Physics 5153 Classical Mechanics
Canonical Transformations

1 Introduction

The choice of generalized coordinates used to describe a physical system is completely arbitrary, but the Lagrangian is invariant to the choice. To transform the Lagrangian from one set of coordinates to a second set requires knowing the relation between the two sets (point transformation)

$$Q_i = Q_i(q_i, t) \tag{1}$$

and calculating the time derivative since the Lagrangian is a function of the coordinates and velocities. There are n transformation equations, one for each degree of freedom. Since a point transformation leaves the Lagrangian unchanged, it must also leave the Hamiltonian unchanged. But in the case of the Hamiltonian there are $2n$ independent variables, n coordinates and n momenta. This leads to a greater degree of freedom for the transformation. We can include all $2n$ variables in the transformation

$$Q_i = Q_i(p_i, q_i, t) \quad \text{and} \quad P_i = P_i(p_i, q_i, t) \tag{2}$$

This enlarges the possible transformations, but not all these transformations lead to the canonical equations of motion

$$\dot{P} = \frac{\partial K}{\partial Q} \quad \text{and} \quad \dot{Q} = \frac{\partial K}{\partial P} \tag{3}$$

We will look for the condition under which the transformation leads to canonical equations of motion. These transformations will be referred to as canonical transformations. We will start by motivating transformations. We will then discuss the Hamiltonian and its relation to a variational principle; Hamilton's principle. This allows us to discuss the general form of a transformation. From here we will discuss the possible forms of a canonical transformation.

1.1 Motivation for Transformations

Before we start the discussion of canonical transformations, we will provide some motivation for transformations in Hamiltonian mechanics. Let's consider a two dimensional harmonic oscillator. In Cartesian coordinates, the Lagrangian is given by

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - \frac{1}{2}k(x^2 + y^2) \tag{4}$$

where x and y correspond to the displacement from the equilibrium position. In this form the Lagrangian has no cyclic variables so that one might conclude that there are no conserved momenta. If we use the following point transformation

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned} \tag{5}$$

the Lagrangian becomes

$$L = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 - \frac{1}{2}kr^2 \tag{6}$$

In this case the Lagrangian has a cyclic variable θ. Therefore, we have a conserved momentum $p_\theta = mr^2\dot{\theta}$. This simplifies solving the Lagrange equations of motion.

Now let's see what this does for Hamilton's equations of motion. If we take the Cartesian form of the Lagrangian we get the following Hamiltonian

$$H = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{1}{2}k(x^2 + y^2) \tag{7}$$

with the canonical equations of motion

$$\begin{aligned} \dot{p}_x &= -kx & \dot{q}_x &= \frac{p_x}{m} \\ \dot{p}_y &= -ky & \dot{q}_y &= \frac{p_x y}{m} \end{aligned} \tag{8}$$

These can be solved, but it is actually easier to use the Lagrange equations of motion.

Let's consider the transformed Lagrangian. In this case, the Hamiltonian is

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{1}{2}kr^2 \tag{9}$$

The canonical equations of motion for this system are

$$\dot{p}_\theta = 0 \qquad \dot{\theta} = \frac{p_\theta}{mr^2} \tag{10}$$

$$\dot{p}_r = -kr \qquad \dot{r} = \frac{p_r}{m} \tag{11}$$

$$\tag{12}$$

If we assume that $r =$ constant, then the equations on the first line have simple solutions

$$p_\theta = \beta \quad \text{and} \quad \theta = \omega t + \alpha \tag{13}$$

where the three constants α, β, and ω are determined by the initial conditions.

1.2 The Hamiltonian and Hamilton's Principle

To derive the Hamiltonian, we used the Legendre transformation to transform the Lagrangian, which is a function of the coordinates and velocities, to a new function, the Hamiltonian, which is a function of the coordinates and the momenta. Using the total derivative of the Hamiltonian, we derived Hamilton's canonical equations of motion. On the other hand, the Lagrange equations of motion were derived from Hamilton's principle. We would like to show that the canonical equations of motion can also be derived from Hamilton's principle, since this will impose conditions on the type of transformations that keep the equations of motion in canonical form.

Hamilton's principle requires the action

$$I = \int_{t_1}^{t_2} L\, dt \tag{14}$$

have an extremum

$$\delta I = \delta \int_{t_1}^{t_2} L\, dt = 0 \tag{15}$$

Note that when we imposed this condition, it was imposed in configuration space[1]. In the case of the Hamiltonian, we will be imposing the condition on the path followed through phase space[2]. Therefore, we need to express the action as a function of q_i and p_i. This is done through the use of the Hamiltonian

$$H(p_i, q_i, t) = \dot{q}_i p_i - L(\dot{q}_i, q_i, t) \quad \Rightarrow \quad I = \int_{t_1}^{t_2} \left[\dot{q}_i p_i - H(p_i, q_i, t)\right] dt \tag{16}$$

Having written the action as a function of the momentum, the coordinates, and the velocity, we can now deduce the Euler-Lagrange equations for this action

$$\delta \int_{t_1}^{t_2} \left[\dot{q}_i p_i - H(p_i, q_i, t)\right] dt = \delta \int_{t_1}^{t_2} f(\dot{q}_i, q_i, \dot{p}_i, p_i)\, dt = 0 \quad \Rightarrow \quad \begin{cases} \dfrac{d}{dt}\left(\dfrac{\partial f}{\partial \dot{q}_i}\right) - \dfrac{\partial f}{\partial q_i} = 0 \\ \dfrac{d}{dt}\left(\dfrac{\partial f}{\partial \dot{p}_i}\right) - \dfrac{\partial f}{\partial p_i} = 0 \end{cases} \tag{17}$$

If we take the first of the Euler-Lagrange equations and carry through the differentiation, we arrive at

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{q}_i}\right) - \frac{\partial f}{\partial q_i} = 0 \quad \Rightarrow \quad \dot{p}_i + \frac{\partial H}{\partial q_i} = 0 \quad \Rightarrow \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{18}$$

The second equation, which does not explicitly depend on $\dot{p}_i$, gives

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{p}_i}\right) - \frac{\partial f}{\partial p_i} = 0 \quad \Rightarrow \quad -\dot{q}_i + \frac{\partial H}{\partial p_i} = 0 \quad \Rightarrow \quad \dot{q}_i = \frac{\partial H}{\partial p_i} \tag{19}$$

Therefore, we have arrived at Hamilton's canonical equations of motion. Notice that this is independent of the generalized coordinates used. In other words, if we apply a point transformation to the Lagrangian, Hamilton's principle still holds, therefore we arrive at the same canonical equations.

A few comments need to be made about the derivation given above: When the Euler-Lagrange equations were first derived, the variation at t_1 and t_2 on the coordinates was required to be zero. The starting and ending points were required to be the same for all paths. In principle this has to be the same for the momentum, since the procedure to derive both sets of Euler-Lagrange equations is the same. But the term in the derivation that requires this condition is zero already since it depends on the partial derivative of f with respect to $\dot{p}_i$. The function f does not depend on $\dot{p}_i$, therefore the requirement is not necessary.

Even though we are not required to have the path of the particle pass through the same momentum end points in phase space, we will impose this condition none-the-less. This gives us a more general principle, one that is independent of the Lagrangian. Imposing this condition on the end points allows total time derivative of a twice differentiable function to be added to the integrand

[1] Configuration space refers the space of coordinates.

[2] Phase space refers to the space of variables composed of the coordinates and the momenta $\{q_i, p_i\}$.

with out changing the equations of motion

$$
\begin{aligned}
&\delta \int_{t_1}^{t_2} \left[\dot{q}_i p_i - H(p_i, q_i, t) + \frac{dF(p_i, q_i, t)}{dt} \right] dt = \\
&\qquad \delta \int_{t_1}^{t_2} \left[\dot{q}_i p_i - H(p_i, q_i, t) \right] dt + \delta F(p_i, q_i, t)|_{t_1}^{t_2} = \\
&\qquad\qquad \delta \int_{t_1}^{t_2} \left[\dot{q}_i p_i - H(p_i, q_i, t) \right] dt = 0 \qquad (20)
\end{aligned}
$$

where we have imposed the condition that the variation on q_i and p_i are zero at the end points. This condition is important for the discussion of canonical transformations.

1.3 The Canonical Transformation

In our discussion of canonical transformations, we will assume holonomic systems only that can be described by the standard form of the Lagrange and Hamilton equations of motion. Hamilton's principle applies to these systems, therefore if we specify the configuration in terms of some set of generalized coordinates q, then we have

$$
\delta \int_{t_0}^{t_1} L(q, \dot{q}, t)\, dt = 0 \qquad (21)
$$

For a given inertial frame, the value of the Lagrangian is $T - V$, independent of the particular set of generalized coordinates used to specify the configuration of the system. It follows that if a new set of coordinates Q is related to the old set by a point transformation

$$
q_i = q_i(Q, t) \qquad (22)
$$

the resulting Lagrangian $L'(Q, \dot{Q}, t)$ is given by

$$
L'(Q, \dot{Q}, t) = L(q, \dot{q}, t) = T - V \qquad (23)
$$

that is they are equal in value even though they may not be equal in form. In addition, Hamilton's principle applies to the new Lagrangian

$$
\delta \int_{t_0}^{t_1} L'(Q, \dot{Q}, t)\, dt = 0 \qquad (24)
$$

But one might ask if there is a more general transformation that might also satisfy Hamilton's principle. For example, one might add a total time derivative of a function of the coordinates

$$
L'(Q, \dot{Q}, t) = L(q, \dot{q}, t) - \frac{d}{dt} F(q, Q, t) \qquad (25)
$$

Applying Hamilton's principle to this Lagrangian gives

$$
\begin{aligned}
&\delta \int_{t_0}^{t_1} L'(Q, \dot{Q}, t) = \delta \int_{t_0}^{t_1} L(q, \dot{q}, t) - \int_{t_0}^{t_1} \frac{d}{dt} F(q, Q, t) \\
&\qquad \Rightarrow \quad \delta \int_{t_0}^{t_1} L'(Q, \dot{Q}, t) = \delta \int_{t_0}^{t_1} L(q, \dot{q}, t) - F(q, Q, t)|_{t_0}^{t_1} \qquad (26)
\end{aligned}
$$

But the last term is zero, since the end points are fixed. Therefore, we are back to where we started with Hamilton's principle

$$\delta \int_{t_0}^{t_1} L'(Q, \dot{Q}, t) = \delta \int_{t_0}^{t_1} L(q, \dot{q}, t) = 0 \tag{27}$$

Note that this new Lagrangian ($L'(Q, \dot{Q}, t)$) does not have the same form nor the same value as the original Lagrangian ($L(q, \dot{q}, t)$) and therefore does not have the value $T - V$ in the original inertial frame.

Now consider the Hamiltonians associated with these Lagrangians

$$H(q, p, t) = \sum_{i=1}^{n} p_i \dot{q}_i - L(q, \dot{q}, t) \tag{28}$$

$$K(Q, P, t) = \sum_{i=1}^{n} P_i \dot{Q}_i - L'(Q, \dot{Q}, t) \tag{29}$$

where the generalized momenta are given by

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \qquad\qquad P_i = \frac{\partial L'}{\partial \dot{Q}_i} \tag{30}$$

Since Hamilton's principle applies to both Lagrangian's, the canonical equations of motion are valid for both descriptions

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad\qquad \dot{q}_i = \frac{\partial H}{\partial p_i} \tag{31}$$

$$\dot{P}_i = -\frac{\partial K}{\partial Q_i} \qquad\qquad \dot{Q}_i = \frac{\partial K}{\partial P_i} \tag{32}$$

We define a canonical transformation, as a transformation that takes the Hamiltonian $H(q, p, t)$ to the Hamiltonian $K(Q, P, t)$ such that they both satisfy the canonical equations of motion.

Let's consider a specific Hamiltonian $H(q, p, t)$. In addition, suppose we are given specific transformations $Q_i = Q_i(q, p, t)$ and $P_i = P_i(q, p, t)$. Using Eq. 25, and writing the Lagrangians in terms of the Hamiltonians, we arrive at the following expression

$$\sum_{i=1}^{n} p_i dq_i - H dt - \sum_{i=1}^{n} P_i dQ_i + K dt = dF \tag{33}$$

This equation provides a sufficient condition for a canonical transformation. The function $F(q, Q, t)$ is called the generating function for this transformation. Let's now expand $dF(q, Q, t)$

$$dF = \sum_{i=1}^{n} \left[\frac{\partial F}{\partial q_i} dq_i + \frac{\partial F}{\partial Q_i} dQ_i \right] + \frac{\partial F}{\partial t} dt \tag{34}$$

Comparing this equation to Eq. 33, we arrive at the following relations

$$p_i = \frac{\partial F}{\partial q_i} \qquad\qquad P_i = -\frac{\partial F}{\partial Q_i} \qquad\qquad K = H + \frac{\partial F}{\partial t} \tag{35}$$

These provide the relation between the generating function and the canonical variables. The function $F(q, Q, t)$ is normally referred to as $F_1(q, Q, t)$.

If one is given the transformation equation, then Eq. 33 provides the condition that the transformation is canonical. Note since this case assume that the generating function is not known, this equation states that the left hand side must be a complete differential, and this is the generating function. On the other hand, if one is given the generating function, then Eq. 35 gives the transformation equations.

1.4 Example

As an example of canonical transformations, and how they can simplify a problem, we will consider the case of the simple harmonic oscillator. Since the system is conservative, the Hamiltonian for this system is the sum of the kinetic and potential energies

$$H = \frac{p^2}{2m} + \frac{1}{2}kq^2 = \frac{1}{2m}\left[p^2 + m^2\omega^2 q^2\right] \tag{36}$$

One of the stated goals of transforming the Hamiltonian to a new set of coordinates is to increase the number of cyclic variables. In this case, the Hamiltonian is quadratic in the momentum and the spatial coordinates therefore one might consider a transformation of the form

$$p = f(P)\cos Q \qquad\qquad q = \frac{f(P)}{m\omega}\sin Q \tag{37}$$

which leads to

$$K = H = \frac{f^2(P)}{2m} \tag{38}$$

The next step is to select $f(P)$ such that the transformation is canonical. For this we will select the generating function to be[3]

$$F_1(q, Q) = \frac{m\omega q^2}{2}\cot Q \tag{39}$$

Since this is the same type of transformation as discussed above ($F_1(q, Q)$), Eq. 35 gives the canonical momenta

$$p = \frac{\partial F_1}{\partial q} = m\omega q\cot Q \qquad\qquad P = -\frac{\partial F_1}{\partial Q} = \frac{m\omega q^2}{2\sin^2 Q} \tag{40}$$

From these two equations we can solve for q and p

$$\left.\begin{aligned} q &= \sqrt{\frac{2P}{m\omega}}\,\sin Q \\ p &= \sqrt{2Pm\omega}\,\cos Q \end{aligned}\right\} \quad\Rightarrow\quad f(P) = \sqrt{2m\omega P} \tag{41}$$

Therefore the Hamiltonian is

$$K = H = \omega P \quad\Rightarrow\quad P = \frac{E}{\omega} \tag{42}$$

[3] We will later see how to arrive at the generating function in a systematic manner.

where we have used the fact that a conservative system is independent of time, and since it is also a natural system, the Hamiltonian is the total energy.

We are now in a position to solve the equations of motion. Since the variable Q is cyclic, the momentum is a constant. Therefore, the only equation of motion that we need to solve is

$$\dot{Q} = \frac{\partial H}{\partial P} = \omega \quad \Rightarrow \quad Q = \omega t + \alpha \tag{43}$$

Substituting this back into Eq. 41, we arrive at

$$q = \sqrt{\frac{2E}{m\omega^2}}\sin(\omega t + \alpha) \qquad\qquad p = \sqrt{2mE}\cos(\omega t + \alpha) \tag{44}$$

1.5 Other Generating Functions

In the derivation given above, we used a generating function made up of the original and new position coordinates. In principle, we could have selected any $2n$ independent variables[4]. Therefore, we could select any of the following generating functions

$$\begin{array}{ccc} F_1(q,Q) & F_2(q,P) & F(q,p) \\ F_3(p,Q) & F_4(p,P) & F(Q,P) \end{array}$$

Since we want to relate the old variables to the new variables, the choice is to take half the old variables and half the new variables. From the functions given above, we take those in the first two columns. Keep in mind that the old and new position or momentum coordinates need not be independent. For instance, the new coordinates Q could be a linear combination of the old position coordinates. In this case one would select $F(q,P)$ or $F(p,Q)$, assuming that p and P are also not linearly independent.

Let's examine the relationship between the generating functions $F_1(q,Q,t)$ and $F_2(q,P,t)$. We will start with the basic differential form

$$dF_1(q,Q,t) = Kdt - Hdt + \sum_{i=1}^{n} p_i dq_i - \sum_{i=1}^{n} P_i dQ_i \tag{45}$$

The goal is to replace Q with P. This can be done by adding the following total differential to Eq. 45

$$d\left(\sum_{i=1}^{n} P_i Q_i\right) = \sum_{i=1}^{n} P_i dQ_i + \sum_{i=1}^{n} Q_i dP_i \quad \Rightarrow \quad dF_2(q,P,t) = Kdt - Hdt + \sum_{i=1}^{n} p_i dq_i + \sum_{i=1}^{n} Q_i dP_i \tag{46}$$

this being exactly what we were looking for. Hence, the two generating functions are related by

$$F_2(q,P,t) = F_1(q,Q,t) + \sum_{i=1}^{n} Q_i P_i \tag{47}$$

[4] $2n$ is the maximum number of independent variable we can have.

Notice, this is a Legendre transform from the variable Q to the variable P. To determine the relation between the old and new coordinates, we write the total derivative of F_2

$$dF_2(q,P,t) = \sum_{i=1}^{n}\left(\frac{\partial F_2}{\partial q_i}dq_i + \frac{\partial F_2}{\partial P_i}dP_i\right) + \frac{\partial F_2}{\partial t}dt \tag{48}$$

and compare the coefficients of the differentials assuming that q_i and P_i are independent

$$p_i = \frac{\partial F_2}{\partial q_i} \qquad Q_i = \frac{\partial F_2}{\partial P_i} \qquad K = H + \frac{\partial F_2}{\partial t} \tag{49}$$

1.6 Poisson Bracket and Canonical Transformations

We previously showed that a transformation is canonical if Eq 33 is satisfied[5]. But the time is not transformed, hence we can consider a variation on F at a fixed time

$$\delta F_1(q,Q,t) = \sum_{i=1}^{n}(p_i\delta q_i - P_i\delta Q_i) \tag{50}$$

Next, since the number of independent variables required to define the system is $2n$ and we can use any set of $2n$ independent variables, we also use any set to test for exactness. Hence, we will use the variables q_i and p_i

$$\delta G(q,p) = \sum_{i=1}^{n}(p_i\delta q_i - P_i\delta Q_i) = \sum_{i=1}^{n}\left(\frac{\partial G}{\partial q_i}\delta q_i + \frac{\partial G}{\partial p_i}\delta p_i\right) \tag{51}$$

under the condition that two functions have equal values $G(q,p) = F(q,Q)$. Next we expand the variations in Eq. 50 in terms of q and p

$$\delta G(q,p) = \sum_{i=1}^{n}\left[p_i\delta q_i - P_i\left(\frac{\partial Q_i}{\partial q_i}\delta q_i + \frac{\partial Q_i}{\partial p_i}\delta p_i\right)\right] \tag{52}$$

Comparing this equation to Eq. 51, we arrive at the following relations

$$\frac{\partial G}{\partial q_i} = p_i - gPi\frac{\partial Q_i}{\partial q_i} \qquad \frac{\partial G}{\partial p_i} = -P_i\frac{\partial Q_i}{\partial p_i} \tag{53}$$

Finally, the condition for an exact differential requires the following

$$\frac{\partial}{\partial p_i}\left(\frac{\partial G}{\partial q_i}\right) = \frac{\partial}{\partial q_i}\left(\frac{\partial G}{\partial p_i}\right) \quad \Rightarrow \quad \frac{\partial}{\partial p_i}\left(p_i - P_i\frac{\partial Q_i}{\partial q_i}\right) = -\frac{\partial}{\partial q_i}\left(P_i\frac{\partial Q_i}{\partial p_i}\right) \tag{54}$$

This is the condition for a transformation to be canonical assuming an F_1 generating function.

A second and more straight forward method is as follows. A canoncal transformation requires that the transformed variables satisfy the canonical equations of motion. W start by looking at the rate of change of the two transformaed variables

$$\begin{aligned}\frac{dQ}{dt} &= [Q,H]_{q,p} = \frac{\partial Q}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial Q}{\partial p_i}\frac{\partial H}{\partial q_i} \\ \frac{dP}{dt} &= [P,H]_{q,p} = \frac{\partial P}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial P}{\partial p_i}\frac{\partial H}{\partial q_i}\end{aligned} \tag{55}$$

[5] In fact, any complete differential of the generating function will do.

The partial derivatives of the Hamiltonian can be rewritten as follows

$$\begin{aligned}\frac{\partial H}{\partial q} &= \frac{\partial K}{\partial Q}\frac{\partial Q}{\partial q} + \frac{\partial K}{\partial P}\frac{\partial P}{\partial p} \\ \frac{\partial H}{\partial p} &= \frac{\partial K}{\partial Q}\frac{\partial Q}{\partial p} + \frac{\partial K}{\partial P}\frac{\partial P}{\partial q}\end{aligned} \tag{56}$$

Substitute this into Eq. 55

$$\begin{aligned}\frac{dQ}{dt} &= \frac{\partial K}{\partial P}\left(\frac{\partial Q}{\partial q_i}\frac{\partial P}{\partial p_i} - \frac{\partial Q}{\partial p_i}\frac{\partial P}{\partial q_i}\right) = \frac{\partial K}{\partial P}\,[Q, P]_{q,p} \\ \frac{dP}{dt} &= -\frac{\partial K}{\partial Q}\left(\frac{\partial Q}{\partial q_i}\frac{\partial P}{\partial p_i} - \frac{\partial Q}{\partial p_i}\frac{\partial P}{\partial q_i}\right) = -\frac{\partial K}{\partial Q}\,[Q, P]_{q,p}\end{aligned} \tag{57}$$

which implies that the Poisson bracket must be equal to 1. Therefore, a canonical transformation must satisfy the standard Poisson bracket relations for the canonically conjugate variables.