Canonical Transformations

Toki Nishikawa

May 2021

1 Abstract

The reader should achieve a general understanding of canonical transformations and some of their most useful applications. The paper will walk through a derivation of the transformation equations associated with some of the generating functions that serve to mediate a particular type of canonical transformation, i.e. those which make use of cyclic coordinates to simplify the Hamiltonian. The remainder of the discussion will be focused on canonical transformations that turn all coordinates into constants given by initial conditions. This is illustrated with a specific generating function known as the *Hamilton action function* which can be solved for using the *Hamilton-Jacobi differential equation*. An example of both of these processes will be given through an analysis of the harmonic oscillator.

2 Introduction

The Hamiltonian formalism is equivalent to the Lagrangian formalism, just expressed in terms of position and momentum as opposed to position and velocity. One benefit of the Hamiltonian is that we can make use of cyclic (or "ignorable") coordinates. Suppose a Lagrangian did not depend on some coordinate q_i. Then, according to the Euler-Lagrange equation, $\dot{p}_i = \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}) = \frac{\partial L}{\partial q_i} = 0$ and the corresponding momentum p_i is conserved. The Lagrangian will then be a function of 2n-1 coordinates, i.e. $L = L(q_1, ..., q_{n-1}; \dot{q}_i)$. If we switch to the Hamiltonian formalism, we see that $\dot{p}_i = -\frac{\partial H}{\partial q_i} = 0$ and again, the corresponding momentum $p_i = k$ is conserved. But, in contrast to the Lagrangian, the Hamiltonian will be a function of 2n-2 coordinates, i.e. $H = H(q_1, ..., q_{n-1}; p_1, ..., p_{n-1}, k)$. Thus, the presence of a cyclic coordinate has simplified the Hamiltonian by a single degree of freedom.

The conversion from Lagrangian to Hamiltonian is not the only way that we can exploit cyclic coordinates. There are cases where the Hamiltonian will have more cyclic coordinates in one coordinate system relative to another (we will see an example of this soon). The coordinate transformations we are looking for are known as *canonical transformations*. Although this term describes a much broader range of transformations, we will limit our discussion to those which yield cyclic coordinates or constants that are given by initial conditions.

3 Canonical Transformations

Let's consider a general Hamiltonian $H = H(q_i, p_i, t)$ with the following Hamilton equations:

$\dot{p}_i = -\frac{\partial H}{\partial q_i}$ and $\dot{q}_i = \frac{\partial H}{\partial p_i}$

We know that, for cyclic coordinates, it must be true that:

$\frac{\partial H}{\partial q_i} = 0$ and $\dot{p}_i = 0$

i.e., the Hamiltonian does not depend on a cyclic coordinate q_i. We can simplify any given mechanical problem by adopting a coordinate system that contains these cyclic coordinates. Consider the following example in 2-dimensional space:

A particle is subject to a central potential. If we choose to describe this motion in Cartesian coordinates, the Lagrangian will be of the form $L = L(x, y, \dot{x}, \dot{y})$ while the associated Hamiltonian will be of the form $H = H(p_x, p_y, x, y)$. However, if we switch to polar coordinates, the the Lagrangian will be

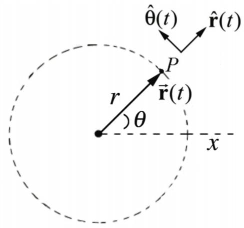

Figure 1: A particle moving in 2-dimensional space represented in polar coordinates where p_θ points in the direction of $\hat{\theta}$ and p_r points in the direction of $\hat{r}$.

of the form $L = L(r, \dot{r}, \dot{\theta})$ while the associated Hamiltonian will be of the form $H = H(p_r, p_\theta, r)$. Since the new Lagrangian and Hamiltonian do not depend on θ, we know that the corresponding momentum p_θ will be a constant. The Hamiltonian can then be written as $H = H(p_r, r)$ and has been effectively reduced to a single degree of freedom. The conservation of angular momentum p_θ can also be viewed as a product of rotation invariance. A central potential implies that the potential energy $U = U(r)$; i.e., the potential energy is only a function of the distance from the particle to the axis of rotation. Therefore, switching to polar coordinates allows us to exploit conservation of angular momentum and simplify our Hamiltonian.

A new set of coordinates

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

where Q_i is cyclic, would greatly reduce the complexity of our Hamiltonian. If we are able to convert all q_i to cyclic coordinates Q_i, then all momenta P_i are constant and the new Hamiltonian can be written as $H' = H'(P_i)$.

We define a canonical transformation of the form shown in **Equation 1** that yields new coordinates P_i and Q_i which satisfy the following canonical Hamilton equations:

$\dot{Q}_i = \frac{\partial H'(P_i)}{\partial P_i} = w_i = const$ and $\dot{P}_i = -\frac{\partial H'(P_i)}{\partial Q_i} = 0$

Integrating these equations with respect to time, we can retrieve the following equations of motion:

$Q_i = w_i t + w_0$ and $P_i = \beta_i = const$

A pair of q_i and p_i are defined as canonically conjugate if they satisfy Hamilton's equations. A canonical transformation is simply a transformation from one canonically conjugate pair to another. We will focus our discussion on canonical transformations that serve to simplify the Hamiltonian by converting q_i's to cyclic coordinates. In order to identify a function that will carry out this transformation between coordinates, we use the Hamilton principle. The next section will be dedicated to deriving the Hamilton Principle from Newton's equations.

Derivation of Hamilton's Principle from Newton's Equations for a Single Particle

Let's consider a particle moving along an arbitrary path $\mathbf{r} = \mathbf{r}(t)$ between $\mathbf{r}(t_1)$ and $\mathbf{r}(t_2)$. Now consider a varied path with displacement $\delta\mathbf{r}$ that has the same start and end points. For the new path, we have that $\mathbf{r} = \mathbf{r}(t) + \delta\mathbf{r}(t)$ and $\delta\mathbf{r}(t_1) = \delta\mathbf{r}(t_2) = 0$. The work required to vary the path by displacement $\delta\mathbf{r}$ can be written as $\delta A = \mathbf{F} \cdot \delta\mathbf{r} = \mathbf{F}^e \cdot \delta\mathbf{r}$ where $\mathbf{F}^e$ represents an external conservative force. For conservative forces, work can be written as the negative of the change in potential energy, i.e.

$F^e \cdot \delta\mathbf{r} = -\delta V$. Using Newton's Second Law, we obtain,

$m\ddot{\mathbf{r}} \cdot \delta\mathbf{r} = -\delta V$. If we consider the following equation:

$\frac{d}{dt}(\dot{\mathbf{r}} \cdot \delta\mathbf{r}) = \dot{\mathbf{r}} \cdot \frac{d}{dt}\delta\mathbf{r} + \ddot{\mathbf{r}} \cdot \delta\mathbf{r} = \dot{\mathbf{r}} \cdot \delta\dot{\mathbf{r}} + \ddot{\mathbf{r}} \cdot \delta\mathbf{r} = \delta(\frac{1}{2}\dot{\mathbf{r}}^2) + \ddot{\mathbf{r}} \cdot \delta\mathbf{r}$ which we can rewrite as $m\ddot{\mathbf{r}} \cdot \delta\mathbf{r} = m\frac{d}{dt}(\dot{\mathbf{r}} \cdot \delta\mathbf{r}) - \delta(\frac{1}{2}m\dot{\mathbf{r}}^2)$

we obtain,

$m\frac{d}{dt}(\dot{\mathbf{r}}\cdot\delta\mathbf{r}) - \delta(m\frac{1}{2}\dot{\mathbf{r}}^2) = -\delta V$. Recognizing that $\delta(m\frac{1}{2}\dot{\mathbf{r}}^2)$ is the variation in kinetic energy, we can simplify further

$\delta(T-V) = \delta L = m\frac{d}{dt}(\dot{\mathbf{r}}\cdot\delta\mathbf{r})$. Integrating with respect to time yields the following:

$\delta\int_{t_1}^{t_2} Ldt = m[\dot{\mathbf{r}}\cdot\delta\mathbf{r}]_{t_1}^{t_2} = 0$ where the last step is computed by recalling that $\delta\mathbf{r}(t_1) = \delta\mathbf{r}(t_2) = 0$.

Although this derivation was based on a single particle, the result can be generalized to systems with many particles. We will now use this result to further our discussion on canonical transformations.

Since these new coordinates must also obey the Hamilton principle, we retrieve the following equations:

$\delta\int L(q_i,\dot{q}_i,t)dt = 0$ and $\delta\int L'(Q_i,\dot{Q}_i,t)dt = 0$

where L corresponds to H and L' corresponds to H'. Therefore,

$\delta\int(L-L')dt = 0$

This result will be true as long as L and L' only differ by a total differential, i.e. $\delta\int_a^b \frac{dF}{dt}dt = \delta(F(b) - F(a)) = 0$ where the last step is computed by recognizing that the variation of a constant will always be zero.

We are searching for a so-called *generating function* F that serves to mediate the canonical transformation. Let's suppose that F will be a function of both the new and old coordinates $F = F(p_i, q_i, P_i, Q_i, t)$. If i=1,2,...,n, then F is a function of 4n+1 coordinates. However, there are only 2n+1 independent variables since Q_i and P_i are already functions of p_i, q_i, and t. As long as F contains an old coordinate and a new coordinate, we will be able to create a relationship between the two coordinate systems. The following generating functions meet this criteria: $F_1 = F(q_i, Q_i, t)$, $F_2 = F(q_i, P_i, t)$, $F_3 = F(p_i, Q_i, t)$, $F_4 = F(p_i, P_i, t)$. Let's see how these functions yield the transformations described in **Equation 1**.

Using the definition of the Hamiltonian $H = \Sigma p_i\dot{q}_i - L$ and the fact that L and L' only differ by a total differential $L = L' + \frac{dF}{dt}$, we see that

$\Sigma p_i\dot{q}_i - H = \Sigma P_i\dot{Q}_i - H' + \frac{dF}{dt}$.

Let's consider that $F = F_1 = F(q_i, Q_i, t)$, then $\frac{F_1}{dt} = \Sigma\frac{\partial F_1}{\partial q_i}\dot{q}_i + \Sigma\frac{\partial F_1}{\partial Q_i}\dot{Q}_i + \frac{\partial F_1}{\partial t}$ by the multivariable chain rule. Thus,

$\Sigma p_i\dot{q}_i - \Sigma P_i\dot{Q}_i - H + H' = \Sigma\frac{\partial F_1}{\partial q_i}\dot{q}_i + \Sigma\frac{\partial F_1}{\partial Q_i}\dot{Q}_i + \frac{\partial F_1}{\partial t}$.

If we compare the coefficients on either side of the equation, we retrieve the following transformation equations associated with F_1:

$$p_i = \frac{\partial F_1}{\partial q_i}(q_j, Q_j, t), P_i = -\frac{\partial F_1}{\partial Q_i}(q_j, Q_j, t), H' = H + \frac{\partial F_1}{\partial t}(q_j, Q_j, t) \quad (2)$$

where the first two equations allow us to create the transformation equations of the form seen in [**Equation 1**] while substitution into the third equation yields the new Hamiltonian. We will see an example of this now.

Example: The Harmonic Oscillator

Considering generalized coordinates q and p for our system, we can represent the kinetic/potential energies and the associated Hamiltonian as follows:

$T = \frac{1}{2m}p^2$, $U = \frac{1}{2}kq^2 = \frac{1}{2}mw^2q^2$, $H = \frac{1}{2m}p^2 + \frac{1}{2}mw^2q^2$

where the potential energy is written in terms of mass and angular velocity by the substitution $w = \sqrt{\frac{k}{m}}$.

Suppose we are given the generating function $F(q,Q) = \frac{m}{2}wq^2cot(Q)$. Recognizing that $F(q,Q)$ is of the form F_1, we can use **Equation 2**. Thus,

$p = \frac{\partial F}{\partial q} = mwqcot(Q)$, $P = -\frac{\partial F}{\partial Q} = \frac{m}{2}\frac{wq^2}{sin^2(Q)}$.

The above equations can be trivially solved for q and p, yielding the following:

$q = \sqrt{\frac{2}{mw}P}sin(Q)$, $p = \sqrt{2mwP}cos(Q)$

By plugging these into H, we obtain $H' = Pw(cos^2(Q) + sin^2(Q)) = wP$. Since $H' = H'(P)$, Q is a cyclic ("ignorable") coordinate. Thus, we have shown that the generating function F produces a new Hamiltonian that only depends on the constant momentum P. Using this information, we can easily formulate the equation of motion as a function of time. Since there is no explicit time dependence in the Hamiltonian, the total energy E is conserved where $E = H = \omega P$. By the Hamilton equation $\dot{Q} = \frac{\partial H'}{\partial P} = \omega$, we find that $Q = \omega t + \phi$. Plugging these results into $q = \sqrt{\frac{2}{mw}P}sin(Q)$, we obtain the following: $q = \sqrt{\frac{2E}{m\omega^2}}sin(\omega t + \phi)$.

Using a similar derivation, we can find the transformation equations associated with $F_2 \equiv S = S(q_i, P_i, t)$:

$$p_i = \frac{\partial S}{\partial q_i}(q_j, P_j, t), Q_i = \frac{\partial S}{\partial P_i}(q_j, P_j, t), H' = H + \frac{\partial S}{\partial t}(q_i, P_i, t) \tag{3}$$

As in the previous case, the first two equations allow us to create the transformation equations of the form seen in [**Equation 1**] while substitution into the third equation yields the new Hamiltonian. We will use [**Equation 3**] in our discussion of Hamilton-Jacobi Theory.

4 Hamilton-Jacobi Differential Equation

So far, we have considered canonical transformations that yield $H' = H'(P_i)$, i.e. a Hamiltonian solely dependent on constant momenta P_i. Now we will consider canonical transformations that produce new coordinates $P_i = p_{i0}$ and $Q_i = q_{i0}$ where p_{i0} and q_{i0} are constants given by initial conditions. Once these coordinates have been identified, we express the transformation equations as follows:

$$q_i = q_i(q_{i0}, p_{i0}, t), p_i = p_i(q_{i0}, p_{i0}, t) \tag{4}$$

Since P_i and Q_i are constants, Hamilton's equations tell us the following:

$\dot{P}_i = 0 = -\frac{\partial H'}{\partial Q_i}$ and $\dot{Q}_i = 0 = \frac{\partial H'}{\partial P_i}$.

Let's define $H' \equiv 0$, since this will definitely satisfy the previous equations. We will need a generating function to carry out the coordinate transformation. Jacobi chose the form $F_2 = S(q_i, P_i, t)$ to be the so-called *Hamilton action function* and we will use this for the remainder of our discussion. Recognizing that **Equation 3** holds for this generating function, we can rewrite the new Hamiltonian as $\frac{\partial S(q_i,P_i,t)}{\partial t} + H(q_i, p_i, t) = 0$. Using the fact that $p_i = \frac{\partial S}{\partial q_i}$ (as seen in equation **Equation 3**) and that $P_i = p_{i0} = \beta_i$ since the new momenta are constant, we retrieve the following equation known as the *Hamilton-Jacobi differential equation*:

$$\frac{\partial S(q_i, P_i = \beta_i, t)}{\partial t} + H(q_i, \frac{\partial S}{\partial q_i}, t) = 0 \tag{5}$$

By solving this first order nonlinear partial differential equation, we can determine the generating function $S = S(q_i, \beta_i, t)$. Using $Q_i = \frac{\partial S}{\partial P_i} = \frac{\partial S}{\partial \beta_i} = \alpha_i$ from **Equation 3**, we can retrieve the transformation equations described in **Equation 4** where q_{i0} is analogous to α_i and p_{i0} is analogous to βi. Let's see an example of this.

Example: Revisiting The Harmonic Oscillator

Recall that $H = \frac{p^2}{2m} + \frac{k}{2}q^2$. Using the Hamilton action function $S = S(q, P, t)$ and its associated transformation equation $p = \frac{\partial S}{\partial q}$ from **Equation 3**, we can write the Hamilton-Jacobi differential equation as:

$$\frac{\partial S}{\partial t} + \frac{1}{2m}(\frac{\partial S}{\partial q})^2 + \frac{k}{2}q^2 = 0$$

Assuming that $S = S_1(t) + S_2(q)$, we can define the following partial derivatives: $\frac{\partial S}{\partial t} = \frac{\partial S_1(t)}{\partial t}$ and $\frac{\partial S}{\partial q} = \frac{\partial S_2(q)}{\partial q}$. We can then simplify the differential equation to

$$-\dot{S}_1(t) = \frac{1}{2m}(\frac{\partial S_2(q)}{\partial q})^2 + \frac{k}{2}q^2 = \beta$$

Since the left-hand-side of the equation only depends on t while the right-hand-side of the equation only depends on q, they will only be equal if they are both some constant β. Thus, for the left-hand-side, we retrieve $-\dot{S}_1(t) = -\beta$ which, by integrating, is equivalent to $S_1(t) = -\beta t$. As for the right-hand-side, we retrieve $\frac{1}{2m}(\frac{\partial S_2(q)}{\partial q})^2 + \frac{k}{2}q^2 = \beta$ which, by isolating $\frac{\partial S_2(q)}{\partial q}$, is equivalent to $\frac{\partial S_2}{\partial q} = \sqrt{2m\beta - mkq^2}$. Solving for S_2, we obtain $S_2 = \sqrt{mk}\int\sqrt{\frac{2\beta}{k} - q^2}dq$. S then becomes

$$S = S_1(t) + S_2(q) = \sqrt{mk}\int\sqrt{\frac{2\beta}{k} - q^2}dq - \beta t = S(q, \beta, t)$$

Earlier we saw that $Q = \frac{\partial S}{\partial \beta} = \alpha$. And so $\alpha = \frac{\sqrt{mk}}{k}\int(\frac{2\beta}{k} - q^2)^{-\frac{1}{2}}dq - t$. Solving for q, we obtain

$q = \sqrt{\frac{2\beta}{k}} sin\omega(t + \alpha)$ where β and α are given by initial conditions. If we compare this equation to the known equation of motion for a harmonic oscillator, we discover that β represents the total energy of the system while α represents the initial time. Thus, we used the Hamilton-Jacobi differential equation to determine a function S that mediates the transformation between p, q and β, α where the latter coordinates are given by initial conditions.

5 Conclusion

We have seen that canonical transformations, specifically those which yield cyclic coordinates or constants that are given by initial conditions, can greatly simplify a Hamiltonian. In the latter case, we showed that the *Hamilton-Jacobi differential equation* can be used to solve for a particular generating function know as the *Hamilton action function* – from which we were able to retrieve an equation of motion dependent on some initial conditions. It turns out that this differential equation is very powerful for other reasons and has applications that reach far beyond the scope of this paper. It is worth mentioning that this differential equation is the only formulation of mechanics that can express particles as waves and is therefore considered the closest classical analogy to quantum mechanics.

6 References

Goldstein, Herbert. Classical Mechanics. 2d ed., Addison-Wesley Pub. Co., 1980.

Greiner, Walter. Classical Mechanics. Springer Berlin / Heidelberg, 2010.

Landau, L. D., and Lifshits, E. M. Mechanics. Addison-Wesley, 1960.