# Classical Mechanics

Ting-Kai Hsu

2023/7/6

## Contents

# 1 Generalized Coordinates and Generalized Velocities

The number of **independent quantities**[1] is called the number of **degree of freedom**. And in system with *s* degrees of freedom, the independent quantities are called **generalized coordinates**, denoted as $q_i$, and the derivative of generalized coordinate $\dot{q}_i$. Then we should introduce an **experience law** which is worth studying, that is, it a system with generalized coordinates and generalized velocities simultaneously determined, then the state of the system would be **completely determined** and we'll be able to describe its motion.

# 2 Introduction to The Principle of Least Action

In this section, we'll see another way to derive **Lagrange's equation** with the *principle of least action* or *Hamilton's principle*, which governs the motion of mechanical system. And one must preview and know what is and the meaning of **calculus of variation**[2] The essence of calculus of variation is that **at minimum**, the deviation of the function is only *second order* if the difference that we go away is in the *first* order. In other word, we'd find out, in the first approximation, no difference at the minimum value.[3]

# 3 Notes on Reading Landau

At the end of section 2 in chapter 1[4], Landau has mentioned that if we shift the Lagrangian $L$ by a **total time derivative of any function of coordinates and time**, that is,

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d}{dt} f(q, t)$$

[1] which can uniquely described the position of any given system.

[2] recommandation: Feyman *vol2 ch19*, and Gelfand *Calculus of Variation*

[3] A short proof is that if we assume at minimum, the approximation has *first order* term. If we first move forward, we should assume the approximation become larger since we're at minimum; however, if we reverse the direction with the same ammount of difference, we'll get the *first order* term becoming smaller, that is, the approximation become smaller when we move backward, which means we're not at the minimum at all.

[4] page 4

There is two ways to prove it, by with the same core of concept. One use the least action principle and the other makes use of the equation of motion.
**One:** first write out the action,

$$S' = \int_{t_1}^{t_2} L'(q, \dot{q}, t)\, dt = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt + \int_{t_1}^{t_2} \frac{d}{dt} f(q, t)\, dt$$

As we can see, the rightest term of the eqaution can be rewritten as:

$$S' = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt + f(q^{(2)}, t_2) - f(q^{(1)}, t_1)$$

Where $q^{(1)}$ is the initial position of the particle, and $q^{(2)}$ is its final position. Then we shall see that $f(q^{(2)}, t_2) - f(q^{(1)}, t_1)$ is constant which would be cancel when calculating the variation of $S$. Therefore, we can see that adding a **total time derivate of any function of generalized coordinates and time** wouldn't affect the equation of motion.

**Second:** Then we'd use another point of veiw to prove it, that is, we just take $L'$ directly into equation of motion and see if it doesn't change.

$$\frac{d}{dt}\left(\frac{\partial (L + \frac{df}{dt})}{\partial \dot{q}_j}\right) - \frac{\partial (L + \frac{df}{dt})}{\partial q_j} = 0$$

To move on, however, we should know the explicit of $\frac{df(q,t)}{dt}$, that is,

$$\frac{df(q, t)}{dt} = \sum_i \frac{\partial f(q, t)}{\partial q_i} \dot{q}_i + \frac{\partial f(q, t)}{\partial t}$$

And we shall see the first term in equation of motion would become

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)+\frac{d}{dt}\left(\frac{\partial f(q,t)}{\partial q_j}\right)$$

and the second term would become

$$\frac{\partial L}{\partial q_j}+\frac{\partial}{\partial q_j}\left(\frac{df(q,t)}{dt}\right)$$

Because the total derivtive and partial derivative can change the order, so the second terms in above will cancel out. So the equation of motion would remain same.

# 4 Legendre Transformation & Hamilton Formulation

When we try to describe the equations of motion by **generalized coordinates** $q$ and **generalized velocities** $\dot{q}$, these $2n$ independent parameters and time $t$ form a special space called **configuation space**. On the other hand, we can also use the other parameter to describe the motion, in Hamilton formulation, we use **generalize coordinates** $q$ and **canonical momentum** $p$ to describe the space called **phase space**.
The relation between these two spaces is that phase space is the cotangent bundle of the configuation space. In the following context, we shall see this is all about math thing, so watch out! First, cotangent bundle is related to a special and important term in mathmetics and physics, **manifold**. And we will go back to this topic later.
To transform the configuation space to the phase space, we should introduce another useful tool called **Legendre transformation**. For instance, if there exists a function $f(x,y)$ that depends on x and y. We can write out its derivative.

$$df = udx \; + \; vdy$$

Where $u=\frac{\partial f}{\partial x}$ and $v=\frac{\partial f}{\partial y}$ and if we want to change $(x,y)$ to $(u,y)$, then we can construct another fucntion $g$

$$g = ux - f$$

Then its differential form would become

$$dg = udx + xdu - df = udx + xdu - udx - vdy = xdu - vdy$$

Therefore, we can see from the differential form that $g$ is a function of $y$ and $u$. And there is following relation

$$x = \frac{\partial g}{\partial u}$$

and

$$v = -\frac{\partial g}{\partial y}$$

By this means, we cah construct a quantity that is a function of $q, p, t$. And the **generalized momentum** is defined as below

$$p_j = \frac{\partial L}{\partial \dot{q}_j}$$

Notice that the above priori defining relationship *form no part of Hamilton Formulation*, momenta is independent varialbe in Hamilton formulation. Then consider the differential form of Lagrangian $L(q, \dot{q}, t)$

$$dL = \sum_i \frac{\partial L}{\partial q_i} dq_i + \sum_i \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial L}{\partial t} dt$$

By the definition of the generalized momentum, we could rewrite the above equation.

$$dL = \sum_i \frac{\partial L}{\partial q_i} dq_i + \sum_i p_i d\dot{q}_i + \frac{\partial L}{\partial t} dt$$

And we do the same thing as above procedure.

$$H(q, p, t) = \dot{q}p - L$$

called **Hamilton**
and its differential form would become

$$dH = \dot{q}dp + pd\dot{q} - dL = \dot{q}dp - \frac{\partial L}{\partial q} dq - \frac{\partial L}{\partial t} dt$$

By Lagrange equation, we can substitute the second coefficient.

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0$$

Therefore, the differential form of the Hamilton would become

$$\begin{aligned} dH &= \dot{q}dp - \dot{p}dq - \frac{\partial L}{\partial t} dt \\ &= \frac{\partial H}{\partial p} dp + \frac{\partial H}{\partial q} dq + \frac{\partial H}{\partial t} dt \end{aligned}$$

Compare each term and then we'd get **Hamilton equation of motion**

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

Thus, there would be $2n + 1$ **first order** differential equations.

## 5 Remark on Lagrangian Formulation and Hamilton Formulation

In the above discussion, however, we need to bear in mind that there is a constraint for both formulatoins, that is,

$$\frac{\partial^2 L}{\partial \dot{q}_i^2} \neq 0$$

Here are some reasons. In Lagrange formulation, if we explictly write out the time derivative term in Lagrange's equation,

$$\frac{\partial L}{\partial q^a} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^a}$$

$$= \left(\frac{\partial}{\partial q^b}\dot{q}^b + \frac{\partial}{\partial \dot{q}^b}\ddot{q}^{\,b} + \frac{\partial}{\partial t}\right)\frac{\partial L}{\partial \dot{q}^a}$$

Here we use **Einstein notation convention**. Then,

$$\ddot{q}^{\,b}\frac{\partial^2 L}{\partial \dot{q}^b \partial \dot{q}^a} = \frac{\partial L}{\partial q^a} - \frac{\partial^2 L}{\partial t \partial \dot{q}^a} - \frac{\partial^2 L}{\partial q^b \partial \dot{q}^a}\dot{q}^b$$

In the above equation, $\frac{\partial^2 L}{\partial \dot{q}^b \partial \dot{q}^a}$ is a matrix. $\ddot{q}^{\,b}$ need to be unique, and in this case, the uniqueness of $\ddot{q}^{\,b}$ requires that the determinant of the matrix shouldn't be zero, that is,

$$det\left(\frac{\partial^2 L}{\partial \dot{q}^b \partial \dot{q}^a}\right) \neq 0$$

Simply put, $\frac{\partial^2 L}{\partial \dot{q}_i^2} \neq 0$.

On the other hand, if we view this with less strick perspective, it merely told us that the mass of the system shouldn't be $0$.

In Hamilton formulation, this condition must be satisfied so that Legendre transformation is **bijective**. In other words, we have to make sure the result of inverse Legendre transformation is well-behaved. We have the transformation between Hamiltonian and Lagrangian be like,

$$H(q^a, p^a, t) = \dot{q}^a p^a - L(q^a, \dot{q}^a, t)$$

And recall that the canonical momentum is defined as,

$$p^a = \frac{\partial L}{\partial \dot{q}^a} = \nabla_{\dot{q}} L$$

And we must ensure that the canonical momenta is uniquely determined by above definition. With the formal definition of Legendre transformation,

$$H(q^a, p^a, t) = max_{\dot{q}} \left( \dot{q}^a p^a - L(q^a, \dot{q}^a, t) \right)$$

That is, we must have,

$$\nabla_{\dot{q}} \left( \dot{q}^a p^a - L(q^a, p^a, t) \right) = 0$$

Which we will get the same equation as the definition of canonical momenta. However, because of the uniqueness of canonical momenta, we must ensure that

$$\nabla_{\dot{q}}^2 L > 0$$

in other words, this means there is **only one maximum point that ensures the uniqueness of canonical momentum**.

# 6 Explict Form of Hamiltonian

First, let's consider applying transformation equation $\mathbf{r} = \mathbf{r}(q_1, ..., q_n; t)$ into kinetic energy $T$,

$$T = \sum_i \frac{1}{2} m_i v_i^2 = \sum_i \frac{1}{2} m_i \dot{r}_i^2$$

$$= \sum_i \frac{1}{2} m_i \left( \sum_j \frac{\partial \mathbf{r_i}}{\partial q_j} d\dot{q}_i + \frac{\partial \mathbf{r_i}}{\partial t} \right)^2$$

Expanding the whole thing out, we'll get,

$$T = M_0 + \sum_j M_j \dot{q}_j + \frac{1}{2} \sum_{j,\, k} M_{jk} \dot{q}_j \dot{q}_k$$

Where $M_0, M_j, M_{jk}$ are definite function of the $\mathbf{r}$'s and $t$ and hence of the $q$'s and $t$, that is,

$$M_0 = \sum_i \frac{1}{2} m_i \left( \frac{\partial \mathbf{r}_i}{\partial t} \right)^2$$

$$M_j = \sum_i m_i \frac{\partial \mathbf{r}_i}{\partial t} \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}$$

$$M_{jk} = \sum_i m_i \frac{\partial \mathbf{r}_i}{\partial q_j} \cdot \frac{\partial \mathbf{r}_i}{\partial q_k}$$

Thus, the kinetic enregy can always be writted as the sum of three homogeneous functions of the generalized velocities, $T = T_0 + T_1 + T_2$, where $T_0$ is independent of generalized velocities, $T_1$ is linear in the velocities, and $T_2$ is quadratic in the velocities.
Simalarly, we can write the Lagrangian into,

$$L(q, \dot{q}, t) = L_0(q_i, t) + L_1(q_i, t)\dot{q}_j + L_2(q_i, t)\dot{q}_j\dot{q}_k$$

Then Hamiltonian will become,

$$H = \sum_i \dot{q}_i p_i - \left[L_0(q_i, t) + \sum_j L_1(q_i, t)\dot{q}_j + \sum_{j,k} L_2(q_i, t)\dot{q}_j\dot{q}_k\right]$$

Following conditions would lead to a powerful conclusion,

- The transformation between coordinates aren't the explict functions of time. Then the kinectic energy becomes $T = \frac{1}{2}\sum_{j,k} M_{jk}\dot{q}_j\dot{q}_k$.
- The potential $V$ is generalized-velocities-independent potential.

When these two conditions, we can conclude that,

$$\sum_{j,k} L_2\dot{q}_j\dot{q}_k = T$$

$$\sum_j L_1\dot{q}_j = 0$$

$$L_0 = -V$$

And then we consider $\sum \dot{q}p = \sum \dot{q}\frac{\partial L}{\partial \dot{q}}$, then,

$$\sum_j \dot{q}_j p_j = \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} = 0 + \sum_j L_1\dot{q}_j + 2\sum_{j,k} L_2\dot{q}_j\dot{q}_k$$

Then,

$$H = 2T + V - T = V + T = E$$

Therefore, we know under these two conditions, we have the Hamiltonian is the total energy of the system.

Next, we can consider another special case that may be confused with the above discussion about total energy of a system. When Lagrangian isn't explict function of time, then automatically Hamiltonian isn't the explict function of time, too. Below gives the reason,

$$\frac{dH}{dt} = \sum_i \frac{\partial H}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial H}{\partial p_i}\dot{p}_i + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}$$

$$= -\frac{\partial L}{\partial t} = 0$$

The first two terms will cancel because of Hamilton's equation. Then, we can conclude that if Lagrangian isn't explict function of time, then **Hamiltonian is a constant of time**.

Second, we should continuely derive the explict form of Hamiltonian. First, we should change the look of Lagrangian,

$$L(q_i, \dot{q}_i, t) = L_0(q_i, t) + \dot{q}_i\ a_i(q_i, t) + \frac{1}{2}\dot{q}_i\dot{q}_j\ T_{ij}(q_i, t)$$

Where the subscript $\dot{q}_i$ that aren't in the function means summing up, that is,

$$\dot{q}_i\ a_i(q_i, t) = \sum_j \dot{q}_j\ a_j(q_i, t)$$

For more, please refer to **Einstein notation convention**.

If we write Lagrangian in symplectic form, as below,

$$L = L_0 + \dot{\mathbf{q}}^T\mathbf{a} + \frac{1}{2}\dot{\mathbf{q}}^T\mathbf{T}\dot{\mathbf{q}}$$

Where the single row matrix $\dot{\mathbf{q}}^T$ is the transpose of a single column matrix $\dot{\mathbf{q}}$, and $\mathbf{a}$ is a column matrix, and $\mathbf{T}$ is square $n \times n$ matrix. Simalarly, the Hamiltonian can be written as

$$H = \dot{\mathbf{q}}^T \mathbf{p} - L = \dot{\mathbf{q}}^T (\mathbf{p} - \mathbf{a}) - \frac{1}{2} \dot{\mathbf{q}}^T \mathbf{T} \dot{\mathbf{q}}$$

By the definition of the transverse momentum, we have,

$$\mathbf{p} = \mathbf{T}\dot{\mathbf{q}} + \mathbf{a}$$

then we have,

$$\dot{\mathbf{q}} = \mathbf{T}^{-1}(\mathbf{p} - \mathbf{a})$$

$$\dot{\mathbf{q}}^T = (\mathbf{p}^T - \mathbf{a}^T)\mathbf{T}^{-1}$$

As we know, in Hamilton formulation, Hamiltonian should be described by generalized coordinates, canonical momentum, and time,

$$H = \frac{1}{2}(\mathbf{p}^T - \mathbf{a}^T)\mathbf{T}^{-1}(\mathbf{p} - \mathbf{a}) - L_0(q_i, t)$$

# 7 Symplectic Form of Hamilton's Equation of Motion

Again, we try to write out the symplectic form of Hamilton's equation of motion,

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

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

If we assign $q_i$ and $p_i$ altogether into a column vector $\boldsymbol{\eta}$ with $2n$ elements, and,

$$\boldsymbol{\eta} = \begin{bmatrix} q_1 \\ \cdots \\ q_n \\ p_1 \\ \cdots \\ p_n \end{bmatrix}$$

In other words, $\eta_1$ to $\eta_n$ are generalized coordinates $q_i$, and $\eta_{n+1}$ to $\eta_{2n}$ are the canonical momentum $p_i$. And then Hamilton's equation of motion would become,

$$\dot{\boldsymbol{\eta}} = \boldsymbol{J}\frac{\partial H}{\partial \boldsymbol{\eta}}$$

Where the $2n \times 2n$ square matrix $\boldsymbol{J}$ is a special matrix,

$$\boldsymbol{J} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{-1} & \mathbf{0} \end{bmatrix}$$

Take 2 variables for instance, we have,

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial H}{\partial q_1} \\ \frac{\partial H}{\partial q_2} \\ \frac{\partial H}{\partial p_1} \\ \frac{\partial H}{\partial p_2} \end{bmatrix}$$

By Hamilton's equation of motion,

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\dot{p}_1 \\ -\dot{p}_2 \\ q_1 \\ q_2 \end{bmatrix}$$

$$\dot{\boldsymbol{\eta}} = \boldsymbol{J}\frac{\partial H}{\partial \boldsymbol{\eta}}$$

There are some properties of symplectic matrix,

$$\boldsymbol{J}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{-1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}$$

$$\boldsymbol{J}^{-1} = \boldsymbol{J}^T$$

$$\det\left(\boldsymbol{J}^{-1}\boldsymbol{J}\right) = \det\left(\begin{bmatrix} \mathbf{0} & \mathbf{-1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{-1} & \mathbf{0} \end{bmatrix}\right) = \det\left(\begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}\right) = 1$$

$$\boldsymbol{J}\boldsymbol{J} = \mathbf{1}$$

# 8 Routhian

Then we should talk about the **cyclic coordinate** in Hamilton foromulation. As we already knew, if a coordinate is cyclic, say is $q_n$, the Lagrangian can be written as

$$L(q_1, q_2, ...q_{n-1}; \dot{q}_i; t)$$

Note that all generalized velocities still occur in Lagrangian, including $\dot{q_n}$. On the other hand, in Hamilton formulation, if we have $q_n$ be cyclic, then by the Hamilton's equation of motion,

$$\dot{p_n} = -\frac{\partial H}{\partial q_n} = 0$$

That is, $p_n$ becomes the constant of time, say is $\alpha$. Then the form of Hamiltonian will be,

$$H(q_1, ..., q_{n-1};\ p_1, ..., p_{n-1};\ \alpha;\ t)$$

The behavior of the cyclic coordinate itself with time is then found by integrating the equation of motion,

$$\dot{q_n} = \frac{\partial H}{\partial \alpha}$$

By the properties shown by the Hamiltonian with cyclic coordinates, we can use it with Lagrangian with non-cyclic coordinates, and this way is designed by Routh. Precisely, we have the Routhian $R(q; p; \xi; \dot{\xi}; t)$, and $\xi$ denote the non-cyclic generalized coordinates, and $q$ denote the cyclic coordinates. If we define the Routhian to be

$$R = \sum \dot{q}p - L(q; \xi; \dot{q}; \dot{\xi}; t)$$

Which can be rewritten as

$$R(q; p; \xi; \dot{\xi}; t) = \sum \dot{q}p - L_{cyclic}(q; \dot{q}; t) - L_{non-cyclic}(\xi; \dot{\xi}; t)$$

$$= H_{cyclic}(q; p; t) - L_{non-cyclic}(\xi; \dot{\xi}; t)$$

For non-cyclic coordinates, we have the Lagrange's equation,

$$\frac{d}{dt}\left(\frac{\partial R}{\partial \dot{\xi}}\right) - \frac{\partial R}{\partial \xi} = 0$$

As for the cyclic coordinates, we have the Hamilton's equation,

$$\dot{p} = -\frac{\partial R}{\partial q} = 0$$

$$\dot{q} = \frac{\partial R}{\partial p}$$

In conclusion, Routhian may help solving some physical and engineering problems, and it combines some of the features of both the Lagrangian and the Hamiltonian pictures. However, it adds no new thing to the physics.

A formal way to deriving Routhian is as shown below. If we try to change the coordinates from $(q, \xi, \dot{q}, \dot{\xi}, t)$ to $(q, \xi, p, \dot{\xi}, t)$, first we consider the differential form of the Lagrangian $L(q, \xi, \dot{q}, \dot{\xi}, t)$,

$$dL = \sum \frac{\partial L}{\partial q} dq + \sum \frac{\partial L}{\partial \dot{q}} d\dot{q} + \sum \frac{\partial L}{\partial \xi} d\xi + \sum \frac{\partial L}{\partial \dot{\xi}} d\dot{\xi} + \frac{\partial L}{\partial t} dt$$

Then we consider,

$$R = \sum \dot{q} p - L$$

and its differential form would be,

$$dR = \sum p d\dot{q} + \sum \dot{q} dp - \left( \sum \frac{\partial L}{\partial q} dq + \sum \frac{\partial L}{\partial \dot{q}} d\dot{q} + \sum \frac{\partial L}{\partial \xi} d\xi + \sum \frac{\partial L}{\partial \dot{\xi}} d\dot{\xi} + \frac{\partial L}{\partial t} dt \right)$$

$$dR = \sum \dot{q} dp - \sum \dot{p} dq - \sum \frac{\partial L}{\partial \xi} d\xi - \sum \frac{\partial L}{\partial \dot{\xi}} d\dot{\xi} - \frac{\partial L}{\partial t} dt$$

which shows that,

$$\frac{\partial R}{\partial p} = \dot{q}$$

$$\frac{\partial R}{\partial q} = -\dot{p}$$

$$\frac{\partial R}{\partial \xi} = -\frac{\partial L}{\partial \xi}$$

$$\frac{\partial R}{\partial \dot{\xi}} = -\frac{\partial L}{\partial \dot{\xi}}$$

The first two equations imply that Routhian conforms to Hamilton's equation with generalized coordinates $q$, on the other hand, the last two equations imply Routhian conforms to Lagrange's equation with generalized coordinates $\xi$. If the generalized coordinates $q$ are cyclic, and $\xi$ isn't, then the dicussion would be like above.

# 9 Hamilton's Equations From Variation Principle

In this section, we're going to derive Hamilton's equation of motion from variation principle. The action can be rewritten as,

$$\delta S = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt = \int_{t_1}^{t_2} \sum \dot{q}p - H(q, p, t)\, dt$$

When we consider more generally, we can view the integrand as a function of generalized coordinates and canonical momentum,

$$\delta S = \int_{t_1}^{t_2} f(q, \dot{q}, p, \dot{p}, t)\, dt$$

Roughly speaking, we'll get $2n$ Euler-Lagrange's equation,

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{q}_i}\right) - \frac{\partial f}{\partial q_i} = 0$$

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{p}_i}\right) - \frac{\partial f}{\partial p_i} = 0$$

In our case, there isn't $\dot{p}$ in $f$, and the above equations would become,

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

$$0 - \left(\dot{q}_i - \frac{\partial H}{\partial p}\right) = 0$$

Which are exacly Hamilton's equation of motion.

# 10 Canonical Transformation-Generating Function

The equal status accorded to coordinates and momenta as independent variables encourages a greater freedom in selecting the pyhsical quantities to be designated as "coordinates" and "momenta". In this section, we'll focus on how to transform from one set of varialbes to another set of variables. First, we shall introduce the most trivial one, *point transformation*,
Let $q_1, ..., q_n$ be a set of independent generalized coordinates for a system of $n$ degrees of freedom, with a Lagrangian $L(q, \dot{q}, t)$. Suppose we transform to another set of independent coordinates $s_1, ..., s_n$ by means of transformation equations

$$q_i = q_i(s_1, ..., s_n, t), \ \ i = 1, ..., n$$

Show that the form of Lagrange's equations is invariant under a point transformation.

**Proof** Our goal is to get,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{s}_i}\right) - \frac{\partial L}{\partial s_i} = 0$$

By chain rule, we have,

$$\frac{\partial L}{\partial s_i} = \frac{\partial L}{\partial q_i}\frac{\partial q_i}{\partial s_i}$$

and the expansion of $\dot{q}_i(s_i, t)$

$$\dot{q}_i = \sum_j \frac{\partial q_i}{\partial s_j}\dot{s}_j + \frac{\partial q_i}{\partial t}$$

$$\frac{\partial L}{\partial \dot{s}_i} = \frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{q}_i}{\partial \dot{s}_i} = \frac{\partial L}{\partial \dot{q}_i}\frac{\partial q_i}{\partial s_i}$$

And then if the original Lagrange's equation of motion times $\frac{\partial q_i}{\partial s_i}$, we will get our goal.
From the above example, we can see what we expect from the transformation is that it has to maintain the equation of motion. In Hamilton formulation, the momenta are also independent variables on the same level as the generalized coordinates, so the concept of transformation must be widened. The simultaneous transformation from original coordinates and momenta $q_i, p_i$ to the new ones $Q_i, P_i$ with the following **invertible transformation equations**,

$$Q_i = Q_i(q, p, t)$$

$$P_i = P_i(q, p, t)$$

Note that not all the transformations are of interested. New $Q, P$ must be canonical coordinates, which requires that there is function $K(Q, P, t)$ such that Hamilton's equations of motion in the new set are satisfied.

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

Function $K$ plays the role of Hamiltonian in the new set, called *transformed Hamiltonian*. Note that the transformation considered be **problem-independent**, in other words, the new coordinates set $(Q, P)$ must be canonical coordinates for all systems with the same degrees of freedom.
How to define canonical coordinates? They must follow Hamilton's principle, that is,

$$\delta \int_{t_1}^{t_2} \left( P_i \dot{Q}_i - K(Q_i, P_i, t) \right) \, dt = 0$$

On the other hand, the old one should satisfy,

$$\delta \int_{t_1}^{t_2} \left( p_i \dot{q}_i - H(q_i, p_i, t) \right) \, dt = 0$$

The above equations don't inply that their integrand should be the same, however, they do follow the following relation,

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

As we have mention in previous section, the last term $\frac{dF}{dt}$ is the trivial term [5] which doesn't affect the equation of motion. One must note that the funciton $F$ would be the funciton of $q_i, p_i$ or $P_i, Q_i$, but not the function of $\dot{q}_i, \dot{p}_i, \dot{Q}_i$ and so on, as we have prove before. And $\lambda$ is the constant independent of time and the canonical coordinates.

In convention, we often assume that $\lambda = 1$, and we must make sure we can always do this assumption. Therefore, we need to consider a transformation related to $\lambda$, *scale transformation*. Consider the following transformation,

[5] In fact, it plays a important role as we'll mention later

$$Q_i = \mu q_i, \quad P_i = \nu p_i$$

Take into Hamilton's equations of motion in the new set are satisfied,

$$\dot{P}_i = \frac{\partial K}{\partial Q_i}, \;\; \dot{Q}_i = -\frac{\partial K}{\partial P_i}$$
$$\nu p_i = \frac{1}{\mu}\frac{\partial K}{\partial q_i}, \;\; \mu \dot{q}_i = -\frac{1}{\nu}\frac{\partial K}{\partial p_i}$$

We may find out that $K = \mu\nu H$, and then,

$$\mu\nu \left[p_i\dot{q}_i - H(q_i, p_i, t)\right] = P_i\dot{Q}_i - K(Q_i, P_i, t)$$

Hence, $\lambda = \mu\nu$, and we can always find a possible *scale transformation* so that we should put emphases on,

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

A transformation of canonical coordinates with $\lambda \neq 1$ is called *extended canonical transformation*. On the other hand, we merely called the transformation with $\lambda = 1$ *canonical transformation*. If the transformation,

$$Q = Q(q, p), \;\; P = P(q, p)$$

don't contain time $t$ explictly, then we called it *restricted canonical transformation*.

Next, we should focus on the last term function $F$, in fact, if half of the parameters of $F$ come from old canonical variables and the other come from the new one, then $F$ would determine the canonical transformation between $(q, p)$ and $(Q, P)$, called **generating function** of the transformation. In the following context, we will show how the generating function to determine the transformation.

In the beginning, consider $F$ is the function of old coordinates and new coordinates,

$$F = F_1(q, Q, t)$$

If we take it into Hamilton equation of transformation, then

$$\begin{aligned} dF_1 &= p_i dq_i - H dt - P_i dQ_i + K dt \\ &= \frac{\partial F_1}{\partial q_i} dq_i + \frac{\partial F}{\partial Q_i} dQ_i + \frac{\partial F}{\partial t} dt \end{aligned}$$

Compare the coefficients, we would get the following relation,

$$p_i = \frac{\partial F_1}{\partial q_i}$$

$$P_i = -\frac{\partial F_1}{\partial Q_i}$$

$$K - H = \frac{\partial F_1}{\partial t}$$

From the first equation above, we can express $n$ old canonical momenta $p_i$ in terms of $q_i$ , $Q_i$ , and $t$. Inversely, we can express new canonical coordinates $Q_i$ in terms of $q_i$ , $p_i$ , and $t$. Then, similarly, from the second equation above, we can express new canonical momenta in terms of $q_i$ , $Q_i$ , and $t$, $P_i = P_i(q, Q, t)$, and by virtue of transformation relation discussed before $Q_i = Q_i(q, p, t)$, we then get the remaining relation $P_i = P_i(q, p, t)$. I think this is just math and provide no new information for physics, since the generating function wouldn't affect the equation of motion anyway.

However, it still worthes going through some basic generating functions of canonical transformation. Then, consider, $F = F_2(q, P, t) - Q_i P_i$, which is a little bit different from the first kind. Yet, let's go through it as above,

$$dF = p_i dq_i - H dt - P_i dQ_i + K dt$$

$$= \frac{\partial F_2}{\partial q_i} dq_i + \frac{\partial F_2}{\partial P_i} dP_i + \frac{\partial F_2}{\partial t} dt - P_i dQ_i - Q_i dP_i$$

Same, by comparing the coefficients, we'd get,

$$p_i = \frac{\partial F_2}{\partial q_i}$$

$$Q_i = \frac{\partial F_2}{\partial P_i}$$

$$K - H = \frac{\partial F_2}{\partial t}$$

Third, if $F = F_3(p, Q, t) + q_i p_i$, by the same way, we'll get,

$$dF = p_i dq_i - Hdt - P_i dQ_i + Kdt$$

$$= \frac{\partial F_3}{\partial p_i} dp_i + \frac{\partial F_3}{\partial Q_i} dQ_i + \frac{\partial F_3}{\partial t} dt + p_i dq_i + q_i dp_i$$

Then we'll get,

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

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

$$K - H = \frac{\partial F_3}{\partial t}$$

Finally, we can move on to the last one, that is, $F = F_4(p, P, t) + q_i p_i - Q_i P_i$, then,

$$dF = p_i dq_i - Hdt - P_i dQ_i + Kdt$$

$$= \frac{\partial F_4}{\partial p_i} dp_i + \frac{\partial F_4}{\partial P_i} dP_i + \frac{\partial F_4}{\partial t} + p_i dq_i + q_i dp_i - P_i dQ_i - Q_i dP_i$$

Then we shall get,

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

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

$$K - H = \frac{\partial F_4}{\partial t}$$

Sometimes the generating function cannoot be expressed as above four equations, and it may be the mixture of them. Next, we're going to do some examples.

**Example 1:** if the generating function has the second type form,

$$F_2(q, P, t) = f_i(q, t)P_i + g(q, t)$$

that is, $F = F_2 - Q_iP_i$, thus,

$$p_i = P_i\frac{\partial f_i}{\partial q_i} + \frac{\partial g}{\partial q_i}$$

$$Q_i = f_i(q, t)$$

From the form of $Q_i$, we can see that it is the canonical transformation we have mentioned before, that is, *point transformation*.

**Example 2:** if the generating function isn't an explict function of time, then we have,

$$F = F(q, p, Q, P)$$

and its differential form would become,

$$dF = p_i dq_i - H dt - P_i dQ_i + K dt$$

$$= \frac{\partial F}{\partial q_i} dq_i + \frac{\partial F}{\partial p_i} dp_i + \frac{\partial F}{\partial Q_i} dQ_i + \frac{\partial F}{\partial P_i} dP_i$$

We can see that the new Hamiltonian and the old one are the same, $K = H$. Furthermore, we have the trasformation between canonical variables like,

$$Q_i = Q_i(q, p)$$

$$P_i = P_i(q, p)$$

It is *restricted canonical transformation*, and we can conclude that Hamiltonian of *restricted canonical transformatoin* is the same.

# 11 Symplectic Approach of Canonical Transformation

In the begining, we focus on *restricted canonical transformation*. One properties of restricted canonical transformation is that the new Hamiltonian $K$ is equal to the old one $H$. We can prove it by what we've just learnt, that is, it is the example of last section, and we use it directly. For restricted canonical transformation, we have,

$$\begin{aligned} Q_i &= Q_i(q, p) \\ P_i &= P_i(q, p) \end{aligned} \tag{11.1}$$

In matrix format,

$$\boldsymbol{\xi} = \boldsymbol{\xi}(\boldsymbol{\eta}) \tag{11.2}$$

Recall the matrix format of Hamilton's equations of motion.

$$\dot{\boldsymbol{\eta}} = \boldsymbol{J} \frac{\partial H}{\partial \boldsymbol{\eta}} \tag{11.3}$$

Then our goal is to find a similar form for new canonical variables. Consider the time derivative of them,

$$\dot{\xi}_j = \sum_i \frac{\partial \xi_j}{\partial \eta_i} \frac{d\eta_i}{dt}$$

If we try to write it into matrix form, we'll get,

$$\dot{\boldsymbol{\xi}} = \boldsymbol{M}\dot{\boldsymbol{\eta}} \tag{11.4}$$

Where transformation matrix $\boldsymbol{M}$ is,

$$M^i{}_j = \frac{\partial \xi_i}{\partial \eta_j}$$

$$\begin{bmatrix} \dot{\xi}_1 \\ \vdots \\ \dot{\xi}_j \\ \vdots \\ \dot{\xi}_n \end{bmatrix} = \begin{bmatrix} \frac{\partial \xi_1}{\partial \eta_1} & \cdots & \frac{\partial \xi_1}{\partial \eta_i} & \cdots & \frac{\partial \xi_1}{\partial \eta_n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{\partial \xi_j}{\partial \eta_1} & \cdots & \frac{\partial \xi_j}{\partial \eta_i} & \cdots & \frac{\partial \xi_j}{\partial \eta_n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{\partial \xi_n}{\partial \eta_1} & \cdots & \frac{\partial \xi_n}{\partial \eta_i} & \cdots & \frac{\partial \xi_n}{\partial \eta_n} \end{bmatrix} \begin{bmatrix} \dot{\eta}_1 \\ \vdots \\ \dot{\eta}_i \\ \vdots \\ \dot{\eta}_n \end{bmatrix}$$

Then the most important part comes, if (11.2) can be inversed, that is,

$$\boldsymbol{\eta} = \boldsymbol{\eta}(\boldsymbol{\xi}) \tag{11.5}$$

We can claim that the old Hamiltonian $H(q, p, t)$ can be expressed by new canonical variables,

$$H(\boldsymbol{\eta}, t) \to H(\boldsymbol{\xi}, t)$$

Then in equation 11.3, we have,

$$\frac{\partial H}{\partial \eta_k} = \sum_l \frac{\partial H}{\partial \xi_l} \frac{\partial \xi_l}{\partial \eta_k}$$

In matrix form,

$$
\begin{bmatrix} \frac{\partial H}{\partial \eta_1} \\ \vdots \\ \frac{\partial H}{\partial \eta_k} \\ \vdots \\ \frac{\partial H}{\partial \eta_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial \xi_1}{\partial \eta_1} & \cdots & \frac{\partial \xi_l}{\partial \eta_1} & \cdots & \frac{\partial \xi_n}{\partial \eta_1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{\partial \xi_1}{\partial \eta_k} & \cdots & \frac{\partial \xi_l}{\partial \eta_k} & \cdots & \frac{\partial \xi_n}{\partial \eta_k} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{\partial \xi_1}{\partial \eta_n} & \cdots & \frac{\partial \xi_l}{\partial \eta_n} & \cdots & \frac{\partial \xi_n}{\partial \eta_n} \end{bmatrix} \begin{bmatrix} \frac{\partial H}{\partial \xi_1} \\ \vdots \\ \frac{\partial H}{\partial \xi_l} \\ \vdots \\ \frac{\partial H}{\partial \xi_n} \end{bmatrix}
$$

As we can see, it can be written in,

$$
\dot{\boldsymbol{\eta}} = \boldsymbol{J}\boldsymbol{M}^T \frac{\partial H}{\partial \boldsymbol{\xi}} \tag{11.6}
$$

with equation 11.4, we have,

$$
\dot{\boldsymbol{\xi}} = \boldsymbol{M}\ \boldsymbol{J}\boldsymbol{M}^T \frac{\partial H}{\partial \boldsymbol{\xi}}
$$

By Hamilton's equations of motion for the new varialbes $\dot{\boldsymbol{\xi}} = \boldsymbol{J}\frac{\partial H}{\partial \boldsymbol{\xi}}$, we have,

$$
\boldsymbol{M}\boldsymbol{J}\boldsymbol{M}^T \frac{\partial H}{\partial \boldsymbol{\xi}} = \boldsymbol{J}\frac{\partial H}{\partial \boldsymbol{\xi}} \tag{11.7}
$$

Therefore, we should have so-called **symplectic condition**

$$
\boldsymbol{M}\boldsymbol{J}\boldsymbol{M}^T = \boldsymbol{J} \tag{11.8}
$$

And then, we can derive,

$$
\begin{gathered}
\boldsymbol{J}\boldsymbol{M}^T = \boldsymbol{M}^{-1}\boldsymbol{J} \\
\text{times } \boldsymbol{J} \text{ from left} \\
\boldsymbol{J}\boldsymbol{J}\boldsymbol{M}^T = \boldsymbol{J}\boldsymbol{M}^{-1}\boldsymbol{J} \\
\text{and then times } \boldsymbol{J} \text{ from right} \\
\boldsymbol{J}\boldsymbol{J}\boldsymbol{M}^T\boldsymbol{J} = \boldsymbol{J}\boldsymbol{M}^{-1}\boldsymbol{J}\boldsymbol{J} \\
\boldsymbol{M}^T\boldsymbol{J} = \boldsymbol{J}\boldsymbol{M}^{-1}
\end{gathered}
$$

Then we'll have,

$$
\boldsymbol{M}^T\boldsymbol{J}\boldsymbol{M} = \boldsymbol{J} \tag{11.9}
$$

which is another form of **symplectic condition**.

## 12 Restricted Canonical Transformations Forms a Group $\mathcal{G}$

We need to notice that formula, $\mathbf{M}^T\mathbf{J}\mathbf{M} = \mathbf{J}$. For starters, we can prove the **closure** of canonical transformation. Consider $\mathbf{M}_1,\ \mathbf{M}_2 \in \mathcal{G}$ ,

$$(\mathbf{M}_1\mathbf{M}_2)^T\,\mathbf{J}\,(\mathbf{M}_1\mathbf{M}_2) = \mathbf{M}_2^T\mathbf{M}_1^T\mathbf{J}\mathbf{M}_1\mathbf{M}_2 = \mathbf{M}_2^T\mathbf{J}\mathbf{M}_2 = \mathbf{J} \tag{12.1}$$

which remains in group $\mathcal{G}$.

Then we can prove the **accosiativity**. Because *restricted canonical transformations can be expressed in matrix form* and that the product of matrice has accosiativity, we can claim that the accosiativity holds.

Then we prove the **indentity**, and consider the identity matrix $\mathbf{I}$,

$$\mathbf{I}^T\mathbf{J}\mathbf{I} = \begin{bmatrix}\mathbf{1} & \mathbf{0}\\ \mathbf{0} & \mathbf{1}\end{bmatrix}\begin{bmatrix}\mathbf{0} & \mathbf{1}\\ -\mathbf{1} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{1} & \mathbf{0}\\ \mathbf{0} & \mathbf{1}\end{bmatrix} = \mathbf{J}$$

Then we can see that $\mathbf{I} \in \mathcal{G}$ and it plays the role of **identity** in canonical transformation[6].

Finally, we have to prove **inverse** exists for every member in group, that is, we need to prove that there is **inverse matrix**, and this is simple. What we have to do is to check the determinant of transformation matrix is not zero.

$$\det\left(\mathbf{M}^T\mathbf{J}\mathbf{M}\right) = \det(\mathbf{M})^2\det(\mathbf{J}) = \det(\mathbf{J}) \tag{12.2}$$

which implies $\det(\mathbf{M}) = \pm 1$.

## 13

[6]this kind of transformation is called *identical transformation*.