PHY354: Advanced Classical Mechanics

QiLin Xue

Fall 2021

Contents

1   Motivation

In previous courses, we focused on $\vec{F} = m\vec{a}$ and its consequences. In this course, we will do it in a more elegant way, i.e. via Euler, Lagrange, Hamilton, Jacobi, Noether. Some disadvantages of Newton's approach are:

- Can be difficult to apply to complex situations and extended objects
- Difficult to see how chaos theory arises
- Obscures relationship between quantum and classical mechanics

2   Lagrangian Mechanics

We first start with a few examples showing Lagrangian Mechanics in action, and we will later discuss generalizations and intricacies:

Example 1: Let us consider a simple system with a mass m on a spring with spring constant k. Instead of solving this by writing $F = -kx$, let us write the following strange combination of kinetic and potential energy, known as the **Lagrangian**.:

$$L = T - V, \tag{2.1}$$

where T is the kinetic energy and V is the potential energy. Using $T = \frac{1}{2}m\dot{x}^2$ and $V = \frac{1}{2}kx^2$, we get:

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2. \tag{2.2}$$

Consider the following equation, known as the Euler-Lagrange Equation:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x} \tag{2.3}$$

which gives

$$\frac{d}{dt}(m\dot{x}) = -kx \implies m\ddot{x} = -kx. \tag{2.4}$$

Example 2: Suppose we have a general potential $V = V(x)$. Then the E-L equation gives

$$m\ddot{x} = -\frac{\partial V(x)}{\partial x}. \tag{2.5}$$

Notice that $-\dfrac{\partial V(x)}{\partial x}$ is just the force!

If we have more than one dimension, i.e. x, y, z, then we'll have 3 E-L equations, and we'll have to solve them separately. The Lagrangian would be

$$L = \frac{1}{2}m\sum_i \dot{x}_i^2 - V(x_i) \tag{2.6}$$

and the 3 different E-L equations would be

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right) = \frac{\partial L}{\partial x_i} \implies m\ddot{x}_i = -\frac{\partial V}{\partial x_i} \tag{2.7}$$

for $i = 1, 2, 3$. These equations are just the 3 components of $m\ddot{\vec{r}} = -\vec{\nabla}V(\vec{r})$. The important idea here is that we can determine the equations of motion without knowing the forces! We just need to figure out the potential energy. We can try a more complex example:

Example 3: Suppose we have a mass on a spring that acts as a pendulum. The spring has equilibrium length ℓ and at

an angle θ, the spring has length $r = \ell + x$. We will work in polar coordinates x, θ. The kinetic energy is

$$T = \frac{1}{2}m\vec{v}^2 = \frac{1}{2}m\left(\dot{x}^2 + (\ell + x)^2\dot{\theta}^2\right),\tag{2.8}$$

and the potential energy is

$$V = -mg(\ell + x)\cos\theta + \frac{1}{2}kx^2.\tag{2.9}$$

The $E - L$ equation for x gives

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}\tag{2.10}$$

$$m\ddot{x} = m(\ell + x)\dot{\theta}^2 + mg\cos\theta - kx.\tag{2.11}$$

The term $m(\ell+x)\dot{\theta}^2$ is the centripetal force, $mg\cos\theta$ is the radial gravitational force, and $-kx$ is the spring force. The θ component is

$$\frac{d}{dt}m(\ell + x)^2\dot{\theta} = mg(\ell + x)\sin\theta\tag{2.12}$$

which corresponds to

$$\vec{\tau} = \frac{d\vec{L}}{dt}\tag{2.13}$$

where $\vec{\tau}$ is torque and $\vec{L}$ is the angular momentum.

3 Principle of Stationary Action

In principle, there are an infinite number of paths a particle can take from point A to point B. Traditionally, we will solve this by analyzing the forces at every single point to find the path it actually takes. We take an alternative approach. We define the action

$$S \equiv \int_{t_0}^{t_1} L(q(t), \dot{q}(t), t)\,\mathrm{d}t\,,\tag{3.1}$$

where q is a generalized coordinate.

We assign each path a particular action, and we claim that the *particle takes the path with the least action*. However, this may be tricky since it depends on $q(t_0), q(t_0 + \epsilon), \ldots, q(t_1)$, an infinite number of variables. The terminology here is that the action S is a functional of $q(t)$, i.e. we write it as

$$S[q(t)] = \lim_{\epsilon \to 0} S(q(t_0), q(t_0 + \epsilon), \ldots, q(t_1)).\tag{3.2}$$

We want to find the extremum of this action. For functionals, we can consider a small perturbation $q_a(t) = q_c(t) + a\beta(t)$. Then we claim that an extremum occurs when

$$\frac{\partial}{\partial a}S[\beta(t)] = 0,\tag{3.3}$$

for any function $\beta(t)$. We will now offer a proof for the E-L equations.

Proof. We can write

$$\begin{aligned}
\frac{\partial}{\partial a}S[q_a(t)] &= \frac{\partial}{\partial a}\int_{t_0}^{t_1} L(q_a(t), \dot{q}_a(t), t)\,\mathrm{d}t \\
&= \int_{t_0}^{t_1} \frac{\partial L}{\partial a}(q_a(t), \dot{q}_a(t), t)\,\mathrm{d}t \\
&= \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial q_a}\frac{\partial q_a}{\partial a} + \frac{\partial L}{\partial \dot{q}_a}\frac{\partial \dot{q}_a}{\partial a}\right]\mathrm{d}t \\
&= \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial q_a}\beta(t) + \frac{\partial L}{\partial \dot{q}_a}\dot{\beta}(t)\right]\mathrm{d}t\,.
\end{aligned}$$

Note that the second term can be simplified using integration by parts

$$\int_{t_0}^{t_1} \frac{\partial L}{\partial \dot{q}_a} \dot{\beta}(t)\, dt = \left. \frac{\partial L}{\partial \dot{q}_a} \beta(t) \right|_{t_0}^{t_1} - \int_{t_0}^{t_1} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) \beta(t)\, dt\,.$$

The first term goes to zero since the end and start points do not change. Our original derivative is then

$$\frac{\partial}{\partial a} S[q_a(t)] = \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right)\right] \beta(t)\, dt\,. \tag{3.4}$$

Since this integral is zero, and β can be arbitrary, then it must be true that

$$\frac{\partial L}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) = 0, \tag{3.5}$$

which is the E-L equation. Note that this easily generalizes to multiple dimensions, and hold for any set of generalized coordinates. □

Remarks: The stationary point always refers to a local minimum or a saddle point, but *never* a maximum.

Furthermore, note that L is not unique. Here are some acceptable modifications:

- We can multiply L by a constant.

- We can add a total time derivative to the Lagrangian: $L \to L + \dfrac{df}{dt}$.

  Proof. The action will pick up a $\displaystyle\int_{t_0}^{t_1} \frac{df}{dt}\, dt$ term, which by the FTC, it adds a constant to the action, and thus the extremums will remain the same. □

3.1 Generalizations

In general, if we have some integral dependent on a path

$$I \equiv \int_{x_1}^{x_2} F(y(x), y'(x))\, dx\,, \tag{3.6}$$

then I is an extremum when

$$\frac{d}{dx}\frac{\partial f}{\partial y'} = \frac{\partial f}{\partial y} \tag{3.7}$$

Example 4: Let's try a trivial problem: What is the shortest distance between 2 points on a plane? Consider a path ℓ that connects (x_1, y_1) and (x_2, y_2). The path length is then

$$\ell = \int_P \sqrt{1 + y'^2}\, dx\,, \tag{3.8}$$

which is an extremum when

$$\frac{d}{dx}\frac{\partial \sqrt{1 + y'^2}}{\partial y'} = \frac{\partial \sqrt{1 + y'^2}}{\partial y} \tag{3.9}$$

$$\frac{d}{dx}\frac{y'}{\sqrt{1 + y'^2}} = 0, \tag{3.10}$$

which is true when $y'(x) = $ constant, which corresponds to a straight line.

3.2   Brachistochrome Problem

Consider a bead sliding down a frictionless wire, whose endpoints are (x_1, y_1) and (x_2, y_2). What shape of wire will minimize the time it takes to slide down?

We can formulate in terms of a variational calculus problem, i.e. by optimizing a "time functional" over all paths $y(x)$. The time it takes is

$$T = \int_1^2 \mathrm{d}t$$

$$= \int_1^2 \frac{\mathrm{d}s}{v} \qquad\qquad = \int_{x_1}^{x_2} \frac{\sqrt{1 + (y')^2}}{v} \, \mathrm{d}x \,,$$

or alternatively we can write it as

$$T = \int_{y_1}^{y_2} \frac{\sqrt{1 + (x')^2}}{v} \, \mathrm{d}y \,, \tag{3.11}$$

which turns out to be easier since $v = \sqrt{2gy}$. Then:

$$T = \frac{1}{\sqrt{2g}} \int_{y_1}^{y_2} \underbrace{\frac{\sqrt{1 + (x')^2}}{\sqrt{y}}}_{F(y,x')} \, \mathrm{d}y \,. \tag{3.12}$$

Using the E-L equation, we have

$$0 = -\frac{d}{dy} \frac{\partial F}{\partial x'} + \frac{\partial F}{\partial x}$$

$$0 = -\frac{d}{dy} \frac{x'}{\sqrt{y}\sqrt{1 + (x')^2}} .$$

Since the derivative is 0, that means $\dfrac{\partial F}{\partial x'} = a$ must be a constant. Rearranging, we get

$$x' = \sqrt{\frac{y}{1/a^2 - y}} = \sqrt{y/(C - y)} \tag{3.13}$$

where $C \equiv 1/a^2$. Now solving for $x(y)$, and using the substitution $y = c\sin^2\theta$, we end up as:

$$y = \frac{c}{2}(1 - \cos 2\theta) \tag{3.14}$$

$$x = \frac{c}{2}(2\theta - \sin(2\theta)) + \beta. \tag{3.15}$$

There are two boundary conditions and two unknowns. If we instead chose to do the problem by integrating the functional with $\mathrm{d}x$, then we'll end up with a very messy second order product rule expression in the E-L equations. However, we can simplify things using a simple theorem.

> **Theorem**: For $S[y] = \displaystyle\int \mathcal{L}(x, y, y') \, \mathrm{d}x$, if $\mathcal{L}(x, y, y') = \mathcal{L}(y, y')$, that is $\mathcal{L}$ is independent of x, then
>
> $$\frac{\partial L}{\partial y'} y' - \mathcal{L} = \text{constant}. \tag{3.16}$$

Proof. Differentiate the quantity

$$\frac{d}{dx}\left(\frac{\partial \mathcal{L}}{\partial y'} y' - \mathcal{L}\right) = \frac{d}{dx}\frac{\partial L}{\partial y'} y' + \frac{\partial \mathcal{L}}{\partial y'} y'' - \frac{\partial \mathcal{L}}{\partial y} y' - \frac{\partial \mathcal{L}}{\partial y'} y'' \tag{3.17}$$

$$= \left(\frac{d}{dx}\frac{\partial \mathcal{L}}{\partial y'} - \frac{\partial \mathcal{L}}{\partial y}\right) y' = 0 \tag{3.18}$$

$$\square$$

Applying this theorem, and using,

$$F(y, y') = \sqrt{\frac{1 + (y')^2}{y}},$$

we get

$$a = \frac{\partial F}{\partial y'} - F$$

$$0 = \frac{y'}{\sqrt{y}\sqrt{1 + (y')^2}} - \sqrt{\frac{1 + (y')^2}{y}}.$$

After rearranging, we get a first order differential equation that is equivalent to the

3.3   Constrained Systems

Sometimes, we have a constraint. For example in a pendulum when working with x and y coordinates, we can introduce a **holonomic constraint:**

$$x^2 + y^2 - l^2 = 0. \tag{3.19}$$

In general, holonomic constraints are in the form of $f(q_1, \ldots, q_n) = 0$.

4   Noether's Theorem

> **Theorem**: If a system has a continuous symmetry property, then there are corresponding quantities whose values are conserved in time.

A **cyclic** coordinate is one that does not explicitly appear in the Lagrangian. If q_i is cyclic, then the **generalized/conjugate momentum** is conserved, i.e.

$$\frac{d}{dt} q_i = 0.$$

In general, generalized momentum is not linear momentum. These forms of translational symmetry implies some conservation law, which is an example of Noether's Theorem.

> **Definition**: A symmetry is a transformation such that the equations of motion do not change.

More precisely, we can do a transformation $q_i(t) \to Q_i(t, \lambda)$ where $q_i(t) = Q_i(t, 0)$. For example, a translation in the x direction is

$$x(t) \to x(t, \lambda) = x(t) + \lambda$$
$$y(t) \to y(t, \lambda) = y(t)$$
$$z(t) \to z(t, \lambda) = z(t).$$

and a time translation would be $q_i(t) \to Q_i(t, \lambda) = q_i(t + \lambda)$. Then, a transformation is a symmetry if the equations of motion for Q_i are the same as q_i. More precisely, we claim that this is a symmetry if

$$\frac{\partial L}{\partial \lambda}(Q_i, \dot{Q}_i, t) = \frac{df}{dt} \tag{4.1}$$

for some function $f(Q_i, \dot{Q}_i, t)$. We can now prove Noether's Theorem now:

Proof. We evaluate $\dfrac{\partial L}{\partial \lambda}$ at $\lambda = 0$ to obtain

$$\frac{\partial L}{\partial \lambda}\Big|_{\lambda=0} = \sum_i \left[\frac{\partial L}{\partial Q_i} \frac{\partial Q_i}{\partial \lambda} + \frac{\partial L}{\partial \dot{Q}_i} \frac{\partial \dot{Q}_i}{\partial \lambda} \right]_{\lambda=0}$$

$$= \sum_i \left[\frac{\partial L}{\partial q_i} \frac{\partial Q_i}{\partial \lambda} + \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{Q}_i}{\partial \lambda} \right]_{\lambda=0}$$

$$= \sum_i \frac{d}{dt} \left[\frac{\partial L}{\partial \dot{q}_i} \frac{\partial Q_i}{\partial \lambda} \right]_{\lambda=0} = \frac{df}{dt}.$$

Therefore,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\frac{\partial Q_i}{\partial \lambda} - f\right) = 0 \qquad (4.2)$$

is conserved.   □

Example 5: Let us look at spatial translations. Consider the Lagrangian of a closed system (no external forces), which is unchanged if all particles are translated by the same amount. For every particle, we apply the transformation

$$\vec{r}_i(t) \to \vec{r}_i(t) + \lambda \hat{n} = \vec{R}_i(t, \lambda).$$

The Lagrangian is

$$L = \frac{1}{2}\sum_i m_i \dot{\vec{r}}_i^2 - V(\{\vec{r}_i - \vec{r}_j\}).$$

Note that L is unchanged for this transformation, so we can set $f = 0$. The conserved property is then

$$\sum_a p_a \frac{\partial Q_a}{\partial \lambda}\bigg|_{\lambda=0} = \sum_i \frac{\partial Q_i}{\partial R_i}\hat{n}$$

$$= \sum_i \vec{p}_i.$$

The total momentum is $\vec{P} = \sum \vec{p}_i$, so Noether's theorem immediately tells us that the total momentum of a closed system is conserved.

Example 6: Let us look at time translations now. This is represented by $q_i(t) \to Q_i(t, \lambda) = q_i(t + \lambda)$. In general, L depends on t in two ways:
   1. Via the generalized coordinate i and velocities.
   2. Explicitly (for example, a time dependent potential energy).
However, if L has no explicit t dependence, then time translation *does* change L, but

$$\frac{\partial L}{\partial \lambda} = \frac{\partial L(Q_i(\lambda), \dot{Q}_i(\lambda))}{\partial \lambda}$$

$$= \frac{\partial L(q_i(t + \lambda), \dot{q}_i(t + \lambda))}{\partial \lambda}$$

$$= \frac{dL(q_i(t + \lambda), \dot{q}_i(t + \lambda))}{dt}$$

since L has the same dependence on λ as on t! Therefore $\frac{dL}{d\lambda} = \frac{dL}{dt}$ so $f = L$, and so time translation is a symmetry. Similarly,

$$\frac{\partial Q_i(t + \lambda)}{\partial \lambda} = \frac{\partial}{\partial \lambda}q_i(t + \lambda) = \frac{dq_i}{dt} = \dot{q}_i$$

and so the conserved quantity is

$$\sum_i \frac{\partial L}{\partial \dot{q}_i}\frac{\partial L}{\partial \lambda}\bigg|_{\lambda=0} - f = \sum_i \dot{q}_i\left[\frac{\partial L}{\partial \dot{q}_i}\right] - L \equiv H,$$

where H is the **Hamiltonian.**

What exactly is the Hamiltonian? We claim that usually, the Hamiltonian is the total energy. We can check that since $L = T - V$, then

$$\sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} = \sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i}.$$

Consider some T that is quadratic in q_i. We use **Euler's Theorem of Homogenous Functions,** which says that if

$$f(\alpha x_1, \ldots, \alpha x_n) = \alpha^n f(x_1, \ldots, x_n)$$

for some n, then $\sum_i x_i \dfrac{\partial f}{\partial x_i} = nf(x)$. Therefore, then T is homogenous with degree 2. Therefore:

$$\sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = \alpha T. \tag{4.3}$$

The Hamiltonian is then

$$H = \sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} - \mathcal{L} = \alpha T - (T - V) = T + V.$$

However, note that the Hamiltonian is not necessarily the total energy. This can occur in an open system where there is some energy transfer, making T not quadratic.

Example 7: Consider a rotationally invariant system. We do so by applying an infinitesimal rotation $\delta\lambda$ about an axis $\hat{n}$. Note that $\delta\vec{r} \perp \hat{r}$ and $\delta\vec{r} \perp \hat{n}$. We can write:

$$\delta\vec{r} = \hat{n} \times \vec{r}\delta\lambda \implies \frac{d\vec{r}}{d\lambda} = \hat{n} \times \vec{r}.$$

Applying Noether's Theorem, we get the conserved quantity to be

$$\frac{d}{dt} \sum_i \vec{p}_i \cdot (\vec{n} \times \vec{r}_i) = 0.$$

However since $a \cdot (b \times c) = b \cdot (a \times c)$, we end up with

$$\frac{d}{dt} \sum_i \vec{p}_i \times \vec{r}_i = 0.$$

This quantity is the **angular momentum,** which we will denote as $\vec{M}$.

Every closed system in three dimensions has 7 conserved quantities, known as integrals of motion. These are the energy, 3 components of linear momentum, and 3 components of angular momentum.