligned}$$

These can be trivially inverted to obtain the new coordinates in terms of the old ones:

$$\begin{aligned}
 Q_i &= p_i \\
 P_i &= -q_i
 \end{aligned}$$

which had the net effect of reversing the coordinates with their conjugate momenta, highlighting how these should really be treated as independent in Hamiltonian mechanics.

EXAMPLE 8-2

Find the transformation equations $Q_i(q_i, p_i, t)$ and $P_i(q_i, p_i, t)$ for the generating function given by $F = \sum_i q_i P_i$

For a canonical transformation of the second type, we have

$$p_i = \frac{\partial F}{\partial q_i} = P_i$$

$$Q_i = \frac{\partial F}{\partial P_i} = q_i$$

Again, these are trivially inverted

$$Q_i = q_i$$

$$P_i = p_i$$

which is the identity transformation.

EXAMPLE 8-3

Show that the harmonic oscillator problem can be solved easily using transformations from the generating function $F = \frac{m\omega}{2} q^2 \cot Q$, where $\omega = \sqrt{\frac{k}{m}}$ is the angular frequency.

The Hamiltonian for the simple harmonic oscillator is:

$$\begin{aligned} H &= \frac{p^2}{2m} + \frac{k}{2} q^2 \\ &= \frac{1}{2m} (p^2 + m^2 \omega^2 q^2) \end{aligned}$$

The transformation equation for the given generating function of the first kind are:

$$\begin{aligned} p &= \frac{\partial F}{\partial q} = m\omega q \cot Q \\ P &= -\frac{\partial F}{\partial Q} = \frac{m\omega q^2}{2 \sin^2 Q} \\ K &= H \end{aligned}$$

Rearranging for p and q :

$$\begin{aligned} q^2 &= \frac{2}{m\omega} P \sin^2 Q \\ p^2 &= 2m\omega P \cos^2 Q \end{aligned}$$

which we can now substitute to get K :

$$\begin{aligned} K = H &= \frac{1}{2m} (p^2 + m^2 \omega^2 q^2) \\ &= \frac{1}{2m} \left(2m\omega P \cos^2 Q + m^2 \omega^2 \frac{2}{m\omega} P \sin^2 Q \right) \\ &= \omega P (\cos^2 Q + \sin^2 Q) \\ &= \omega P \end{aligned}$$

and since Q is clearly cyclic, we have that P is constant. In this case, the total Hamiltonian is the energy, E , so we can write:

$$P = \frac{E}{\omega}$$

The canonical equation for Q gives:

$$\dot{Q} = \frac{\partial K}{\partial P} = \omega$$

$$\therefore Q(t) = \omega t + \beta$$

where β is an integration constant that depends on the initial conditions. We can now transform this back to the original variables:

$$q(t) = \sqrt{\frac{2P}{m\omega}} \sin Q$$

$$= \sqrt{\frac{2E}{m\omega^2}} \sin(\omega t + \beta)$$

as expected.

8.2 Poisson Brackets and canonical transformations

Consider time-invariant canonical transformations to new coordinates $Q_i(q_i, p_i)$, $P_i(q_i, p_i)$, which preserve the Hamiltonian, H . By definition, Hamilton's canonical equations are satisfied:

$$\dot{Q}_i = \frac{\partial H}{\partial P_i} = \sum_j \left(\frac{\partial H}{\partial q_j} \frac{\partial q_j}{\partial P_i} + \frac{\partial H}{\partial p_j} \frac{\partial p_j}{\partial P_i} \right)$$

$$\dot{P}_i = -\frac{\partial H}{\partial Q_i} = -\sum_j \left(\frac{\partial H}{\partial q_j} \frac{\partial q_j}{\partial Q_i} + \frac{\partial H}{\partial p_j} \frac{\partial p_j}{\partial Q_i} \right)$$

Consider the time derivative of the Q_i and P_i :

$$\begin{aligned} \dot{Q}_i &= \sum_j \left(\frac{\partial Q_i}{\partial q_j} \dot{q}_j + \frac{\partial Q_i}{\partial p_j} \dot{p}_j \right) = \sum_j \left(\frac{\partial Q_i}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial Q_i}{\partial p_j} \frac{\partial H}{\partial q_j} \right) \\ \dot{P}_i &= \sum_j \left(\frac{\partial P_i}{\partial q_j} \frac{\partial H}{\partial p_j} + \frac{\partial P_i}{\partial p_j} \dot{p}_j \right) = \sum_j \left(\frac{\partial P_i}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial P_i}{\partial p_j} \frac{\partial H}{\partial q_j} \right) \end{aligned} \quad (8.12)$$

By comparing terms, we can get the “Direct Conditions” for the transformation to be canonical:

$$\begin{aligned} \frac{\partial p_j}{\partial P_i} &= \frac{\partial Q_i}{\partial q_j} \\ \frac{\partial q_j}{\partial P_i} &= -\frac{\partial Q_i}{\partial p_j} \\ \frac{\partial q_j}{\partial Q_i} &= \frac{\partial P_i}{\partial p_j} \\ \frac{\partial p_j}{\partial Q_i} &= -\frac{\partial P_i}{\partial q_j} \end{aligned}$$

Here, we obtained these relations by assuming that the transformation is canonical (that is, that we could get the time derivative of Q and P from the Hamiltonian). It is also possible to show that the Direct Conditions also hold if the transformations are time-dependent.

Now, consider the Poisson Brackets of the q_i and p_i evaluated in the coordinates Q_i and P_i , where we use the Direct conditions to simplify:

$$\begin{aligned}
 \{q_i, q_j\}_{Q,P} &= \sum_k \left(\frac{\partial q_i}{\partial Q_k} \frac{\partial q_j}{\partial P_k} - \frac{\partial q_i}{\partial P_k} \frac{\partial q_j}{\partial Q_k} \right) = \sum_k \left(-\frac{\partial q_i}{\partial Q_k} \frac{\partial Q_k}{\partial p_j} - \frac{\partial q_i}{\partial P_k} \frac{\partial P_k}{\partial p_j} \right) \\
 &= -\frac{\partial q_i}{\partial p_j} = 0 \\
 \{p_i, p_j\}_{Q,P} &= \sum_k \left(\frac{\partial p_i}{\partial Q_k} \frac{\partial p_j}{\partial P_k} - \frac{\partial p_i}{\partial P_k} \frac{\partial p_j}{\partial Q_k} \right) = \sum_k \left(\frac{\partial p_i}{\partial Q_k} \frac{\partial Q_k}{\partial q_j} + \frac{\partial p_i}{\partial P_k} \frac{\partial P_k}{\partial q_j} \right) \\
 &= \frac{\partial p_i}{\partial q_j} = 0 \\
 \{q_i, p_j\}_{Q,P} &= \sum_k \left(\frac{\partial q_i}{\partial Q_k} \frac{\partial p_j}{\partial P_k} - \frac{\partial q_i}{\partial P_k} \frac{\partial p_j}{\partial Q_k} \right) = \sum_k \left(\frac{\partial q_i}{\partial Q_k} \frac{\partial Q_k}{\partial q_j} + \frac{\partial q_i}{\partial P_k} \frac{\partial P_k}{\partial q_j} \right) \\
 &= \frac{\partial q_i}{\partial q_j} = \delta_{ij}
 \end{aligned} \tag{8.14}$$

Thus it is clear that the Poisson Brackets between the coordinates is independent of the coordinate system. From the exact same algebra, it follows that:

$$\begin{aligned}
 \{Q_i, P_j\} &= \delta_{ij} \\
 \{Q_i, Q_j\} &= \{P_i, P_j\} = 0
 \end{aligned} \tag{8.15}$$

which are equivalent to the Direct Conditions for checking that a transformation is canonical.

EXAMPLE 8-4

Show that the transformation $Q = p$, $P = -q$ is canonical

We can test this by checking the Poisson Brackets between Q and P :

$$\begin{aligned}
 \{Q, Q\} &= \frac{\partial Q}{\partial q} \frac{\partial Q}{\partial p} - \frac{\partial Q}{\partial p} \frac{\partial Q}{\partial q} = 0 \\
 \{P, P\} &= \frac{\partial P}{\partial q} \frac{\partial P}{\partial p} - \frac{\partial P}{\partial p} \frac{\partial P}{\partial q} = 0 \\
 \{Q, P\} &= \frac{\partial Q}{\partial q} \frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p} \frac{\partial P}{\partial q} = 1
 \end{aligned}$$

thus the transformation is canonical. Note that checking the first two is unnecessary, since they are identically zero.

Now, consider two functions in the old coordinates, $U(q_i, p_i)$ and $V(q_i, p_i)$, with Poisson Bracket:

$$\{U, V\}_{q,p} = \sum_i^n \left(\frac{\partial U}{\partial q_i} \frac{\partial V}{\partial p_i} - \frac{\partial U}{\partial p_i} \frac{\partial V}{\partial q_i} \right) \quad (8.16)$$

Consider the same Poisson Bracket, expressed in terms of the new coordinates, Q_i, P_i , where the functions are expressed as $U(Q_i, P_i)$ and $V(Q_i, P_i)$:

$$\begin{aligned} \{U, V\}_{Q,P} &= \sum_i^n \left(\frac{\partial U}{\partial Q_i} \frac{\partial V}{\partial P_i} - \frac{\partial U}{\partial P_i} \frac{\partial V}{\partial Q_i} \right) \\ &= \sum_i^n \left[\left(\sum_j \frac{\partial U}{\partial q_j} \frac{\partial q_j}{\partial Q_i} + \frac{\partial U}{\partial p_j} \frac{\partial p_j}{\partial Q_i} \right) \left(\sum_j \frac{\partial V}{\partial q_j} \frac{\partial q_j}{\partial P_i} + \frac{\partial V}{\partial p_j} \frac{\partial p_j}{\partial P_i} \right) \right. \\ &\quad \left. - \left(\sum_j \frac{\partial U}{\partial q_j} \frac{\partial q_j}{\partial P_i} + \frac{\partial U}{\partial p_j} \frac{\partial p_j}{\partial P_i} \right) \left(\sum_j \frac{\partial V}{\partial q_j} \frac{\partial q_j}{\partial Q_i} + \frac{\partial V}{\partial p_j} \frac{\partial p_j}{\partial Q_i} \right) \right] \\ &= \sum_j^n \left[\frac{\partial U}{\partial q_j} \frac{\partial V}{\partial q_j} \{q_j, q_j\}_{Q,P} + \frac{\partial U}{\partial q_j} \frac{\partial V}{\partial p_j} \{q_j, p_j\}_{Q,P} \right. \\ &\quad \left. + \frac{\partial U}{\partial p_j} \frac{\partial V}{\partial p_j} \{p_j, p_j\}_{Q,P} + \frac{\partial U}{\partial p_j} \frac{\partial V}{\partial q_j} \{p_j, q_j\}_{Q,P} \right] \\ &= \sum_j^n \left[\frac{\partial U}{\partial q_j} \frac{\partial V}{\partial p_j} - \frac{\partial U}{\partial p_j} \frac{\partial V}{\partial q_j} \right] \\ &= \{U, V\}_{q,p} \end{aligned} \quad (8.17)$$

We thus find that the Poisson Bracket of two quantities is preserved in a canonical transformation. Poisson Brackets are canonical invariants.

8.3 Infinitesimal canonical transformations

In an infinitesimal canonical transformation, we have:

$$Q_i = q_i + \delta q_i \quad (8.18)$$

$$P_i = p_i + \delta p_i \quad (8.19)$$

where the δq_i and δp_i are small changes in the variables that are consistent with a canonical transformation (rather than an arbitrary virtual displacement). This transformation can be related to the identity transformation by:

$$F = \sum_i q_i P_i + \epsilon G(q_i, P_i) \quad (8.20)$$

where G is some arbitrary function and ϵ is small. The transformation for a canonical transformation of the second type are:

$$\begin{aligned} p_i &= \frac{\partial F}{\partial q_i} = P_i + \epsilon \frac{\partial G}{\partial q_i} \\ Q_i &= \frac{\partial F}{\partial P_i} = q_i + \epsilon \frac{\partial G}{\partial P_i} \end{aligned} \quad (8.21)$$

We can thus identify:

$$\begin{aligned} \delta q_i &= \epsilon \frac{\partial G}{\partial P_i} \\ \delta p_i &= -\epsilon \frac{\partial G}{\partial q_i} \end{aligned} \quad (8.22)$$

The first term can be written as:

$$\delta q_i = \epsilon \frac{\partial G}{\partial P_i} = \epsilon \frac{\partial G}{\partial (p_i + \delta p_i)} = \epsilon \frac{\partial G}{\partial p_i} \quad (8.23)$$

to first order in ϵ . That is, for a small value of ϵ , the derivative with respect to P_i is almost the same as with respect to p_i . From the properties of Poisson Brackets, this can also be written as:

$$\begin{aligned} \delta q_i &= \epsilon \frac{\partial G}{\partial p_i} = \epsilon \{q_i, G\} \\ \delta p_i &= -\epsilon \frac{\partial G}{\partial q_i} = \epsilon \{p_i, G\} \end{aligned} \quad (8.24)$$

if G is taken to be the Hamiltonian, then we have:

$$\begin{aligned} \delta q_i &= \epsilon \{q_i, H\} = \epsilon \frac{dq_i}{dt} \\ \delta p_i &= \epsilon \{p_i, H\} = \epsilon \frac{dp_i}{dt} \end{aligned} \quad (8.25)$$

and the displacement, $\epsilon \frac{dq(p)}{dt}$ are in the direction of time. That is, the Hamiltonian is the generator of the transformation that takes the variable q_i and p_i to variables Q_i and P_i an infinitesimal time later.

8.4 Hamilton-Jacobi Equation

Consider the case of a canonical transformation of the second type

$$\begin{aligned} F &= F_2(q_i, P_i, t) - \sum_i Q_i P_i \\ \frac{dF}{dt} &= \sum_i \left(\frac{\partial F_2}{\partial q_i} \dot{q}_i + \frac{\partial F_2}{\partial P_i} \dot{P}_i - Q_i \dot{P}_i - P_i \dot{Q}_i \right) + \frac{\partial F_2}{\partial t} \end{aligned} \quad (8.26)$$

Substituting into Equation 8.4:

$$\begin{aligned} \sum_i \dot{q}_i p_i - H &= \sum_i \dot{Q}_i P_i - K + \sum_i \left(\frac{\partial F_2}{\partial q_i} \dot{q}_i + \frac{\partial F_2}{\partial P_i} \dot{P}_i - Q_i \dot{P}_i - P_i \dot{Q}_i \right) + \frac{\partial F_2}{\partial t} \\ \sum_i \left(p_i - \frac{\partial F_2}{\partial q_i} \right) \dot{q}_i - H &= \sum_i \left(\frac{\partial F_2}{\partial P_i} - Q_i \right) \dot{P}_i + \frac{\partial F_2}{\partial t} - K \end{aligned} \quad (8.27)$$

And we obtain the transformation equations:

$$\begin{aligned} p_i &= \frac{\partial F_2}{\partial q_i} \\ Q_i &= \frac{\partial F_2}{\partial P_i} \\ K &= H + \frac{\partial F_2}{\partial t} \end{aligned} \quad (8.28)$$

As a notational convention, we will use $S(q_i, P_i, t)$ instead of F_2 (as will see that it is equal to the action). We can also take $S(q_i, P_i, t)$ as the generator of a canonical transformation with the above equations, since this will guarantee that the canonical equations are satisfied. In fact, we can choose any form that we wish for S , since the above equations guarantee that the canonical equations are satisfied. Let's then choose the particular case that gives $K = 0$. This is a nice choice, since K will then be independent of all the variables, and all Q_i and P_i are cyclic:

$$\begin{aligned} \dot{Q}_i &= 0 \\ \dot{P}_i &= 0 \end{aligned} \quad (8.29)$$

We have thus changed the problem of dynamics to one of finding a transformation that gives all cyclic coordinates. Note that this is not necessarily more straightforward mathematically (in fact, it is usually much harder).

Imposing that $K = 0$ gives us:

$$\begin{aligned} H(q_i, p_i, t) + \frac{\partial S}{\partial t} &= 0 \\ H\left(q_i, \frac{\partial S}{\partial q_i}, t\right) + \frac{\partial S}{\partial t} &= 0 \end{aligned} \quad (8.30)$$

which is a partial differential equation for S and is called the “Hamilton-Jacobi equation”. The function S is called “Hamilton’s principal function”. In the transformed coordinates, the solutions are trivial constants:

$$\begin{aligned} P_i &= \alpha_i \\ Q_i &= \beta_i = \frac{\partial S}{\partial \alpha_i} \end{aligned} \quad (8.31)$$

Is is clear that by combining this with the transformation equations, the problem is solved if all of the α_i and β_i are known. Typically, one does not need to refer to P_i and Q_i in the Hamilton-Jacobi formalism, since these are constants of motion. To highlight this, they are usually called α_i and β_i :

$$\begin{aligned} S &= S(q_i, \alpha_i, t) \\ H(q_i, \frac{\partial}{\partial q_i} S(q_i, \alpha_i, t), t) + \frac{\partial}{\partial t} S(q_i, \alpha_i, t) &= 0 \\ \alpha_i &= \text{const.} \\ \beta_i &= \frac{\partial S}{\partial \alpha_i} \end{aligned} \quad (8.32)$$

Connection to the Lagrangian and action

Noting that S depends on the q_i and the constants α_i , we can ask how S changes with time along the path of the system (the path where the P_i are constant and equal to α_i):

$$\begin{aligned} S &= S(q_i, \alpha_i, t) \\ \therefore \frac{dS}{dt} &= \sum_i \frac{\partial S}{\partial q_i} \dot{q}_i + \frac{\partial S}{\partial t} \end{aligned} \quad (8.33)$$

Using the transformation equations:

$$\begin{aligned} p_i &= \frac{\partial S}{\partial q_i} \\ K &= H + \frac{\partial S}{\partial t} = 0 \\ \frac{dS}{dt} &= \sum_i p_i \dot{q}_i - H = L \\ S &= \int L dt + \text{const} \end{aligned} \quad (8.34)$$

It is thus clear that S is equal to the action within an additive constant.

If the Hamiltonian does not depend on time, then H is a constant of the motion, call it E . If this is the case, one can assume a form for S where the time variable is separated out:

$$\begin{aligned}
H(q_i, \frac{\partial S}{\partial q_i}, t) + \frac{\partial S}{\partial t} &= 0 \\
S(q_i, \alpha_i, t) &= W(q_i, \alpha_i) - Et \\
\therefore H(q_i, \frac{\partial W}{\partial q_i}) &= E
\end{aligned} \tag{8.35}$$

Where W is called “Hamilton’s characteristic function”.

EXAMPLE 8-5

Use the Hamilton-Jacobi formalism to solve the simple harmonic oscillator problem

The Hamiltonian is given by:

$$H = \frac{p^2}{2m} + \frac{k}{2}q^2$$

We can write this in terms of Hamilton's principal function to get the Hamilton-Jacobi equation:

$$\begin{aligned} H(q, \frac{\partial S}{\partial q}, t) + \frac{\partial S}{\partial t} &= 0 \\ \frac{1}{2m} \left(\frac{\partial S}{\partial q} \right)^2 + \frac{k}{2}q^2 + \frac{\partial S}{\partial t} &= 0 \end{aligned}$$

Since the Hamiltonian does not depend on time, we can try a solution using separation of variables:

$$S(q, \alpha, t) = W(q, \alpha) - \alpha t$$

which we substitute back into the Hamilton Jacobi equation:

$$\frac{1}{2m} \left(\frac{\partial W}{\partial q} \right)^2 + \frac{k}{2}q^2 = \alpha$$

Since the left side only depends on q the partial derivatives can be made into total derivatives. It is also clear that the constant α is equal to energy, since it is equal to the Hamiltonian.

$$\begin{aligned} \frac{1}{2m} \left(\frac{dW}{dq} \right)^2 + \frac{k}{2}q^2 &= \alpha \\ W &= \int \sqrt{2m\alpha - m^2\omega^2 q^2} dq \\ S &= -\alpha t + \int \sqrt{2m\alpha - m^2\omega^2 q^2} dq \end{aligned}$$

We also know that we have a second constant of motion, β :

$$\begin{aligned} \beta &= \frac{\partial S}{\partial \alpha} \\ &= -t + \int \frac{2m}{2\sqrt{2m\alpha - m^2\omega^2 q^2}} dq \\ &= -t + \frac{1}{\omega} \sin^{-1} \left(q \sqrt{\frac{m\omega^2}{2\alpha}} \right) \end{aligned}$$

Which we can invert to get q :

$$q(t) = \sqrt{\frac{2\alpha}{m\omega^2}} \sin(\omega(\beta + t))$$

Similarly, the momentum is given by:

$$\begin{aligned} p &= \frac{\partial S}{\partial q} = \frac{\partial}{\partial q} \left(-\alpha t + \int \sqrt{2m\alpha - m^2\omega^2 q^2} dq \right) \\ &= \sqrt{2m\alpha - m^2\omega^2 q^2} \\ &= \sqrt{2m\alpha} \sqrt{1 - \sin^2(\omega(\beta + t))} \\ &= \sqrt{2m\alpha} \cos(\omega(\beta + t)) \end{aligned}$$

8.5 Action-angle variables

Action-angle variables are an extension to the Hamilton-Jacobi method, when the Hamiltonian does not depend on time. This formalism will allow us to determine certain fundamental frequencies of the system without explicitly solving the equations of motion. This method also allows one to transition from classical to quantum mechanics. Recall that in the case where the Hamiltonian does not depend on time, we can use Hamilton's Characteristic function, W in lieu of S :

$$\begin{aligned} S(q_i, \alpha_i, t) &= W(q_i, \alpha_i) - Et \\ p_i &= \frac{\partial W}{\partial q_i} \\ H(q_i, \frac{\partial W}{\partial q_i}) &= E \end{aligned} \tag{8.36}$$

We will consider the case where W is “separable” (at least in some of the coordinates):

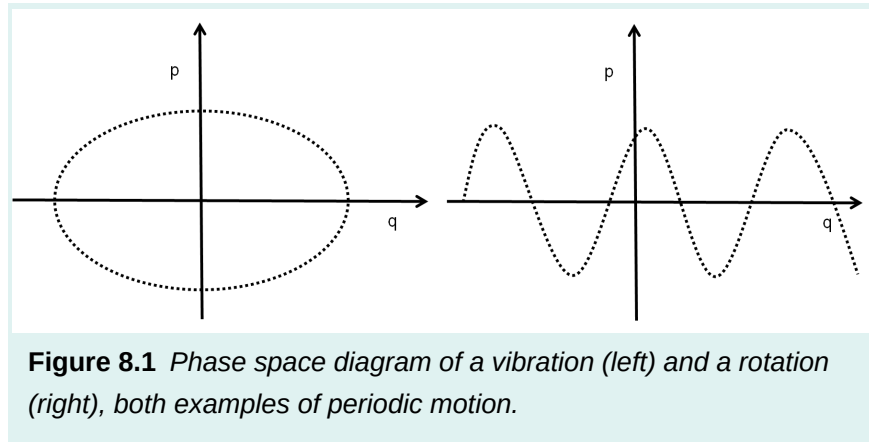
$$W(q_1, \dots, q_n, \alpha_i) = W_1(q_1, \alpha_i) + \dots + W_n(q_n, \alpha_i) \tag{8.37}$$

that is, where the motion in each separable coordinate can be treated as independent from the others coordinates. The conjugate momentum for the separable coordinates are given by:

$$p_i = \frac{\partial W_i(q_i, \alpha_1, \dots, \alpha_n)}{\partial q_i}$$

Furthermore, we will consider the case of “periodic” motion. Two such cases are considered. First, “vibrations” where the coordinates and their momenta, p_i, q_i , will, after a certain period of time, return to their original values. This is the case, for example, for a simple harmonic oscillator. Second, “rotations”

where the momentum, p_i , is periodic as a function of its generalized coordinate, q_i . This is the case, for example, for a pendulum that has enough energy to “go over the top” (the angular position of the pendulum increases indefinitely with time and the momentum is periodic). The two cases are illustrated in phase space in Figure 8.1.



We start by introducing the “action-variables”, J_i , obtained by integration over 1 period of the generalized coordinate:

$$J_i \equiv \oint p_i dq_i \quad (8.39)$$

There is one action variable per separable coordinate. Given the relations for the conjugate momenta for separable variables, we have:

$$\begin{aligned} J_i &= \oint p_i dq_i \\ &= \oint \frac{\partial W_i(q_i, \alpha_1, \dots, \alpha_n)}{\partial q_i} dq_i \end{aligned}$$

and the J_i thus only depend on the α , which are constants of the motion. The J_i are thus constants of the motion as well. We assume that the equations relating the J_i and α_i are invertible:

$$\begin{aligned} J_i &= J_i(\alpha_1, \dots, \alpha_n) \\ \alpha_i &= \alpha_i(J_1, \dots, J_n) \end{aligned} \quad (8.41)$$

We can thus re-write Hamilton’s characteristic function, W , and the Hamiltonian in terms of the J instead of the α :

$$\begin{aligned}
W(q_i, \alpha_i) &\rightarrow W(q_i, J_i) \\
H(q_i, \frac{\partial W}{\partial \alpha_i}) &\rightarrow H(q_i, \frac{\partial W}{\partial J_i}) \\
S(q_i, \frac{\partial W}{\partial \alpha_i}) &\rightarrow W(q_i, J_i) - Et = W(q_i, J_i) - Ht
\end{aligned} \tag{8.42}$$

In effect, we have chosen a new canonical transformation where J_i are the momenta, P_i . In terms of the generating function S , we have:

$$\begin{aligned}
p_i &= \frac{\partial S}{\partial q_i} \\
Q_i &= \frac{\partial S}{\partial J_i}
\end{aligned} \tag{8.43}$$

where S still satisfies the Hamilton-Jacobi equation. The new canonical momenta, J_i , are thus still constants of motion, and we have:

$$\begin{aligned}
P_i &\equiv J_i = \text{const.} \\
Q_i &\equiv \beta_i = \text{const.}
\end{aligned} \tag{8.44}$$

We introduce the “angle variables”, w_i :

$$w_i \equiv \frac{\partial W(q_1, \dots, q_n, J_1, \dots, J_n)}{\partial J_i} \tag{8.45}$$

and consider the constants β_i in terms of the angle variables:

$$\begin{aligned}
\beta_i &= \frac{\partial S}{\partial J_i} \\
&= \frac{\partial}{\partial J_i} (W(q_i, J_i) - Ht) \\
&= w_i - \frac{\partial H}{\partial J_i} t
\end{aligned} \tag{8.46}$$

We write this as:

$$w_i = \beta_i + \nu_i t \tag{8.47}$$

where we have introduced the “frequency”, ν_i (we call it frequency only because of its units for the moment):

$$\nu_i \equiv \frac{\partial H}{\partial J_i} \quad (8.48)$$

Now, assume that the system has a period of T_i for the i th degree of freedom (recall that we imposed that the system was periodic). In the amount of time T_i , the angle variable will have changed by an amount:

$$\Delta w_i \equiv w_i(t = T_i) - w_i(t = 0) = \nu_i T_i \quad (8.49)$$

We can also calculate the amount that w_i changes as the system goes through one period:

$$\begin{aligned} \Delta w_i &= \oint dw_i \\ &= \oint \sum_j \frac{\partial w_i}{\partial q_j} dq_j \\ &= \oint \sum_j \frac{\partial}{\partial q_j} \frac{\partial W}{\partial J_i} dq_j \\ &= \frac{\partial}{\partial J_i} \oint \sum_j \frac{\partial W}{\partial q_j} dq_j \\ &= \frac{\partial}{\partial J_i} \oint \sum_j p_j dq_j \\ &= \frac{\partial}{\partial J_i} \sum_j \oint p_j dq_j \\ &= \frac{\partial}{\partial J_i} \sum_j J_j \\ &= 1 \end{aligned} \quad (8.50)$$

where we have used the definition of the angle variables, the fact that the conjugate momenta can be obtained from W , and the definition of the action variables. Equating the two expressions for Δw_i , it is clear that ν_i is indeed the frequency with which the system is periodic in the i th degree of freedom:

$$\nu_i = \frac{1}{T_i} \quad (8.51)$$

Although the derivation was not particularly intuitive, we have in fact derived a rather elegant result. Namely that for a periodic system, one can re-write the Hamiltonian in terms of the action variable and the derivative of the Hamiltonian with respect to that action variable is the frequency of the motion in the associated coordinate.

EXAMPLE 8-6

Determine the frequencies of a two-dimensional simple harmonic oscillator with different spring constants in the two dimensions

The Hamiltonian for this system is written as:

$$H = \frac{1}{2m}(p_1^2 + p_2^2) + \frac{1}{2}k_1q_1^2 + \frac{1}{2}k_2q_2^2$$

and since it does not depend explicitly on time, is a constant of the motion, which we will call E . The Hamilton-Jacobi equation is given by:

$$\frac{1}{2m} \left(\frac{\partial W}{\partial q_1} \right)^2 + \frac{1}{2m} \left(\frac{\partial W}{\partial q_2} \right)^2 + \frac{1}{2}k_1q_1^2 + \frac{1}{2}k_2q_2^2 = E$$

Here, the variables q_1 and q_2 can be separated, and the Hamilton Jacobi equation written as two independent equations:

$$W = W_1(q_1) + W_2(q_2)$$

$$\frac{1}{2m} \left(\frac{dW_1}{dq_1} \right)^2 + \frac{1}{2m} \left(\frac{dW_2}{dq_2} \right)^2 + \frac{1}{2}k_1q_1^2 + \frac{1}{2}k_2q_2^2 = E$$

Since the total is a constant, the terms depending on only one of the q must be a constant, so we can write:

$$\frac{1}{2m} \left(\frac{dW_1}{dq_1} \right)^2 + \frac{1}{2}k_1q_1^2 = \alpha_1$$

$$\frac{1}{2m} \left(\frac{dW_2}{dq_2} \right)^2 + \frac{1}{2}k_2q_2^2 = \alpha_2$$

where α_1 and α_2 are constants. Solving for the differentials, we can write:

$$p_i = \frac{dW_i}{dq_i} = \sqrt{2m(\alpha_i - \frac{1}{2}k_iq_i^2)}$$

Consider expressing q_i using a new variable, θ_i , given by:

$$q_i = \sqrt{\frac{2\alpha_i}{k_i}} \sin \theta_i$$

$$\therefore dq_i = \sqrt{\frac{2\alpha_i}{k_i}} \cos \theta_i d\theta_i$$

This gives:

$$\begin{aligned} p_i &= \sqrt{2m \left(\alpha_i - \frac{1}{2} k_i \left(\frac{2\alpha_i}{k_i} \sin^2 \theta_i \right) \right)} \\ &= \sqrt{2m\alpha_i \cos^2 \theta_i} \end{aligned}$$

We are now set to evaluate the action variables:

$$\begin{aligned} J_i &= \oint p_i dq_i \\ &= \oint \sqrt{2m\alpha_i \cos^2 \theta_i} \sqrt{\frac{2\alpha_i}{k_i}} \cos \theta_i d\theta_i \\ &= 2\alpha_i \sqrt{\frac{m}{k_i}} \int_0^{2\pi} \cos^2 \theta_i d\theta_i \\ &= 2\pi\alpha_i \sqrt{\frac{m}{k_i}} \end{aligned}$$

This is trivially inverted to get α in terms of J :

$$\alpha_i = \frac{J_i}{2\pi} \sqrt{\frac{k_i}{m}}$$

We can now obtain the Hamiltonian in terms of the action variables:

$$\begin{aligned} H &= \alpha_1 + \alpha_2 \\ &= \frac{1}{2\pi} \left(\sqrt{\frac{k_1}{m}} J_1 + \sqrt{\frac{k_2}{m}} J_2 \right) \end{aligned}$$

The fundamental frequencies for the two degrees of freedom are thus:

$$\begin{aligned} \nu_1 &= \frac{\partial H}{\partial J_1} = \frac{1}{2\pi} \sqrt{\frac{k_1}{m}} \\ \nu_2 &= \frac{\partial H}{\partial J_2} = \frac{1}{2\pi} \sqrt{\frac{k_2}{m}} \end{aligned}$$

as expected.

Connection to Quantum Mechanics

The Hamiltonian description of classical mechanics is important as it is connected with Quantum Mechanics. In this section, we look at a few similarities between the two formalisms and how they are connected.

Recall a few key points about Quantum Mechanics:

- In Quantum Mechanics, the state of a system is described by a wave function: $|\psi\rangle$
- Observables are obtained by operating on the wave function: $\hat{X}|\psi\rangle = x|\psi\rangle$
- Operators do not, in general, commute: $\hat{X}\hat{Y}|\psi\rangle \neq \hat{Y}\hat{X}|\psi\rangle$
- In the Heisenberg representation, the state vector (wave function) is constant in time and the operators change with time
- In the time dependent Schrödinger representation, the operators are constant in time and the wave-function changes with time

Poisson brackets and commutators

You may have noticed a similarity between the use of Poisson Brackets in Classical Mechanics and Commutator relations in Quantum Mechanics. In fact, the connection between Classical and Quantum Mechanics can be made by the prescription that the Poisson Brackets be replaced with commutators:

$$\{U, V\} \rightarrow \frac{1}{i\hbar} [\hat{U}, \hat{V}] \quad (8.52)$$

where $\hbar$ is Planck's constant divided by 2π . Now, consider the evolution of a system in classical mechanics; in particular, recall how the time-variation of some quantity, $F(q_i, p_i, t)$, is given by its Poisson Bracket with the Hamiltonian:

$$\frac{dF}{dt} = \{F, H\} + \frac{\partial F}{\partial t} \quad (8.53)$$

Now consider the prescription from equation 8.52 to go to Quantum Mechanics:

$$\frac{d\hat{F}}{dt} = \frac{1}{i\hbar} [\hat{F}, \hat{H}] + \frac{\partial \hat{F}}{\partial t} \quad (8.54)$$

This gives precisely the time evolution of a quantum mechanical operator in the Heisenberg formulation. This “Canonical Prescription” is the most general for going from Classical to Quantum Mechanics.

The Hamilton-Jacobi equation and the Schrödinger equation

Consider the time-dependent Schrödinger equation:

$$\begin{aligned}
H(q_i, p_i, t)\psi &= i\hbar \frac{\partial \psi}{\partial t} \\
p_i &= \frac{\hbar}{i} \frac{\partial}{\partial q_i}
\end{aligned} \tag{8.55}$$

and consider a solution of the form:

$$\begin{aligned}
\psi &= Ae^{\frac{i}{\hbar}S(q_i, t)} \\
\therefore \frac{\hbar}{i} \frac{\partial \psi}{\partial q_i} &= \hbar \psi \frac{\partial S}{\partial q_i} \\
\therefore i\hbar \frac{\partial \psi}{\partial t} &= -\hbar \psi \frac{\partial S}{\partial t}
\end{aligned} \tag{8.56}$$

Substituting into the Schrödinger equation:

$$\left(H(q_i, \frac{\partial S}{\partial q_i}, t) + \frac{\partial S}{\partial t} \right) \hbar \psi = 0 \tag{8.57}$$

Since ψ cannot vanish, we obtain the Hamilton-Jacobi equation:

$$H(q_i, \frac{\partial S}{\partial q_i}, t) + \frac{\partial S}{\partial t} = 0 \tag{8.58}$$

There is thus a connection between the Hamilton-Jacobi equation and the Schrödinger equation. Quantum and classical mechanics meet in the limit where $\hbar \rightarrow 0$. The function S is the phase of the wave-function. In the case where we separated out the energy term from S , we are equivalently searching for stationary states in the quantum mechanics formulation.

Sommerfeld and Wilson prescription

The earliest prescription for going from Classical to Quantum Mechanics is due to Sommerfeld and Wilson who postulated that requiring that the action variables be quantized is a sufficient condition to obtain a quantum mechanical description:

$$J_i = \oint p_i dq_i = nh \tag{8.59}$$

where n is an integer and h is Planck's constant. Consider for example the simple harmonic oscillator, where we have explicitly calculated the action variable in [Example 8-6](#):

$$J = 2\pi\alpha \sqrt{\frac{m}{k}} = nh \tag{8.60}$$

where we can identify α with the energy. Writing this in terms of the energy, we have:

$$\alpha = E = n \frac{h}{2\pi} \sqrt{\frac{k}{m}} = n\hbar\omega \quad (8.61)$$

which is the correct quantization for a simple harmonic oscillator (apart for the ground state).

Consider a particle in a central force field, described in polar coordinates, where we have seen that the momentum p_ϕ is conserved and equal to the z-component of angular momentum, L_z . The action variable is easily determined:

$$J_\phi = \oint p_\phi d\phi = \int_0^{2\pi} L_z d\phi = 2\pi L_z \quad (8.62)$$

Applying the Sommerfeld-Wilson quantization prescription, we find that:

$$L_z = n\hbar \quad (8.63)$$

which is the correct quantization for the angular momentum along a specific direction. One should however note that in Classical Mechanics, the total momentum is also equal to L_z and one could be wrongly tempted to conclude that the total angular momentum is quantized in units of $\hbar$.

Finally, consider the fact that the energy can be written in terms of the quantized action variables:

$$E = H(J_1, \dots, J_n) = H(n_1\hbar, \dots, n_n\hbar) \quad (8.64)$$

And consider the change in energy when one of the action variables changes by one unit:

$$J_i \rightarrow J_i + \hbar \quad (8.65)$$

The change in energy of the system is:

$$\Delta E = H(J_1, \dots, J_i + \hbar, \dots, J_n) - H(J_1, \dots, J_i, \dots, J_n) \quad (8.66)$$

If we write the first term as a Taylor series (since $\hbar$ is small):

$$H(J_1, \dots, J_i + \hbar, \dots, J_n) = H(J_1, \dots, J_i, \dots, J_n) + \frac{\partial H}{\partial J_i} \hbar \quad (8.67)$$

hence:

$$\Delta E = \frac{\partial H}{\partial J_i} h = \nu_i h \quad (8.68)$$

where ν_i is the frequency of the corresponding degree of freedom. It should be clear that this is exactly equivalent to the prescription for the Bohr atom.

It should be noted that the Sommerfeld-Wilson prescription is part of the "old quantum mechanics" and was derived in an ad-hoc fashion before a more self-consistent formulation was obtained by Schrödinger, Heisenberg, Dirac and others.

8.6 Problems

Problem 8-1: Invariance of Poisson Brackets

Show that the Poisson Bracket of two functions, U, V , is invariant even under time-dependent canonical transformations, $Q_i(q_i, p_i, t), P_i(q_i, p_i, t)$:

$$\{U, V\}_{q,p} = \{U, V\}_{Q,P}$$

Problem 8-2: Canonical prescription and angular momentum

Use the canonical prescription for quantization to show that the angular momentum in the z-direction is quantized.

Problem 8-3: Canonical transformation types

Show that the transformation equations for type 3 and type 4 canonical transformations are given by the following relations, respectively:

$$F = F_3(p_i, Q_i, t) + \sum_i p_i q_i$$

$$q_i = -\frac{\partial F_3}{\partial p_i}$$

$$P_i = -\frac{\partial F_3}{\partial Q_i}$$

and:

$$F = F_4(p_i, P_i, t) + \sum_i p_i q_i - \sum_i Q_i P_i$$

$$q_i = -\frac{\partial F_4}{\partial p_i}$$

$$Q_i = \frac{\partial F_4}{\partial P_i}$$

Problem 8-4: Identifying canonical transformations

Determine and show which of the following transformations are canonical:

a):

$$Q = \frac{1}{2}(q^2 + p^2)$$

$$P = -\tan^{-1}\left(\frac{q}{p}\right)$$

b):

$$Q = \sqrt{2q}e^t \cos(p)$$

$$P = \sqrt{2q}e^{-t} \sin(p)$$

c):

$$Q = \ln \frac{\sin(p)}{q}$$

$$P = q \cot(p)$$

Problem 8-5: Canonical transformation of a Hamiltonian

a) Show that the following transformation is canonical:

$$Q = \frac{1}{2}(q^2 + p^2)$$

$$P = -\tan^{-1}\left(\frac{q}{p}\right)$$

b) If the Hamiltonian, $H(q, p)$, is given by:

$$H = \frac{1}{2}(q^2 + p^2)$$

Write an expression of the new Hamiltonian, $K(Q, P)$, and write the equations of motion for the new canonical variables (Q, P) .

c) Show that the transformed Hamiltonian, $K(Q, P)$, is a constant of the motion

d) Write an expression for the Lagrangian of this system in terms of q and $\dot{q}$

e) Write out the equation of motion for $\ddot{q}$, and give an example of a physical system that is described by this Lagrangian/Hamiltonian