7 Canonical Transformation

The Hamiltonian mechanics does not reduce the difficulty of solving a given problem. The main advantage of Hamiltonian is that it gives a deeper insight to the structure of mechanics. It sets the stage for constructing of modern theories – statistical mechanics and quantum mechanics to name a few. There is one scenario where the use of Hamiltonian considerably reduces difficulty of the problem – ignorable coordinates. Let say we have a system with two coordinates q_1, q_2. Their conjugate momentum are p_1 and p_2. So the Hamiltonian is

$$H = H(q_1, q_2, p_1, p_2)\,.$$

If q_2 is ignorable then $p_2 = k$ is a constant. So the Hamiltonian is

$$H = H(q_1, p_1, k)\,.$$

So the Hamiltonian depends on just two variables. Let see the same system from the Lagrangian point of view. The generalized velocities are $\dot{q}_1$ and $\dot{q}_2$. Since, q_2 is ignorable, the Lagrangian is

$$L = L(q_1, \dot{q}_1, \dot{q}_2)\,.$$

While the Lagrangian depends on three coordinates the Hamiltonian is a function of two coordinates making it simpler to work with. The Hamiltonian formulation allows us to go from one set of coordinates to another set where the generalized coordinates are cyclic.

In the Kepler problem we may chose a set of generalzed coordinates

$$q_1 = x\,, \quad q_2 = y\,,$$

or

$$q_1 = r\,, \quad q_2 = \theta\,,$$

Both these choices are valid. But in the second choice θ is a cyclic coordinate. The cyclic coordinate depends on the choice of generalized coordinates. In the next sections we develop a method to go from one set of coordinates to another.

7.1 Point Transformation

Suppose there is a Hamiltonian that depends on generalized coordinates q_i and the generalized momentum p_i: $H(q_i, p_i)$. The *canonical equations of motion* or the *Hamiltonian equations of motion* are (following ([80]))

$$\boxed{\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} \end{aligned}}$$

$\dot{p} = -\frac{\partial H}{\partial q_i}$

$\dot{q} = \frac{\partial H}{\partial p_i}$

Any set of coordinates (q_i, p_i) that satisfies the above equation is called *canonical coordinates*. Our aim is to describe the same problem using another set of coordinates (Q_i, P_i). The only requirement to satisfy for (Q_i, P_i) is that they must be canonical – in other words there must be a function $K(Q_i, P_i)$ such that the canonical equations are satisfied

$$\boxed{\begin{aligned} \dot{Q}_i &= \frac{\partial K}{\partial P_i} \\ \dot{P}_i &= -\frac{\partial K}{\partial Q_i} \end{aligned}}$$

The K can be identified with a Hamiltonian.

$q_1 = Q_1(q_1, P_1, t)$

As we saw in the Kepler problem there are relations between (x, y) and (r, θ). In the Hamiltonian formulation coordinates and momentum have the same status, so the most general relations between (q_i, p_i) and (Q_i, P_i) can be written as

$$Q_i = Q_i(q_i, p_i, t), \qquad P_i = P_i(q_i, p_i, t) \tag{81}$$

$P_1 = P_1(q_1, P_1, t)$

These are called *point transformations*.

To find a relation between the old Hamiltonian H and the new Hamiltonian K, we resort to the principle of least action. According to this, for the old set of coordinates

$$\delta S = \delta \int L dt = \delta \int \Big(\sum_i p_i \dot{q}_i - H(q_i, p_i, t) \Big) dt = 0 \tag{82}$$

$\delta S = \delta \int L dt = \delta \int (\Sigma p \dot{q} - H) dt$

In the new coordinates

$$\delta S = \delta \int L' dt = \delta \int \Big(\sum_i P_i \dot{Q}_i - K(Q_i, P_i, t) \Big) dt = 0 \tag{83}$$

The simultaneous validity of the preceding equations mean that the integrands are related by the following relation

$$\lambda \Big(\sum_i p_i \dot{q}_i - H(q_i, p_i, t) \Big) = \Big(\sum_i P_i \dot{Q}_i - K(Q_i, P_i, t) \Big) + \frac{dF}{dt} \tag{84}$$

The F is any function of the phase space coordinates with continuous second derivatives, and λ is independent if the coordinates and time. Though λ can be anything, for which $\lambda = 1$, the transformations (81) are called canonical.

The last term (84) contributes to the variation of the action at the end points and therefore vanishes if F is a function of some combination of old and new coordinates. It acts as a bridge between the old and the new Hamiltonin and is called the *generating function*. To see how a generating function specifies equation of transformation consider F as

$$F = F_1(q_i, Q_i, t) .$$

$F :- f(q_1, Q_1, t)$

From (84) we get

$$\begin{aligned} \sum_i p_i \dot{q}_i - H(q_i, p_i, t) &= \sum_i P_i \dot{Q}_i - K(Q_i, P_i, t) + \frac{dF_1}{dt} , \\ &= \sum_i P_i \dot{Q}_i - K(Q_i, P_i, t) + \frac{\partial F_1}{\partial dt} + \frac{\partial F_1}{\partial q_i} \dot{q}_i + \frac{\partial F_1}{\partial Q_i} \dot{Q}_i . \end{aligned}$$

As the old and new coordinates q_i and Q_i are separately independent the above relation holds identically if the coefficients of $\dot{q}_i$ and $\dot{Q}_i$ vanishes

$$\begin{aligned} p_i &= \frac{\partial F_1}{\partial q_i} , \\ P_i &= -\frac{\partial F_1}{\partial Q_i} , \end{aligned} \tag{85}$$

and

$$K = H + \frac{\partial F_1}{\partial t} . \tag{86}$$

If the F_1 is known the first half of the equations (85) give n relations defining p_i in terms of q_i, Q_i and t. Inverting them the n. Q_i's can be obtained in terms of q_i, p_i and t. So the new Q_i's are completely known in terms of old set (q_i, p_i) and t. These relations can be substituted in the second half of (85) to obtain the n P_i's in terms of q_i, p_i and t. Finally, equation (86) gives the expressions for the new Hamiltonian K.

Example: The Hamiltonian for simple Harmonic oscillator is

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 q^2 \,.$$

Consider a generating function

$$F_1(q, Q, t) = \frac{1}{2} m\omega q^2 \cot Q \,. \tag{87}$$

Then we have

$$\begin{aligned} p_i &= \frac{\partial F_1}{\partial q_i} = m\omega q \cot Q \,, \\ P_i &= -\frac{\partial F_1}{\partial Q_i} = \frac{m\omega q^2}{2\sin^2 Q} \,, \\ K &= H \end{aligned} \tag{88}$$

Using these relations we express the old coordinates in terms of the new coordinates. From the second equation

$$q = \sqrt{\frac{2P}{m\omega}} \sin Q \,.$$

Substituting in the second we get

$$p = \sqrt{2m\omega P} \cos Q \,.$$

Substituting in the Hamiltonian H we get the new Hamiltonian as

$$K = H = \omega P \,.$$

So in the new Hamiltonian the new coordinate Q is cyclic. So the momentum is P constant. As the total energy of the system E is constant, $E = K$

$$P = E/\omega \,.$$

From the Hamilton's EoM

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

which gives a solution

$$Q = \omega t + \phi \,.$$

Substituting P and Q in q we get

$$q = \sqrt{\frac{2E}{m\omega^2}} \sin(\omega t + \phi) \,.$$

This is the usual expression for oscillator.

7.2 Four types of Canonical Transformations

Generating Function	Generating Function Derivatives		Trivial Special Case		
$F = F_1(q, Q, t)$	$p_i = \frac{\partial F_1}{\partial q_i}$	$P_i = -\frac{\partial F_1}{\partial Q_i}$	$F_1 = q_i Q_i$,	$Q_i = p_i$,	$P_i = -q_i$
$F = F_2(q, P, t) - Q_i P_i$	$p_i = \frac{\partial F_2}{\partial q_i}$	$Q_i = \frac{\partial F_2}{\partial P_i}$	$F_2 = q_i P_i$,	$Q_i = q_i$,	$P_i = p_i$
$F = F_3(p, Q, t) + q_i p_i$	$q_i = -\frac{\partial F_3}{\partial p_i}$	$P_i = -\frac{\partial F_3}{\partial Q_i}$	$F_3 = p_i Q_i$,	$Q_i = -q_i$,	$P_i = -p_i$
$F = F_4(p, P, t) + q_i p_i - Q_i P_i$	$q_i = -\frac{\partial F_4}{\partial p_i}$	$Q_i = \frac{\partial F_4}{\partial P_i}$	$F_4 = p_i P_i$,	$Q_i = p_i$,	$P_i = -q_i$

Figure 12: Four basic CTs.

7.3 Conditions for Canonical Transformation

Suppose we have $F = F(q_i, Q_i)$ as a generating function for a canonical transformation. Then $\partial F/\partial t = 0$ and $K = H$ and

$$p_i = \frac{\partial F_1}{\partial q_i}\,, \qquad P_i = -\frac{\partial F_1}{\partial Q_i}\,.$$

$\frac{\partial F}{\partial t} = 0 \quad K = H$

Now

$$dF = \sum_i \frac{\partial F}{\partial q_i} dq_i + \sum \frac{\partial F}{\partial Q_i} dQ_i = \sum_i p_i dq_i - \sum P_i dQ_i\,. \tag{89}$$

This is an exact differential. Therefore, for a function F to generate a canonical transformation, it must be an exact differential. The same can be shown for other three generating functions $F = F_2(q, P) - Q_i P_i$, $F = F_3(p_i, Q_i, t) + q_i p_i$, $F = F_4(p_i, p_i, t) + q_i p_i - Q_i P_i$ – they are exact differential with respect to the coordinates q_i, Q_i.

7.4 Poisson Bracket

If two functions F, G depend on canonical coordinates (q_i, p_i) and time t, then the Poisson's bracket of F and G is defined as

$$[F, G]_{q,p} = \sum_i \left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i} \right).$$

Many books also denote $\{\}$ to denote PB. Also, for brevity superscripts q, p often omitted to indicate PB. From the fundamental definition, following *fundamental PB* can be proven

$$[q_j, q_k] = 0 = [p_j, p_k]\,,$$
$$[q_j, p_k] = \delta_{jk} = -[p_j, q_k]\,,$$

$[u, v] = -[v, u]$

where in the last line we have used the anti-symmetry property

$$[u, v] = -[v, u]\,,$$

where u, v are any two functions of coordinates and momentum. Two other properties of PB include linearity

$$[au + bv, w] = a[u, w] + b[v, w]\,, \quad [uv, w] = u[v, w] + [u, w]v\,,$$

and the Jacobi identity

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$$

7.5 Invariance of PB under CT

In this section we will show that if $(q_i, p_i) \to (Q_i, P_i)$ is a CT then for any two function F, G

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

meaning PB is invariant under CT. To prove it we derive the fundamental PB for the transformed variables

$$[Q_k, Q_l] = [P_k, P_l] = 0\,, \quad \text{and} \quad [Q_k, P_l] = \delta_{kl}\,,$$

From the definition of PB

$$[Q_k, Q_l]_{q,p} = \sum_i \left(\frac{\partial Q_k}{\partial q_i}\frac{\partial Q_l}{\partial p_i} - \frac{\partial Q_k}{\partial p_i}\frac{\partial Q_l}{\partial q_i} \right).$$

From the generating functions F_1 and F_3 we can obtain the following relations

$$\frac{\partial p_k}{\partial Q_l} = \frac{\partial}{\partial Q_l}\frac{\partial F_1}{\partial q_k} = \frac{\partial}{\partial q_k}\frac{\partial F_1}{\partial Q_l} = -\frac{\partial P_l}{\partial q_k}\,,$$
$$\frac{\partial q_k}{\partial Q_l} = -\frac{\partial}{\partial Q_l}\frac{\partial F_3}{\partial p_k} = -\frac{\partial}{\partial p_k}\frac{\partial F_3}{\partial Q_l} = \frac{\partial P_l}{\partial p_k}\,, \tag{90}$$

Substituting (90) in the PB we get

$$\begin{aligned}
[Q_k, Q_l]_{q,p} &= \sum_i \left(-\frac{\partial Q_k}{\partial q_i}\frac{\partial q_i}{\partial P_l} - \frac{\partial Q_k}{\partial p_i}\frac{\partial p_i}{\partial P_l} \right) . \\
&= -\frac{\partial Q_k}{\partial P_l} = 0 \,,
\end{aligned}$$

Since Q, P are independent variable. For the same reason

$$[Q_k, Q_l]_{Q,P} = \sum_i \left(\frac{\partial Q_k}{\partial Q_i}\frac{\partial Q_l}{\partial P_i} - \frac{\partial Q_k}{\partial P_i}\frac{\partial Q_l}{\partial Q_i} \right) = 0 \,.$$

Similarly we can show that

$$[P_k, P_l]_{q,p} = [P_k, P_l]_{Q,P} = 0 \,,$$

Finally using (90) we also get

$$\begin{aligned}
[Q_k, P_l]_{q,p} &= \sum_i \left(\frac{\partial Q_k}{\partial q_i}\frac{\partial P_l}{\partial p_i} - \frac{\partial Q_k}{\partial p_i}\frac{\partial P_l}{\partial p_i} \right) , \\
&= \sum_i \left(\frac{\partial Q_k}{\partial q_i}\frac{\partial q_i}{\partial Q_l} + \frac{\partial Q_k}{\partial p_i}\frac{\partial p_i}{\partial Q_l} \right) = \frac{\partial Q_k}{\partial Q_l} = \delta_{kl} \,,
\end{aligned}$$

and

$$[Q_k, P_l]_{Q,P} = \delta_{k,l} \,,$$

from definition. Therefore, all the fundamental PBs for Q, P are proved. The rest of the proof is given as exercise.