Particle Physics

Preliminaries for Particle Physics

Last Update: July 26, 2025

Preface

Author and Contact Information

- Written by: **Jai Nasu**
- Contact: **Discord Account**

Preface

This is a introductory text for preliminaries of particle physics and covers the basics of:

- Analytical Mechanics
- Field Theory
- Special Relativity
- Quantum Mechanics
- Some Mathematics
- Extra topics

The text is intended to be a self-contained introduction to the topics, and is meant to be a mere introduction - a preparation for not freaking out before reading more advanced texts. By that I mean the understanding on:

- Greek letters (may be most important, really.)
- High School calculus + partial derivatives
- Vectors and (some matrix algebra)
- Newtonian Mechanics and Electromagnetism
- Some basic knowledge on Lagrangian and Hamiltonian mechanics

Contents

Chapter 1 Notations and Conventions

1.1 General Notations

1.1.1 Vectors and Functions

A vector (in a non-relativistic context) is denoted by an arrow:

$$\vec{x}, \vec{r}, \vec{p}, \vec{F}, \vec{v}, \vec{a}, \ldots \tag{1.1}$$

If a function depends on multiple variables (e.g., $\{q_i\} = \{q_1, q_2, q_3, \ldots\}$), we may denote it in a short-handed way:

$$f(q_1, q_2, q_3, \ldots) \iff f(q_i) \tag{1.2}$$

A partial derivative by the variable x is denoted by:

$$\frac{\partial f}{\partial x} = \partial_x f, \tag{1.3}$$

For a general coordinate system $(x_1, x_2, x_3, \ldots)$, we denote the partial derivative by:

$$\frac{\partial f}{\partial x_i} = \partial_i f \tag{1.4}$$

1.2 Relativity Notations

1.2.1 Vectors(in Relativity)

For 4-vectors and tensors in relativity, sans-serif font is used:

$$\mathsf{x}, \mathsf{p}^\mu, \mathsf{g}_{\mu\nu} \tag{1.5}$$

Each component of a 4-vector is denoted by a superscript or subscript:

$$\mathsf{x}^\mu = x^0, x^1, x^2 \text{ or } x^3 \tag{1.6}$$

and without an index, it represents the entire vector:

$$\mathsf{x} = (x^0, x^1, x^2, x^3) = (ct, \vec{x}) \tag{1.7}$$

1.2.2 Einstein's Summation Convention

If an index such as i, j, k, μ, ν, etc., appears twice in a single term, it implies summation over that index. For example:

Example 1.2.1 (Examples)

$$x_i y_i = \vec{x} \cdot \vec{y} = \sum_{i=1}^{3} x_i y_i \tag{1.8}$$

$$g_{\mu\nu} u^\mu v^\nu = \sum_{\mu,\nu=0}^{3} g_{\mu\nu} u^\mu v^\nu \tag{1.9}$$

Note that the greek index such as μ, ν usually runs from 0 to 3, while the latin index such as i, j, k usually runs from 1 to 3.

1.2.3 Metric Tensor

As per the convention in particle physics, we take the $(+, -, -, -)$ metric signature, and the Minkowski metric $\eta_{\mu\nu}$ is denoted by:

$$\eta_{\mu\nu} = \begin{pmatrix} \eta_{00} & \eta_{01} & \eta_{02} & \eta_{03} \\ \eta_{10} & \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{20} & \eta_{21} & \eta_{22} & \eta_{23} \\ \eta_{30} & \eta_{31} & \eta_{32} & \eta_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.10}$$

1.2.4 Units

In principle, we use SI units:

$$\text{Length} = [L] = \text{m} \tag{1.11}$$

$$\text{Time} = [T] = \text{s} \tag{1.12}$$

$$\text{Mass} = [M] = \text{kg} \tag{1.13}$$

$$\text{Charge} = [C] = \text{C} \tag{1.14}$$

However, for particle physics, we often use natural units, where:

$$\hbar = c = 1 \tag{1.15}$$

and the unit of energy in electronvolts (eV) is used.

Chapter 2

Analytical Mechanics

2.1 Lagrange Formalism

2.1.1 Newtonian Mechanics

In Newtonian mechanics, the motion of a particle is described through a few important quantities: for a particle of (inertial) mass m, position $\vec{r}$, we have

$$\text{velocity} : \vec{v} = \frac{d\vec{r}}{dt} \tag{2.1}$$

$$\text{acceleration} : \vec{a} = \frac{d\vec{v}}{dt} = \frac{d^2\vec{r}}{dt^2} \tag{2.2}$$

$$\text{momentum} : \vec{p} = m\vec{v} = m\frac{d\vec{r}}{dt} \tag{2.3}$$

and the relations between these quantities, in the presence of external forces $\vec{F}^{(i)}_{\text{ext}}$ acting on the particle, are given by *Newton's second law*:

$$\frac{d\vec{p}}{dt} = m\frac{d^2\vec{r}}{dt^2} = \sum_i \vec{F}^{(i)}_{\text{ext}} \tag{2.4}$$

The work done by such forces is given by

$$W_{\text{total}} = \sum_i \int_l d\vec{x} \cdot \vec{F}^{(i)}_{\text{ext}}, \quad \text{where } l \text{ is the path of the particle.} \tag{2.5}$$

This is the energy change of the particle through the motion:

$$\begin{aligned} W_{\text{total}} &= \sum_i \int_l d\vec{x} \cdot \vec{F}^{(i)}_{\text{ext}} && (2.6) \\ &= \int_{t_i}^{t_f} dt\, \vec{v} \cdot m\frac{d\vec{v}}{dt} && (2.7) \\ &= \int_{t_i}^{t_f} dt\, \frac{m}{2}\frac{d}{dt}\vec{v}^{\,2} && (2.8) \\ &= \frac{m}{2}\vec{v}_f^{\,2} - \frac{m}{2}\vec{v}_i^{\,2} && (2.9) \end{aligned}$$

meaning that $m\vec{v}^2/2$ is the energy due to the motion of the particle: the **kinetic energy** T:

$$T = \frac{m}{2}\vec{v}^2 \tag{2.10}$$

Now, often, the external force acting on the particle is due to a potential V:

$$\vec{F}_{\text{ext}} = -\nabla V \tag{2.11}$$

For example, for a 1D spring, the potential is given by

Example 2.1.1 (1D spring/ Harmonic potential)

$$V = \frac{1}{2}kx^2 \implies F_{\text{ext}} = -kx \tag{2.12}$$

or the electrostatic potential:

Example 2.1.2 (Electrostatic potential)

$$V = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r} \implies F_{\text{ext}} = -\nabla V = -\frac{q_1 q_2}{4\pi\varepsilon_0 r^2}\hat{r} \tag{2.13}$$

Now, for a particle whose the external forces are given by a potential:

$$m\frac{d^2\vec{r}}{dt^2} = -\nabla V \iff -m\frac{d^2\vec{r}}{dt^2} - \nabla V = 0 \tag{2.14}$$

This looks as if the forces $-\nabla V$ and $m\ddot{\vec{r}}$ are in equilibrium. So, if we move a particle by an infinitisimal distance $\delta\vec{r}$, the total work done by these forces must be zero:

$$\left(-m\frac{d^2\vec{r}}{dt^2} - \nabla V\right) \cdot \delta\vec{r} = 0 \tag{2.15}$$

at any time t. We want to apply this for entire path of the motion of the particle, from t_i to t_f. Then, the integral of this equation over the time interval $[t_i, t_f]$ gives

$$\delta I = \int_{t_i}^{t_f} dt \left(-m\frac{d^2\vec{r}}{dt^2} - \nabla V\right) \cdot \delta\vec{r} = 0 \tag{2.16}$$

now, we can apply integration by parts:

$$\frac{d}{dt}\left[\dot{\vec{r}}\cdot\delta\vec{r}\right] = \ddot{\vec{r}}\cdot\delta\vec{r} + \dot{\vec{r}}\cdot\frac{d\,\delta\vec{r}}{dt} \tag{2.17}$$

$$= \ddot{\vec{r}}\cdot\delta\vec{r} + \dot{\vec{r}}\cdot\delta\dot{\vec{r}} \tag{2.18}$$

$$\Longleftrightarrow \quad -m\ddot{\vec{r}} = -\frac{d}{dt}\left(\dot{\vec{r}}\cdot\delta\vec{r}\right) + \dot{\vec{r}}\cdot\delta\dot{\vec{r}} \tag{2.19}$$

$$= -\frac{d}{dt}\left(m\dot{\vec{r}}\cdot\delta\vec{r}\right) + \delta\left(\frac{m}{2}\left(\frac{d\vec{r}}{dt}\right)^2\right), \tag{2.20}$$

where we used the commutitivity of $\frac{d(\cdot)}{dt}$ and $\delta(\cdot)$. Then the integral becomes

$$\delta I = \int_{t_i}^{t_f} dt\left(\delta\left[\frac{m}{2}\left(\frac{d\vec{r}}{dt}\right)^2\right] - \delta V\right) \tag{2.21}$$

$$= \int_{t_i}^{t_f} dt\,(\delta T - \delta V) = 0 \tag{2.22}$$

$$\Longleftrightarrow \quad \delta I = \delta\int_{t_i}^{t_f} dt\,(T - V) = 0 \tag{2.23}$$

[1]

2.1.2 Lagrangian and Variational Principle

In Lagrangian mechanics, we will use a different approach to describe the motion of a particle than the Newtonian mechanics. Instead of using the usual Euclidean space, we will use a *configuration space* $\mathcal{C}$, which is the space of all possible positions of the particle. This space is spanned by the so-called **generalized coordinates** q_i, and their time derivatives (or generalized velocity) $\dot{q}_i := \frac{dq_i}{dt}$. Now, let us define quantities called the **Lagrangian** L and **action** S:

Definition 2.1: Lagrangian L and Action S

The Lagrangian L is defined as the difference between the kinetic energy T and potential energy V of a particle:

$$L := T - V, \qquad S[L] := \int_{t_i}^{t_f} dt\,L \tag{2.24}$$

Now, **variation** of the action δS is given by

$$\delta S = \delta\int_{t_i}^{t_f} dt\,(T - V) \tag{2.25}$$

which is an identical expression to (2.23). So, we postulate that the motion of the particle is such that the action is stationary:

Principle 2.1: Variational Principle

The motion of a particle is such that the action S is stationary, i.e., $\delta S = 0$.

$$\delta S = \delta \int_{t_i}^{t_f} dt\, L(q_i(t), \dot{q}_i(t), t) = 0 \tag{2.26}$$

As to show why this maybe useful, let us compare between Eq. (2.16):

$$\delta I = \delta S[L] = \int_{t_i}^{t_f} dt \left(-m\frac{d^2\vec{r}}{dt^2} - \nabla V \right) \cdot \delta\vec{r} = 0 \tag{2.27}$$

and Eq. (5.27):

$$\delta I = \delta S[f] = \int_B^A dx \left(\frac{\partial F}{\partial f} - \frac{d}{dx}\frac{\partial F}{\partial \frac{df}{dx}} \right) \delta f = 0 \tag{2.28}$$

which immidiately gives that the EoM is the Euler-Lagrange equation, if we set $F = L(q_i, \dot{q}_i, t)$:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i} \right) = 0 \iff m\frac{d^2\vec{r}}{dt^2} = -\nabla V \tag{2.29}$$

The generalization to multiple particles is straightforward:

$$L = \sum_n \frac{m_n}{2}\dot{\vec{r}}_n^2 - V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \tag{2.30}$$

The Euler-Lagrange equation is given by

$$\frac{\partial L}{\partial \vec{r}_i^{\,(N)}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\vec{r}}_i^{\,(N)}} \right) = 0 \tag{2.31}$$

2.1.3 Harmonic Oscillator

Let us consider a particle of mass m in a 1D harmonic potential:

$$V = \frac{1}{2}kx^2 \implies F_{\text{ext}} = -\nabla V = -kx \tag{2.32}$$

then the Lagrangian is given by

$$L = T - V = \frac{m}{2}\dot{x}^2 - \frac{1}{2}kx^2 \tag{2.33}$$

The Euler-Lagrange equation yields

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0 \implies \frac{\partial L}{\partial x} = -kx, \quad \frac{\partial L}{\partial \dot{x}} = m\dot{x} \tag{2.34}$$

$$\implies \qquad -kx - \frac{d}{dt}(m\dot{x}) = 0 \tag{2.35}$$

$$\implies \qquad m\frac{d^2x}{dt^2} = -kx \tag{2.36}$$

which is the equation of motion of a harmonic oscillator in Newtonian mechanics. Here, notice that

$$\frac{\partial L}{\partial x} = -kx = -\nabla V = F_{\text{ext}} \tag{2.37}$$

and

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x} = p \tag{2.38}$$

these derivatives of the Lagrangian are the **generalized force** and **conjugate momentum**, respectively.

$$\frac{\partial L}{\partial q} = F_{\text{g}}, \quad \frac{\partial L}{\partial \dot{q}} = p_{\text{g}} \tag{2.39}$$

2.2 Hamiltonian Formalism

2.2.1 Legendre Transform

Consider a function $f(x,y)$. The total differential of f is given by

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy \tag{2.40}$$

Often we want to find another function g such that

$$g = p \cdot y - f(x,y) \tag{2.41}$$

and the total differential of g is given by

$$\begin{aligned} dg &= \frac{\partial g}{\partial p}\,dp + \frac{\partial g}{\partial y}\,dy - df && (2.42)\\ &= y\,dp + p\,dy - \frac{\partial f}{\partial x}\,dx - \frac{\partial f}{\partial y}\,dy && (2.43)\\ &= y\,dp - \frac{\partial f}{\partial x}\,dx + \left(p - \frac{\partial f}{\partial y}\right)dy && (2.44) \end{aligned}$$

for this function g to be a function of x and p, we need to have

$$p = \frac{\partial f}{\partial y}, \quad \frac{\partial g}{\partial x} = -\frac{\partial f}{\partial x}. \tag{2.45}$$

This is called the **Legendre transform** of f.

Definition 2.2: Legendre transform(Analytical Mechanics)

The Legendre transform of a function $f(x, y)$ is defined as

$$g(p, x) = p \cdot y - f(x, y) \tag{2.46}$$

where

$$p = \frac{\partial f}{\partial y}, \quad \frac{\partial g}{\partial x} = -\frac{\partial f}{\partial x} \tag{2.47}$$

.

[2]

2.2.2 Hamiltonian and Canonical Equations

Now, we define Hamiltonian $H = H(q, p)$ as the Legendre transform of Lagrangian $L = L(q, \dot{q})$:

Definition 2.3: Hamiltonian H

The Hamiltonian $H(q, p)$ is defined as the Legendre transform of the Lagrangian $L(q, \dot{q})$ $(\dot{q} \to p)$:

$$H(q, p) = p \cdot \dot{q} - L(q, \dot{q}) \tag{2.48}$$

where

$$p = \frac{\partial L}{\partial \dot{q}}, \quad \frac{\partial H}{\partial q} = -\frac{\partial L}{\partial q} \tag{2.49}$$

Note:
In Lagrange formalism, generalized coordinate span the configuration space, while in Hamilton formalism, generalized coordinates and **generalized momenta** or **conjugate momenta** span the **phase space**.

This is actually a physically intuitive quantity. Since p is defined as the derivative of Lagrangian L:

$$p = \frac{\partial L}{\partial \dot{q}} = \frac{\partial}{\partial \dot{q}}\left(\frac{1}{2}m\dot{q}^2 - V(q)\right) = m\dot{q} \tag{2.50}$$

and we can rewrite the Hamiltonian as

$$\begin{aligned} H(q,p) &= p \cdot \dot{q} - L(q,\dot{q}) && (2.51)\\ &= m \cdot \dot{q}^{\,2} - \left(\frac{1}{2}m\dot{q}^2 - V(q)\right) && (2.52)\\ &= \frac{1}{2}m\dot{q}^{\,2} + V(q) = T + V && (2.53) \end{aligned}$$

This is the total energy of the system, for the case of non-velocity dependent potential $V(q)$.

This relation is useful because we can obtain Lagrangian L from the Hamiltonian H as well:

$$L(q,\dot{q}) = \dot{q} \cdot p - H(q,p) \implies \dot{q} = \frac{\partial H}{\partial p} \tag{2.54}$$

and the Euler-Lagrange equation can be rewritten in terms of Hamiltonian:

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0 \implies -\frac{\partial H}{\partial q} - \frac{d}{dt}(p) = 0 \iff \quad \dot{p} = -\frac{\partial H}{\partial q} \tag{2.55}$$

These two equations are called the **canonical equations** or **Hamilton's equations**:

Theorem 2.2.1 Canonical Equations

The relationship between a mechanical variable q and its canonical conjugate variable p is given by the canonical equations:

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q} \tag{2.56}$$

For multiple degrees of freedom, we can write the Hamiltonian as

$$H(\{q_i\},\{p_i\},t) = \sum_i p_i\dot{q}_i - L(\{q_i\},\{\dot{q}_i\},t) \tag{2.57}$$

and the canonical equations as

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{2.58}$$

where

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \quad \frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i} \tag{2.59}$$

2.2.3 Poisson Bracket

Now, consider how a physical quantity $X(q_i, p_i, t)$ changes with time:

$$\frac{dX(q_i, p_i, t)}{dt} = \frac{\partial X}{\partial t} + \frac{\partial X}{\partial q_i}\dot{q}_i + \frac{\partial X}{\partial p_i}\dot{p}_i \tag{2.60}$$

from the canonical equations, the second part becomes:

$$\frac{\partial X}{\partial q_i}\dot{q}_i + \frac{\partial X}{\partial p_i}\dot{p}_i = \frac{\partial X}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial X}{\partial p_i}\frac{\partial H}{\partial q_i} \tag{2.61}$$

Since this term only depend on X and H, given q_i and p_i, we can define a new quantity called the **Poisson bracket**:

Definition 2.4: Poisson Bracket

The Poisson bracket of two physical quantities A and B is defined as

$$\{A,\, B\} := \sum_i \left(\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i} \right) \tag{2.62}$$

If a physical quantity X does not explicitly depend on time, we can rewrite the time derivative as:

$$\frac{dX(q_i, p_i)}{dt} = \{X,\, H\} \tag{2.63}$$

Due to this property, we often call the Hamiltonian H the **generator** of time evolution.

Now, what happens if X happens to be q_i or p_i? A physical quantity q_i or p_i can be written as a function of q_i and p_i - simply as itself (noting that no explicit time dependence is present):

$$q_i(q_i, p_i) = q_i, \quad p_i(q_i, p_i) = p_i \tag{2.64}$$

Then, the Poisson bracket of q_i and H is given by

$$\{q_i,\, H\} = \sum_j \left(\frac{\partial q_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial q_i}{\partial p_j}\frac{\partial H}{\partial q_j} \right) \tag{2.65}$$

$$= \sum_j \left(\delta_{ij}\frac{\partial H}{\partial p_j} - 0 \right) = \frac{\partial H}{\partial p_i} \tag{2.66}$$

where Eq. (5.31) is used. Similarly, the Poisson bracket of p_i and H is given by

$$\{p_i,\, H\} = \sum_j \left(\frac{\partial p_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial p_i}{\partial p_j}\frac{\partial H}{\partial q_j} \right) \tag{2.67}$$

$$= \sum_j \left(0 - \delta_{ij}\frac{\partial H}{\partial q_j} \right) = -\frac{\partial H}{\partial q_i} \tag{2.68}$$

Thus, we can rewrite the canonical equations in terms of Poisson bracket:

Theorem 2.2.2 Canonical Equations with Poisson Bracket

The canonical equations can be rewritten in terms of Poisson bracket as follows:

$$\dot{q}_i = \{q_i, H\}, \quad \dot{p}_i = \{p_i, H\} \tag{2.69}$$

Finally, note that the Poisson bracket of q_i and p_j gives:

$$\{q_i, p_j\} = \sum_k \frac{\partial q_i}{\partial q_k}\frac{\partial p_j}{\partial p_k} - \frac{\partial q_i}{\partial p_k}\frac{\partial p_i}{\partial q_k} \tag{2.70}$$

$$= \sum_k \delta_{ik}\delta_{jk} - 0 = \delta_{ij} \tag{2.71}$$

Theorem 2.2.3 Canonical Conjugate Relation in Analytical Mechanics

The Poisson bracket of the generalized coordinate q_i and its conjugate momentum p_j is given by

$$\{q_i, p_j\} = \delta_{ij} \tag{2.72}$$

2.2.4 Generators

Let us remark on some more important properties of the Poisson bracket: We know that for q_i and p_j, the Poisson bracket is given by

$$\{q_i, p_j\} = \delta_{ij} \tag{2.73}$$

but what if we replace q_i to some other physical quantity $f(q_i, p_j)$?

Let us write down the Poisson bracket explicitly:

$$\{f(q_i, p_j), p_k\} = \sum_l \frac{\partial f(q_i, p_j)}{\partial q_l}\frac{\partial p_k}{\partial p_l} - \frac{\partial f(q_i, p_j)}{\partial p_l}\frac{\partial p_k}{\partial q_l} \tag{2.74}$$

the second term is always zero, so

$$= \sum_l \frac{\partial f(q_i, p_j)}{\partial q_l}\delta_{kl} \tag{2.75}$$

$$= \frac{\partial f(q_i, p_j)}{\partial q_k} \tag{2.76}$$

so we have seen that the taking the derivative of a physical quantity with respect to q_k is equivalent to taking the Poisson bracket with p_k. By linearity,

$$\{f(q_i, p_j), \epsilon p_k\} = \epsilon \frac{\partial f(q_i, p_j)}{\partial q_k} \tag{2.77}$$

2.3 Symmetry of a System

2.3.1 Symmetry

A fundamental concept in physics is **symmetry**. Let us see an example of symmetry in geometry:

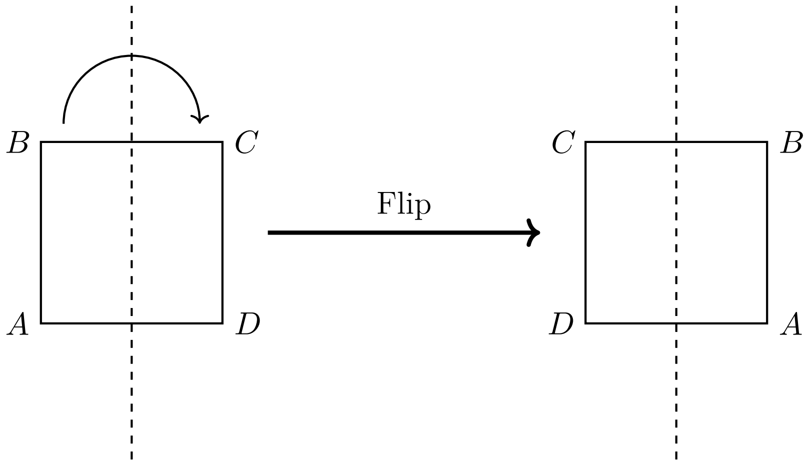

Figure 2.1: A square and its reflection.

Notice how we flipped the square, and it looks the same without the labels. In a more formal sense, we can say that the system(square) is **invariant** under the transformation(flip).

2.3.2 Transformations

For example, Newtonian mechanics(or Newton's equation of motion) has some symmetries.

Example 2.3.1 (Equation of Motion)
For example, the equation of motion is unchanged under Galilean transformation.

$$\sum \vec{F} = m\vec{a} \tag{2.78}$$

Definition 2.5: Galilean Transformation

The **Galilean transformation** is a transformation of coordinates from a stationary observer to a moving observer with constant velocity v.

$$\vec{\xi}(\vec{x}, t) = \vec{x} - \vec{v}t \tag{2.79}$$

where $\vec{\xi}$ is the new coordinate, $\vec{x}$ is the old coordinate, $\vec{v}$ is the velocity of the observer, and t is time.

and space/ time reversal:

Definition 2.6: Space/Time Reversal

The **space reversal** is a transformation of coordinates that flips the sign of the position vector:

$$\vec{x}' = -\vec{x} \tag{2.80}$$

The **time reversal** is a transformation of time that flips the sign of time:

$$t' = -t \tag{2.81}$$

In the context of analytical mechanics, the system is described by the action S. If the variation of the action δS is invariant under a transformation, we say that the system has a symmetry under that transformation.

Notice that the Lagrangian has a degree of freedom(that is, the Lagrangian can be modified by adding a total derivative and variation of action remains the same):

Theorem 2.3.1 Degree of Freedom in Lagrangian

Adding a total derivative of time to the Lagrangian does not change the action:

$$L'(q_i, \dot{q}_i, t) = L(q_i, \dot{q}_i, t) + \frac{d}{dt} f(q(t), t) \implies \delta S'[L'] = \delta S[L] \tag{2.82}$$

Proof: The action is given by

$$\begin{aligned} S'[L'] &= \int_{t_i}^{t_f} dt\, L'(q_i(t), \dot{q}_i(t), t) && (2.83) \\ &= \int_{t_i}^{t_f} dt\, L(q_i(t), \dot{q}_i(t), t) + \int_{t_i}^{t_f} dt\, \frac{d}{dt} f(q(t), t) && (2.84) \\ &= S[L] + \Big[f(q_i, t) \Big]_{t_i}^{t_f} = S[L] + f(q_i(t_f), t_f) - f(q(t_i), t_i) && (2.85) \end{aligned}$$

Since by taking the variation, $\delta q(t_f) = \delta q(t_i) = 0$, the last term is zero. Thus

$$\delta S'[L'] = \delta S[L] \tag{2.86}$$

□

2.3.3 Point Transformation

In Analytical Mechanics, there is a general family of transformations called **point transformation**.

Definition 2.7: Point Transformation

A **point transformation** is a transformation of the generalized coordinates q_i and time t to new coordinates Q_i

$$Q_j = Q_j(q_i, t) \tag{2.87}$$

Note: if there are N generalized coordinates q_i, then there must be N new coordinates Q_j.

Example 2.3.2 (Point Transformation)

For example, changing from Cartesian coordinates to polar coordinates is a point transformation:

$$(x, y) \to (r, \theta) \qquad \Longleftrightarrow \begin{cases} r & = \sqrt{x^2 + y^2} \\ \theta & = \arctan(\frac{y}{x}) \end{cases} \tag{2.88}$$

$$(r, \theta) \to (x, y) \qquad \Longleftrightarrow \begin{cases} x & = r \cos\theta \\ y & = r \sin\theta \end{cases} \tag{2.89}$$

notice how each new coordinate only depends on the old coordinates, not the velocity.

Point transformation does not change the action, i.e., the variation of the action is invariant under point transformation:

Theorem 2.3.2 Invariance of Euler-Lagrange Equation

The Euler-Lagrange equation is invariant under point transformation:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0 \implies \frac{\partial L}{\partial Q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_j}\right) = 0 \tag{2.90}$$

proof is given in Sec. 6.1.

Now, using the point transformation, we can give the definition of symmetry, in the context of analytical mechanics:

Definition 2.8: Symmetry of the System

A system is said to have a **symmetry** under a point transformation $q_i \to Q_i$ if the change in Lagrangian is up to the degree of freedom:

$$L(Q_i, \dot{Q}_i, t) = L(q_i, \dot{q}_i, t) + \underbrace{\frac{d}{dt} f(q_i, t)}_{\text{change in } L} \tag{2.91}$$

where L is the Lagrangian of the system before the point transformation. Note that the Lagrangian of LHS has the same functional form as RHS(i.e. Q nad $\dot{Q}$ are substituted for q and $\dot{q}$). Or equivalently, the action is invariant under the point transformation:

$$S[Q_i, \dot{Q}_i, t] = S[q_i, \dot{q}_i, t] + \int_{t_i}^{t_f} dt\, \frac{d}{dt} f(q_i, t) = S[q_i, \dot{q}_i, t] \tag{2.92}$$

Importantly, the point transformation can be either discrete or continous, and collection of point transformations that preserve the action is called a **group**. In short, it means three things:

1. There is an identity transformation (i.e. no change)
2. If there are multiple transformations, there is a single transformation that is the result of applying those transformations in sequence.
3. If there is a transformation, there is an **inverse transformation** that undoes the transformation.

This concept is very important in particle physics.

2.3.4 Symmetry of the System including Time

$$t' = t'(q_i, t) \tag{2.93}$$
$$Q_j(t') = Q_j(q_i, t) \tag{2.94}$$

and define

$$\dot{Q}_j := \frac{dQ(t')}{dt'} \tag{2.95}$$

For this transformation to be a symmetry of the system, the change in action must keep the variation of action invariant:

$$\begin{aligned} S[Q_i] &= \int_{t'_i}^{t'_f} dt'\, L(Q_i(t'), \dot{Q}_i(t'), t') && (2.96)\\ &= \int_{t_i}^{t_f} dt \frac{dt'}{dt} L(Q_i(t'), \dot{Q}_i(t'), t') && (2.97) \end{aligned}$$

For the variation of action to be invariant, the change in Lagrangian must be up to the degree of freedom:

$$\frac{dt'}{dt}L(Q_i(t'),\dot{Q}_i(t'),t') = L(q_i,\dot{q}_i,t) + \frac{d}{dt}f(q_i,t) \tag{2.98}$$

Thus the symmetry under the transformation including time is defined as:

Definition 2.9: Symmetry of the System including Time

A system is said to have a **symmetry** under a transformation including time $t' = t'(q_i,t)$ and $Q_j(t') = Q_j(q_i,t)$ if the change in Lagrangian can be written as follows:

$$\frac{dt'}{dt}L(Q_i(t'),\dot{Q}_i(t'),t') = L(q_i,\dot{q}_i,t) + \frac{d}{dt}f(q_i,t) \tag{2.99}$$

2.3.5 Noether's Theorem

Here, I introduce Noether's theorem, highlights the direct correspondence between symmetries and conservation laws:

Theorem 2.3.3 Noether's Theorem under Point Transformation

Every real paramter of a continuous symmetry of the action corresponds to a conserved quantity $\mathcal{Q}$. If the continous symmetry is given by an infinitesimal transformation: $Q_i(t) = q_i(t) + \epsilon F_i(q(t),t) + \mathcal{O}(\epsilon^2)$,

$$\mathcal{Q} = \sum_i F_i(q(t),t)\frac{\partial L(Q_i(t),\dot{Q}_i(t),t)}{\partial \dot{q}_i} - \Lambda(q(t),t) \tag{2.100}$$

where $\Lambda(q(t),t)$ is the change in the Lagrangian under the infinitesimal transformation.

This might be quite abstract, so let us see an example of Noether's theorem in action. First, take the generalized coordinates $\vec{q}(t)$ to be the position vector $\vec{r}(t)$ of a particle in space. Then the new coordinates $Q_i(t)$ can be defined as:

$$Q_i(t) = r_i(t) + \epsilon F_i(\vec{r}(t),t) + \mathcal{O}(\epsilon^2) \tag{2.101}$$

Example 2.3.3 (Translation in Space)

If we consider a constant translation in space, we can write the infinitesimal transformation as:

$$Q_i(t) = r_i(t) + \epsilon a_i + \mathcal{O}(\epsilon^2) \tag{2.102}$$

(2.103)

If Lagrangian only depends on $\dot{\vec{r}}_i(t)$ or $\vec{r}_n - \vec{r}_m$ (often called internal forces, which satisfy action-reaction law), the effect of constant translation is cancelled out (i.e. $\Lambda = 0$):

$$\implies \quad \mathcal{Q} = \sum_i a_i \frac{\partial L}{\partial \dot{r}_i} \implies \dot{\mathcal{Q}} = \sum_i a_i \frac{d}{dt}\frac{\partial L}{\partial \dot{r}_i} = 0 \tag{2.104}$$

since a_i is arbitrary, the only possibility is that

$$\frac{\partial L}{\partial \dot{r}_i} = \text{const.} \tag{2.105}$$

which is the momentum of the particle!

There are many more symmetries and corresponding conserved quantities, which I summarize below:

Table 2.1: Symmetries and Corresponding Conserved Quantities

Symmetry	Transformation	Conserved Quantity $\mathcal{Q}$	Group G
Spatial Translation	$\vec{r}' = \vec{r} + \epsilon\vec{a}$	Linear Momenum $\sum \vec{p}$	$\mathbb{R}^3$
Temporal translation	$t' = t + \epsilon T(q_i(t), t)$	Energy E	$\mathbb{R}$
Spatial Rotation	$\vec{r}' = R(\theta)\vec{r}$	Angular Momentum $\sum \vec{L}$	$SO(3)$
Spin Rotation	$\vec{r}' = R(\theta)\vec{r}$	Spin $\sum \vec{S}$	$SO(3)$
Galilean Transformation	$\vec{r}' = \vec{r} - \vec{v}t$	$\sum \vec{p} - t\vec{P}_{\text{obs}}$	$\mathbb{R}^3$
Phase Transformation	$\psi' = e^{i\theta}\psi$	Number of particles N / Charge Q	$U(1)$

Note:
Proofs are in Section 6.2.

Theorem 2.3.4 Noether's Theorem under Transformation Including Time
If the system is symmetric under the transformation includes time translation:

$$t' = t + \epsilon T(q_i(t), t) + \mathcal{O}(\epsilon^2) \tag{2.106}$$
$$Q_i(t') = q_i(t) + \epsilon F_i(q_i(t), t) + \mathcal{O}(\epsilon^2) \tag{2.107}$$

There is a conserved quantity $\mathcal{Q}$ given by:

$$\begin{aligned} \mathcal{Q} = \sum_i \left[F_i(q_i(t), t) \frac{\partial L(q_i(t), \dot{q}_i(t), t)}{\partial \dot{q}_i} \right] - \Lambda(q_i, t) \\ - T(q_i(t), t) \sum_i \left[\dot{q}_i(t) \frac{\partial L(q_i(t), \dot{q}_i(t), t)}{\partial \dot{q}_i} - L(q_i(t), \dot{q}_i(t), t) \right] \end{aligned} \tag{2.108}$$

[3]

Chapter 3 Physics of Fields

3.1 Introduction: 1D String

3.1.1 Newtonian Derivation

Now let us consider the motion of 1D wave on a very long string.

- Tension: T
- Line density (mass per unit length): μ

Notice, that any point on the string is not moving in x direction, otherwise the string itself will be moving. Also, if we produce a "uniform" wave, then at all points on the string, the wave will have the same amplitude A , wavelength λ, and frequency f and velocity v. This means

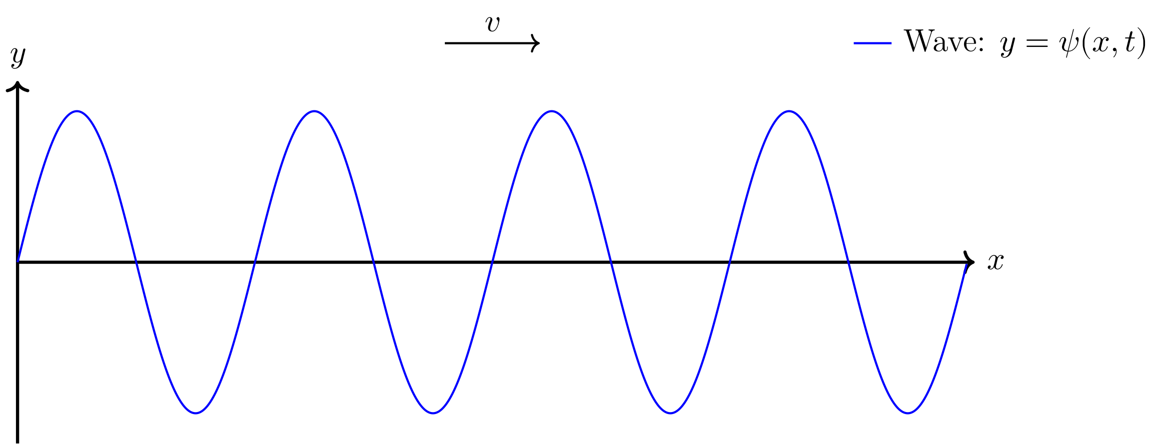

Figure 3.1: A plot of 1D wave on a string

that we can choose convinient point on the string, and calculate A, λ, f, and v, then the same values will be valid for all points on the string. For a moving wave, a stationary point of view is not very useful, so we will use a moving point of view: the Lagrange description in fluid mechanics.

Consider moving in the x direction, at the same constatn velocity v, as the wave. Then, our new x coordinate is given by the Galilean transformation:

$$\xi(x,t) = x - vt \tag{3.1}$$

In this new perspective, we do not have to worry about the motion of the wave, since we are moving with the wave. In short, the wave is stationary in the new coordinate system. This means that vertical displacement of the wave at a point ξ only depends on the ξ coordinate, and not on time:

$$\psi(x,t) = \psi(\xi) \tag{3.2}$$

and any point on the string seems to move at velocity $-v$.

Then immidiately,

$$\frac{\partial\psi(\xi)}{\partial x} = \frac{\partial\psi(\xi)}{\partial\xi}\frac{\partial\xi}{\partial x} = \frac{\partial\psi(\xi)}{\partial\xi}, \tag{3.3}$$

$$\frac{\partial\psi(\xi)}{\partial t} = \frac{\partial\psi(\xi)}{\partial\xi}\frac{\partial\xi}{\partial t} = -v\frac{\partial\psi(\xi)}{\partial\xi} \tag{3.4}$$

$$\implies \quad \frac{\partial^2\psi(\xi)}{\partial x^2} = \frac{\partial^2\psi(\xi)}{\partial\xi^2}, \quad \frac{\partial^2\psi(\xi)}{\partial t^2} = v^2\frac{\partial^2\psi(\xi)}{\partial\xi^2} \tag{3.5}$$

$$\implies \quad \frac{\partial^2\psi(x,t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\psi(x,t)}{\partial t^2} \tag{3.6}$$

we obtain the wave equation, but we should find its physical meaning.

Theorem 3.1.1 Wave Equation in 1D

The wave equation in 1D is given by:

$$\frac{\partial^2\psi(x,t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\psi(x,t)}{\partial t^2} \tag{3.7}$$

where $\psi(x,t)$ is the vertical displacement of the wave at point x and time t, and v is the velocity of the wave.

Now, let us focus on one of the peaks of the wave, at the point ξ_0. The tension on the

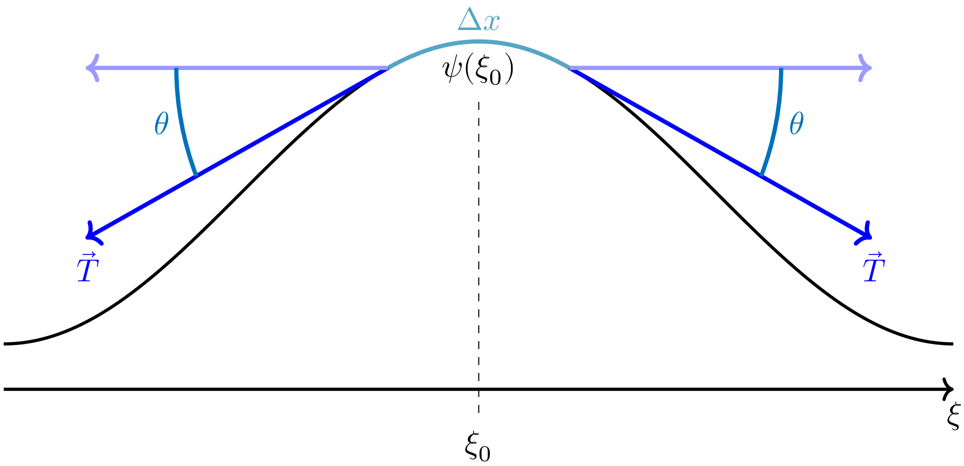

Figure 3.2: A peak of the wave

string at ξ_0 cancels out horizontally, but adds vertically:

$$F_x = T\cos\theta - T\cos\theta = 0 \tag{3.8}$$

$$F_y = F_{-y} + F_{+y} = -T\partial_x\psi\Big(\xi_0 + \frac{\Delta x}{2}\Big) + T\partial_x\psi\Big(\xi_0 + \frac{\Delta x}{2}\Big) \tag{3.9}$$

Thus the equation of motion for the infinitisimal segment of the string at ξ_0 is given by:

$$\vec{F} = m\ddot{\vec{r}} \iff T\Big(\partial_x\psi\Big(\xi_0 + \frac{\Delta x}{2}\Big) - \partial_x\psi\Big(\xi_0 - \frac{\Delta x}{2}\Big)\Big) = \mu\Delta x\,\partial_t^2\psi(\xi_0) \tag{3.10}$$

by rearranging,

$$\frac{1}{\Delta x}\Big(\partial_x\psi\Big(\xi_0 + \frac{\Delta x}{2}\Big) - \partial_x\psi\Big(\xi_0 - \frac{\Delta x}{2}\Big)\Big) = \frac{\mu}{T}\partial_t^2\psi(\xi_0) \tag{3.11}$$

as $\Delta x \to 0$, LHS becomes the derivative of the first derivative:

$$\partial_x^2\psi(\xi_0) = \frac{\mu}{T}\partial_t^2\psi(\xi_0) \tag{3.12}$$

Now, note that the infinitisimal section experiences force in y direction, while it moves in x direction, This is equivalent to a circular motion with radius $r = \psi(\xi_0)$:

$$F_c = \frac{mv^2}{r} = \frac{\mu\Delta x v^2}{\psi(\xi_0)} = F_y \tag{3.13}$$

In this limit of $\Delta x \to 0$, $\Delta x = 2\psi(\xi_0)\theta$,

$$\implies \quad 2\theta\mu v^2 = 2T\sin\theta \approx 2T\theta \tag{3.14}$$

$$\implies \quad v^2 = \frac{T}{\mu} \tag{3.15}$$

Thus, the wave equation can be written as:

$$\frac{\partial^2\psi(x,t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\psi(x,t)}{\partial t^2}, \quad \text{where } v = \sqrt{\frac{T}{\mu}} \tag{3.16}$$

3.2 Lagrange Formalism

3.2.1 Potential and Momentum Density

Now, notice that the force on the infinitisimal section is given by

$$\dot{p} = \mu\,\partial_t^2\psi\Delta x \tag{3.17}$$

since the extension of the string in $\Delta x/2$ is given by

$$\Delta l - \frac{\Delta x}{2} = \frac{\Delta x}{2}\sqrt{1 + \left[\partial_x \psi\left(\xi_0 - \frac{\Delta x}{2}\right)\right]^2} - \frac{\Delta x}{2} \tag{3.18}$$

$$\approx \frac{\Delta x}{2}\,\partial_x \psi\left(\xi_0 - \frac{\Delta x}{2}\right) \tag{3.19}$$

The force on the infinitisimal section due to the extension is then given by

$$F = -k\left(\Delta l - \frac{\Delta x}{2}\right) = -k\frac{\Delta x}{2}\,\partial_x \psi\left(\xi_0 - \frac{\Delta x}{2}\right) = -T\partial_x \psi\left(\xi_0 - \frac{\Delta x}{2}\right) \tag{3.20}$$

$$\implies \quad V = \frac{k}{2}\left(\Delta l - \frac{\Delta x}{2}\right)^2 = \frac{k}{2}\left(\frac{\Delta x}{2}\,\partial_x \psi\left(\xi_0 - \frac{\Delta x}{2}\right)\right)^2 = \frac{T}{2}\frac{\Delta x}{2}\left(\partial_x \psi\left(\xi_0 - \frac{\Delta x}{2}\right)\right)^2 \tag{3.21}$$

We should define these quantity as the potential energy density and the momentum density:

Definition 3.1: Momentum and Potential Energy Density

The **momentum density** and **potential energy density** is defined as the momentum and potential energy per unit length of the string:

$$\pi = \mu\,\partial_t \psi, \quad \tilde{V} = \frac{T}{2}(\partial_x \psi)^2 \tag{3.22}$$

Now, since we have defined our potential energy (density) and momentum (density), we can write the Lagrangian:

$$L = T - V \tag{3.23}$$

$$= \int dx\,\tilde{T} - \int dx\,\tilde{V} \tag{3.24}$$

$$= \int dx\,\frac{\mu}{2}(\partial_t \psi)^2 - \frac{T}{2}(\partial_x \psi)^2 \tag{3.25}$$

thus we should define the Lagrangian density $\mathcal{L}$ as:

Definition 3.2: Lagrangian Density for 1D String

The **Lagrangian density** is defined as the Lagrangian per unit length of the string:

$$\mathcal{L}[\partial_x \psi, \partial_t \psi] = \frac{\mu}{2}(\partial_t \psi)^2 - \frac{T}{2}(\partial_x \psi)^2 \tag{3.26}$$

Then the Euler-Lagrange equation for the Lagrangian density $\mathcal{L}$ is given by:

$$\partial_t\left(\frac{\partial \mathcal{L}}{\partial(\partial_t\psi)}\right)+\partial_x\left(\frac{\partial \mathcal{L}}{\partial(\partial_x\psi)}\right)=\frac{\partial \mathcal{L}}{\partial \psi} \tag{3.27}$$

$$\implies \quad \partial_t(\mu\,\partial_t\psi)+\partial_x(-T\partial_x\psi)=0 \tag{3.28}$$

$$\implies \quad \mu\,\partial_t^2\psi-T\partial_x^2\psi=0 \tag{3.29}$$

$$\implies \quad \partial_x^2\psi=\frac{\mu}{T}\partial_t^2\psi \tag{3.30}$$

which is the wave equation in 1D. Notice that our mechanical variable is now the field $\psi(x,t)$ and its derivatives, rather than the position of a particle or the velocity of a particle.

3.2.2 Generalization to General Field

Now, we would like to apply Lagrange formalism to a general field $\psi(\vec{x},t)$. We can define the Lagrangian density for a general field $\psi(\vec{x},t)$ as:

Definition 3.3: Lagrangian Density of a Field

The **Lagrangian density** of a field $\psi(\vec{x},t)$ is defined such that the Lagrangian is given by the integral of the Lagrangian density over the whole space:

$$L[\psi,\partial_t\psi,\partial_i\psi]=\int d\vec{x}\,\mathcal{L}(\psi,\partial_\mu\psi) \tag{3.31}$$

where for $\mu=0,1,2,3$, $x^{0,1,2,3}=(t,\vec{x})$

$$\partial_\mu\psi:=\frac{\partial\psi}{\partial x^\mu}=\left(\frac{\partial\psi}{\partial t},\frac{\partial\psi}{\partial x^1},\frac{\partial\psi}{\partial x^2},\frac{\partial\psi}{\partial x^3}\right) \tag{3.32}$$

Definition 3.4: Action of a Field

The **action** of a field $\psi(\vec{x},t)$ is defined as the integral of the Lagrangian over time:

$$S[\psi]=\int dt\,L[\psi,\partial_t\psi,\partial_i\psi]=\int dt\int d^3\vec{x}\,\mathcal{L}(\psi,\partial_\mu\psi) \tag{3.33}$$

where $\mathcal{L}$ is the Lagrangian density of the field $\psi(\vec{x},t)$.

and the stationary point of the action is given by:

$$\delta S[\mathcal{L}]=0 \implies \int dt\int d^3\vec{x}\,\frac{\partial\mathcal{L}}{\partial\psi}\delta\psi+\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\delta(\partial_\mu\psi)=0 \tag{3.34}$$

where $\delta\psi=0$ at the boundary of the integration region. The second term can be integrated by parts:

$$\int dt \int d^3\vec{x}\, \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \delta(\partial_\mu \psi) = \int dt \int d^3\vec{x}\, \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)}\, \partial_\mu(\delta\psi) \tag{3.35}$$

$$= \int dt \int d^3\vec{x}\, \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \delta\psi \right) - \int dt \int d^3\vec{x}\, \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) \delta\psi \tag{3.36}$$

Since $\delta\psi = 0$ at the boundary of the integration region, the first term vanishes:

$$= - \int dt \int d^3\vec{x}\, \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) \delta\psi \tag{3.37}$$

Thus, the stationary point of the action is given by:

$$\int dt \int d^3\vec{x} \left(\frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) \right) \delta\psi = 0 \tag{3.38}$$

Since time and space are arbitrary, the integrand must be zero:

$$\frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) = 0 \iff \partial_t \left(\frac{\partial \mathcal{L}}{\partial(\partial_t \psi)} \right) + \sum_i \partial_i \left(\frac{\partial \mathcal{L}}{\partial(\partial_i \psi)} \right) = \frac{\partial \mathcal{L}}{\partial \psi} \tag{3.39}$$

This is the **Euler-Lagrange equation** for a field $\psi(\vec{x}, t)$:

Theorem 3.2.1 Euler-Lagrange Equation for a Field

The **Euler-Lagrange equation** for a field $\psi(\vec{x}, t)$ is given by:

$$\frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) = 0 \tag{3.40}$$

where $\mathcal{L}$ is the Lagrangian density of the field $\psi(\vec{x}, t)$.

Additionally, we define the **canonical conjugate field** $\pi(\vec{x}, t)$ as:

Definition 3.5: Canonical Conjugate Field

The **canonical conjugate field** or **canonical momentum density** $\pi(\vec{x}, t)$ is defined as the derivative of the Lagrangian density with respect to the time derivative of the field ψ:

$$\pi(\vec{x}, t) := \frac{\partial \mathcal{L}}{\partial(\partial_t \psi(\vec{x}, t))} \tag{3.41}$$

3.3 Hamilton Formalism

3.3.1 Legendre Transformation

Now, remember that we obtained the **Hamiltonian** H of the system through the **Legendre transformation** of the Lagrangian L:

$$H(q_i, p_i) = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i) \tag{3.42}$$

where i represents each degree of freedom of the system.

In the field, the degree of freedom is infinite - each indexed by the spatial position $\vec{x}$, then,

Definition 3.6: Hamiltonian of a Field

The **Hamiltonian** of a field $\psi(\vec{x}, t)$ (whose canonical conjugate field is $\pi(\vec{x}, t)$) is defined as

$$H[\psi, \partial_i \psi, \pi] = \int d^3\vec{x} \left[\pi \cdot \partial_t \psi\right] - L = \int d^3\vec{x} \left[\pi \cdot \partial_t \psi - \mathcal{L}\right] \tag{3.43}$$

thus, we should define the **Hamiltonian density** $\mathcal{H}$ as:

Definition 3.7: Hamiltonian Density

Hamiltonian density $\mathcal{H}$ is defined as the Hamiltonian per unit volume of the field:

$$\mathcal{H}(\psi, \partial_i \psi, \pi, t) = \pi \cdot \partial_t \psi - \mathcal{L}(\psi, \partial_i \psi, \partial_t \psi, t) \tag{3.44}$$

which satisfies:

$$\int d^3\vec{x}\, \mathcal{H}(\psi, \partial_i \psi, \pi, t) = H[\psi, \partial_i \psi, \pi, t] \tag{3.45}$$

which makes the Hamiltonian H a functional of the field ψ, its spatial derivatives $\partial_i \psi$, and the conjugate field π.

Note:

Similarly to the discrete case, we can write the Lagrangian densty $\mathcal{L}$ in terms of the Hamiltonian density $\mathcal{H}$:

$$\mathcal{L}[\psi, \partial_i \psi, \pi, t] = \pi \cdot \partial_t \psi - \mathcal{H}[\psi, \partial_i \psi, \pi, t] \tag{3.46}$$

where

$$\partial_t \psi = \frac{\partial \mathcal{H}[\psi, \partial_i \psi, \pi, t]}{\partial \pi} \tag{3.47}$$

3.3.2 Canonical Equations of Fields

We are also interested in the **canonical equations** of fields, which are derived from the variational principle:

$$\delta S = 0 \iff \delta \int dt \int d^3\vec{x}\mathcal{L} = \delta \int dt \int d^3\vec{x}\,\pi \cdot \partial_t \psi - \mathcal{H} \tag{3.48}$$

The variation on this integral can be expanded as follows:

$$\begin{aligned}
\delta S &= \int dt \int d^3\vec{x}\, \delta\pi\, \partial_t \psi + \pi \delta(\partial_t \psi) - \delta\mathcal{H} && (3.49)\\
&= \int dt \int d^3\vec{x}\, \delta\pi\, \partial_t \psi + \pi\, \partial_t(\delta\psi) - \frac{\partial\mathcal{H}}{\partial\psi}\delta\psi - \frac{\partial\mathcal{H}}{\partial\pi}\delta\pi - \frac{\partial\mathcal{H}}{\partial(\partial_i\psi)}\delta(\partial_i\psi) && (3.50)\\
&= \int dt \int d^3\vec{x} \left(\partial_t\psi - \frac{\partial\mathcal{H}}{\partial\pi}\right)\delta\pi + \left(\pi\, \partial_t(\delta\psi) - \frac{\partial\mathcal{H}}{\partial\psi}\delta\psi - \frac{\partial\mathcal{H}}{\partial(\partial_i\psi)}\, \partial_i(\delta\psi)\right) && (3.51)
\end{aligned}$$

the $\delta\psi$ term can be integrated by parts:

$$\begin{aligned}
&\int dt \int d^3\vec{x}\, \pi\, \partial_t(\delta\psi) - \frac{\partial\mathcal{H}}{\partial(\partial_i\psi)}\, \partial_i(\delta\psi) && (3.52)\\
&= \int dt \int d^3\vec{x}\ \partial_t(\pi\delta\psi) - \int dt \int d^3\vec{x}\ \partial_t\pi \cdot \delta\psi && (3.53)\\
&\quad - \int dt \int d^3\vec{x}\ \partial_i\left(\frac{\partial\mathcal{H}}{\partial(\partial_i\psi)}\delta\psi\right) + \int dt \int d^3\vec{x}\ \partial_i\left(\frac{\partial\mathcal{H}}{\partial(\partial_i\psi)}\right)\delta\psi\\
&= \int d^3\vec{x}\, [\pi\delta\psi]_{\text{boundary}} - \int dt \int d^3\vec{x}\ \partial_t\pi\delta\psi && (3.54)\\
&\quad - \int dt \int_{\text{boundary}} dS\, \frac{\partial\mathcal{H}}{\partial(\partial_i\psi)}\delta\psi + \int dt \int d^3\vec{x}\ \partial_i\left(\frac{\partial\mathcal{H}}{\partial(\partial_i\psi)}\right)\delta\psi
\end{aligned}$$

the boundary terms vanish, so we end up with:

$$\int dt \int d^3\vec{x}\, \pi\, \partial_t(\delta\psi) - \frac{\partial\mathcal{H}}{\partial(\partial_i\psi)}\, \partial_i(\delta\psi) = -\int dt \int d^3\vec{x}\ \partial_t\pi\delta\psi + \int dt \int d^3\vec{x}\ \partial_i\left(\frac{\partial\mathcal{H}}{\partial(\partial_i\psi)}\right)\delta\psi \tag{3.55}$$

$$= -\int dt \int d^3\vec{x} \left[\partial_t\pi - \nabla \cdot \left(\frac{\partial\mathcal{H}}{\partial(\nabla\psi)}\right)\right]\delta\psi \tag{3.56}$$

and thus the variation of the action becomes:

$$\delta S = \int dt \int d^3\vec{x} \left(\partial_t\psi - \frac{\partial\mathcal{H}}{\partial\pi}\right)\delta\pi - \left(\partial_t\pi - \nabla \cdot \left(\frac{\partial\mathcal{H}}{\partial(\nabla\psi)}\right)\right)\delta\psi \tag{3.57}$$

for the action to be stationary, the integrand must vanish:

$$\begin{cases} \partial_t \psi - \dfrac{\partial \mathcal{H}}{\partial \pi} & = 0 \\ \partial_t \pi + \dfrac{\partial \mathcal{H}}{\partial \psi} - \partial_i \left(\dfrac{\partial \mathcal{H}}{\partial(\partial_i \psi)} \right) & = 0 \end{cases} \iff \begin{cases} \dfrac{\partial \psi(\vec{x},t)}{\partial t} & = \dfrac{\partial \mathcal{H}[\psi,\pi]}{\partial \pi} \\ \dfrac{\partial \pi(\vec{x},t)}{\partial t} & = -\dfrac{\partial \mathcal{H}[\psi,\pi]}{\partial \psi} + \partial_i \left(\dfrac{\partial \mathcal{H}}{\partial(\partial_i \psi)} \right) \end{cases} \tag{3.58}$$

Thus we have the **canonical equations of fields**:

Theorem 3.3.1 Canonical Equations of Fields

The variational principle in Hamilton formalism leads to the **canonical equations of fields**:

$$\frac{\partial \psi(\vec{x},t)}{\partial t} = \frac{\partial \mathcal{H}(\psi, \partial_i \psi, \pi)}{\partial \pi}, \qquad \frac{\partial \pi(\vec{x},t)}{\partial t} = -\frac{\partial \mathcal{H}(\psi, \partial_i \psi, \pi)}{\partial \psi} + \partial_i \left(\frac{\partial \mathcal{H}(\psi, \partial_i \psi, \pi)}{\partial(\partial_i \psi)} \right) \tag{3.59}$$

Corollary 3.3.1 Canonical Equations of Fields using H

Using Theorem 5.1.1, the **canonical equations of fields** can be re-written using the Hamiltonian H instead of the Hamiltonian density $\mathcal{H}$:

$$\frac{\partial \psi(\vec{x},t)}{\partial t} = \frac{\delta H}{\delta \pi}, \qquad \frac{\partial \pi(\vec{x},t)}{\partial t} = -\frac{\delta H}{\delta \psi} \tag{3.60}$$

3.3.3 Poisson Bracket

In the discrete case, the time evolution of a physical quantity $X(q_i, p_i, t)$, $\dot{X}$ can be written as:

$$\dot{X} = \frac{dX}{dt} = \frac{\partial X}{\partial t} + \sum_i \left(\frac{\partial X}{\partial q_i} \dot{q}_i + \frac{\partial X}{\partial p_i} \dot{p}_i \right) = \frac{\partial X}{\partial t} + \{X, H\} \tag{3.61}$$

In the continous case, a physical quantity X should be an integral of "density" $\tilde{X}$ over some volume:

$$X = \int_V d^3\vec{x}' \, \tilde{X}(\vec{x}', t) \tag{3.62}$$

and assume that $\tilde{X}$ is a function of the field $\psi(\vec{x}, t)$ and its conjugate field $\pi(\vec{x}, t)$ (which makes X a functional of the fields):

$$\implies \quad X[\psi, \partial_i \psi, \pi, t] = \int_V d^3\vec{x}' \, \tilde{X}(\psi(\vec{x}', t), \pi(\vec{x}', t), t) \tag{3.63}$$

Then the time evolution of X can be written as:

$$\frac{dX[\psi,\partial_i\psi,\pi,t]}{dt} = \frac{d}{dt}\int_V d^3\vec{x}'\,\tilde{X}(\psi(\vec{x}',t),\pi(\vec{x}',t),t) = \int_V d^3\vec{x}'\,\frac{d\tilde{X}(\psi(\vec{x}',t),\pi(\vec{x}',t),t)}{dt} \tag{3.64}$$

$$= \int_V d^3\vec{x}'\,\frac{\partial\tilde{X}(\psi,\pi,t)}{\partial t} + \frac{\partial\tilde{X}(\psi,\pi,t)}{\partial\psi}\,\partial_t\psi + \frac{\partial\tilde{X}(\psi,\pi,t)}{\partial\pi}\,\partial_t\pi \tag{3.65}$$

$$= \int_V d^3\vec{x}'\,\frac{\partial\tilde{X}}{\partial t} + \int_V d^3\vec{x}'\left(\frac{\partial\tilde{X}}{\partial\psi}\frac{\partial\mathcal{H}}{\partial\pi} - \frac{\partial\tilde{X}}{\partial\pi}\frac{\partial\mathcal{H}}{\partial\psi}\right) \tag{3.66}$$

Using Theorem 5.1.1, the partial derivatives can be replaced with functional derivatives:

$$\frac{\partial\tilde{X}}{\partial t} = \frac{\delta X}{\delta t}, \quad \frac{\partial\mathcal{H}}{\partial\psi} = \frac{\delta H}{\delta\psi}, \quad \frac{\partial\mathcal{H}}{\partial\pi} = \frac{\delta H}{\delta\pi} \tag{3.67}$$

Thus, we can write the time evolution of X as:

$$\frac{dX[\psi,\partial_i\psi,\pi,t]}{dt} = \frac{\partial X[\psi,\partial_i\psi,\pi,t]}{\partial t} + \int d^3\vec{x}'\left(\frac{\delta X}{\delta\psi}\frac{\delta H}{\delta\pi} - \frac{\delta X}{\delta\pi}\frac{\delta H}{\delta\psi}\right) \tag{3.68}$$

If we were to write the coordinates explicitly,

$$\frac{dX[\psi,\partial_i\psi,\pi,t]}{dt} = \frac{\partial X}{\partial t} + \int d^3\vec{x}'\left(\frac{\delta X[\psi,\partial_i\psi,\pi,t]}{\delta\psi(\vec{x}',t)}\frac{\delta H[\psi,\partial_i\psi,\pi,t]}{\delta\pi(\vec{x}',t)} - \frac{X[\psi,\partial_i\psi,\pi,t]}{\delta\pi(\vec{x}',t)}\frac{\delta H[\psi,\partial_i\psi,\pi,t]}{\delta\psi(\vec{x}',t)}\right) \tag{3.69}$$

Comparing with the discrete case, we can define the **Poisson bracket** of two physical quantities X and Y as:

Definition 3.8: Poisson Bracket of a Field

The **Poisson bracket** of two physical quantities $X[\psi,\partial_i\psi,\pi,t]$ and $Y[\psi,\partial_i\psi,\pi,t]$ is defined as:

$$\{X,\,Y\} = \int d^3\vec{x}'\left(\frac{\delta X}{\delta\psi(\vec{x}',t)}\frac{\delta Y}{\delta\pi(\vec{x}',t)} - \frac{\delta X}{\delta\pi(\vec{x}',t)}\frac{\delta Y}{\delta\psi(\vec{x}',t)}\right) \tag{3.70}$$

Theorem 3.3.2 Time Evolution of a Physical Quantity

The time evolution of a physical quantity $X[\psi,\partial_i\psi,\pi,t]$ can be written as:

$$\frac{dX}{dt} = \frac{\partial X}{\partial t} + \{X,\,H\} \tag{3.71}$$

Theorem 3.3.3 Poisson Brackets of Fields

For ψ and its conjugate field π, the Poisson bracket satisfies the following properties:

$$\{\psi(\vec{x},t),\, \pi(\vec{x}',t)\} = \delta^3(\vec{x}-\vec{x}') \tag{3.72}$$

$$\{\psi(\vec{x},t),\, \psi(\vec{x}',t)\} = 0 \tag{3.73}$$

$$\{\pi(\vec{x},t),\, \pi(\vec{x}',t)\} = 0 \tag{3.74}$$

Chapter 4 Quantum Mechanics

4.1 Basic Probability Theory

4.1.1 Discrete Probability

A fundamental interpretation of quantum mechanics is based on the principles of probability theory. Thus, let us review some basics of probability theory.

Let Ω be the collection of all possible outcomes of an experiment (called the **sample space**). Subsets of Ω are called events. Then the probability $\mathbb{P}$ is a rule to assign each event a number between 0 and 1, called the **probability** of an event, in a REASONABLE way.

Such "probability" must satisfy the following conditions:

Principle 4.1: Requirements for Probability

1. $\mathbb{P}(\Omega) = 1$ (the probability of any outcome is 1),
2. $0 \leq \mathbb{P}(A) \leq 1$ for all events A (the probability of an event is non-negative),
3. If $A_1, A_2, ...$ are disjoint events (mutually exclusive), then

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i). \tag{4.1}$$

Sometimes, it is useful to know the expectation value of dice when you need to make a bet, let's say. It can be defined as a weighted average:

Definition 4.1: Expectation Value

The expectation value (or expected value) of a random variable X is defined as:

$$\mathbb{E}\left[X\right] = \langle X \rangle = \sum_{i=1}^{n} x_i \cdot \mathbb{P}(x_i) \tag{4.2}$$

where x_i are the possible outcomes and $\mathbb{P}(x_i)$ is the probability of observing the outcome x_i

Example 4.1.1 (A fair dice)

Consider a fair dice - which means each face is equally likely to appear on top every time the dice is rolled. The sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$. The probability of each face appearing is $\mathbb{P}(i) = \frac{1}{6}$ for $i = 1, 2, 3, 4, 5, 6$. The expectation value of the dice is:

$$\langle \omega \rangle = \sum_{i=1}^{6} i \cdot \mathbb{P}(i) = \frac{21}{6} = 3.5. \tag{4.3}$$

This means that if you roll the dice many times, the average outcome will be around 3.5.

4.1.2 Continous Probability

Sometimes, we deal with continous data/ random variables, such as height, weight, or temperature. For such cases, we notice that the probability of observing a single value is zero, since there are infinitely many possible values. We need to consider a range of values. For example, we may not be able to find a person with exactly 1.74012959278 ... m tall, but we can find people who are between 1.74 m and 1.75 m tall. So, instead of probabilities for single value, we assign a **probability density function** (PDF) to a certain range of values:

Definition 4.2: Probability Density Function

A probability density function (PDF) $P(x)$ is a function whose integral describes the likelihood of a random variable X taking on a value in a certain range. The probability of X being in the interval $[a, b]$ is given by:

$$\mathbb{P}(a \leq X \leq b) = \int_a^b dx\, P(x). \tag{4.4}$$

Note:

The PDF must satisfy the following conditions:

1. $P(x) \geq 0$ for all x (the PDF is non-negative),
2. $\int_{-\infty}^{\infty} dx P(x) = 1$ (the total probability is 1).

which is very much an extension of the discrete case.

4.2 Braket Notation

4.2.1 Discrete Probability and Vectors

For a discrete distribution, we can also use vectors to represent expectation values. Using the dice as an example again, let us study how we can use vectors to represent probabilites. First, define a vector $|\omega\rangle$ as follows (maybe a not familiar notation, but just follow me for now):

$$|\omega\rangle = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_6 \end{pmatrix} \tag{4.5}$$

Just assume that ω_i is a number between 0 and 1. Now, consider what happens when we calculate the norm (or length) of this vector:

$$\||\omega\rangle\| = \sqrt{\sum_{i=1}^{6} \omega_i^2} = \sqrt{\omega_1^2 + \omega_2^2 + \ldots + \omega_6^2} \tag{4.6}$$

Now, what happens if $\omega_i = \sqrt{\mathbb{P}(i)}$, the probability of observing the outcome i? Then, we have:

$$\||\omega\rangle\| = \sqrt{\sum_{i=1}^{6} \mathbb{P}(i)} = \sqrt{1} = 1. \tag{4.7}$$

This means that the vector $|\omega\rangle$ is a **normalized vector** or **unit vector**, which is a vector whose length is 1.

This vector notation is very useful because when the dice is not fair, we can just change the values of ω_i to represent the probabilities of each outcome, while maintaining the normalization condition $\||\omega\rangle\| = 1$:

Example 4.2.1 (Not so fair dice)

A dice where the probability of rolling a 6 is twice as likely as rolling any other number can be represented by a vector:

$$|\omega\rangle = \frac{1}{\sqrt{7}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \sqrt{2} \end{pmatrix} \tag{4.8}$$

And notice that this vector contains the information about probability distribution of the dice: how likely each outcome is to occur, but not the outcome itself.

4.2.2 Matrixs and Expectation Values

Thus, to truly apply linear algebra to probabilistic theory, we need to define another quantity, a **matrix**:

Definition 4.3: Matrix

A matrix is a 2 dimensional array of numbers:

$$\hat{M} = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mn} \end{pmatrix} \tag{4.9}$$

We can specify each element (or entry) of the matrix using two indices M_{ij}, where i is the row index and j is the column index. If a matrix has m rows and n columns, we say that the matrix is of size $m \times n$. If $i = j$, we call the element M_{ii} a **diagonal element** of the matrix.

and we define the multiplication of a vector and a matrix as follows:

$$|\omega\rangle = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{pmatrix}, \quad \hat{M} = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mn} \end{pmatrix} \tag{4.10}$$

$$\hat{M}|\omega\rangle = \sum_{j=1}^{n} \begin{pmatrix} M_{1j}\omega_j \\ M_{2j}\omega_j \\ \vdots \\ M_{mj}\omega_j \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} M_{1j}\omega_j \\ \sum_{j=1}^{n} M_{2j}\omega_j \\ \vdots \\ \sum_{j=1}^{n} M_{mj}\omega_j \end{pmatrix} \tag{4.11}$$

(we will use Einstein convention, but still keep the summation as a reminder for now). and if the matrix is an $1 \times n$ matrix, it is equivalent to a row vector, which we denote as $\langle\psi|$ (the transpose of a column vector $|\psi\rangle$. for now):

$$\langle\psi| = \begin{pmatrix} \psi_1 & \psi_2 & \cdots & \psi_n \end{pmatrix} = (|\psi\rangle)^T \tag{4.12}$$

following the matrix multiplication rules, row vector and column vector, we can define the **inner product** as follows:

Definition 4.4: Inner product

The inner product of a row vector $\langle\psi|$ and a column vector $|\omega\rangle$ is defined as:

$$\langle\psi|\omega\rangle = \begin{pmatrix} \psi_1 & \psi_2 & \cdots & \psi_n \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{pmatrix} = \sum_{i=1}^{n} \psi_i \cdot \omega_i \tag{4.13}$$

which is a linear operation for both $\langle\psi|$ and $|\omega\rangle$ (we call it bi-linear):

$$\langle\psi|(a|\omega_1\rangle + b|\omega_2\rangle) = a\langle\psi|\omega_1\rangle + b\langle\psi|\omega_2\rangle \tag{4.14}$$

In physics, we often call the row vector $\langle\psi|$ a **bra** and the column vector $|\omega\rangle$ a **ket**, so we will be calling them as such from now on.

Now, consider what happens when we replace $|\omega\rangle$ with $\hat{M}|\omega\rangle$:

$$\langle\psi\,|\,(\hat{M}\,|\,\omega\rangle) =: \langle\psi\,|\,\hat{M}\,|\,\omega\rangle = \psi_i \cdot M_{ij} \cdot \omega_j = M_{ij}\psi_i\omega_j \tag{4.15}$$

and further, let's say that $|\psi\rangle = |\omega\rangle$, then we have:

$$\langle\psi\,|\,\hat{M}\,|\,\omega\rangle = \sum_{i,j=1}^{n} M_{ij}\omega_i\omega_j \tag{4.16}$$

now compare this with the definition of expectation value:

$$\langle X\rangle = \sum_{i=1}^{n} x_i \cdot \mathbb{P}(x_i) \tag{4.17}$$

Here, let us define a quantity called **Kronecker delta** δ_{ij} as follows:

Definition 4.5: Kronecker Delta

The Kronecker delta δ_{ij} is defined as:

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{4.18}$$

we see that if $M_{ij} = x_i\delta_{ij}$, we can see that the expression in Eq.(4.16) is equivalent to the expectation value of a random variable X because $\mathbb{P}(x_i) = \omega_i^2$.

This means that our matrix is a diagonal matrix, which only has diagonal elements, and the off-diagonal elements are all zero:

$$\hat{M} = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{pmatrix} \tag{4.19}$$

notice how this matrix contains all the outcomes x_i of the random variable X on the diagonal, and the probabilities $\mathbb{P}(x_i) = \omega_i^2$ are encoded in the vector $|\omega\rangle$. Thus given a vector $|\omega\rangle$ in which all probabilities are encoded, and a matrix $\hat{M}$ in which all outcomes are encoded, we can calculate the expectation value of a random variable X as follows:

$$\langle X\rangle = \langle\omega\,|\,\hat{M}\,|\,\omega\rangle = \sum_{i=1}^{n} M_{ii}\omega_i^2 \tag{4.20}$$

4.2.3 Specific Outcome: Eigenvectors and Eigenvalues

Now, let us come back to the example of a fair dice:

$$|\omega\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\1\\1\\1\\1\end{pmatrix}, \quad \hat{X} = \begin{pmatrix}1&0&0&0&0&0\\0&2&0&0&0&0\\0&0&3&0&0&0\\0&0&0&4&0&0\\0&0&0&0&5&0\\0&0&0&0&0&6\end{pmatrix} \tag{4.21}$$

and again, let me remind you that $|\omega\rangle$ contains the probabilities of each outcome, and $\hat{X}$ contains the outcomes themselves.

Now, how can we calculate the probability of observing a specific outcome, say 3? To do that, we have to define the **basis vectors**:

$$|1\rangle = \begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}, \quad |2\rangle = \begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix}, \cdots, |6\rangle = \begin{pmatrix}0\\0\\0\\0\\0\\1\end{pmatrix} \tag{4.22}$$

just like the 3-dimensional Cartesian coordinates, these basis vectors are unit vectors and orthogonal to one another:

$$\langle i|j\rangle = \delta_{ij} \tag{4.23}$$

Note:
A set of basis vectors satisfying $\langle i|j\rangle = \delta_{ij}$ is called an **orthonormal basis**.

(refer to Eq. (4.12) for $\langle i|$) This means that we can express the vector $|\omega\rangle$ in terms of these basis vectors:

$$|\omega\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\1\\1\\1\\1\end{pmatrix} = \sum_{j=1}^{6}\omega_j|j\rangle \tag{4.24}$$

where $\omega_i = \sqrt{\mathbb{P}(x_i)} = \frac{1}{\sqrt{6}}$ for $i = 1, 2, 3, 4, 5, 6$. By taking the inner product with $\langle i|$, we get:

$$\langle i|\omega\rangle = \langle i|\sum_{j=1}^{6}\omega_j|j\rangle = \sum_{j=1}^{6}\omega_j\langle i|j\rangle = \sum_{j=1}^{6}\omega_j\delta_{ij} = \omega_i = \sqrt{\mathbb{P}(x_i)} \tag{4.25}$$

thus, we can write the $|\omega\rangle$ using $\langle i|\omega\rangle$ as follows:

$$|\omega\rangle = \sum_{i=1}^{6} |i\rangle\langle i|\omega\rangle \tag{4.26}$$

if we define an **identity matrix** $\hat{\mathbf{I}}$ as follows:

Definition 4.6: Identity Matrix

The identity matrix $\hat{\mathbf{I}}$ is a square matrix with ones on the diagonal and zeros elsewhere:

$$\hat{\mathbf{I}} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \tag{4.27}$$

it elements can be written simply as $\hat{\mathbf{I}}_{ij} = \delta_{ij}$, where i and j are the row and column indices, respectively.

we can re-write Eq.(4.26) as:

$$\hat{\mathbf{I}}|\omega\rangle = \sum_{i=1}^{6} |i\rangle\langle i|\omega\rangle \implies \left(\hat{\mathbf{I}} - \sum_{i=1}^{6} |i\rangle\langle i|\right)|\omega\rangle = 0 \tag{4.28}$$

since all ω_i are non-zero, the term inside bracket must be zero:

$$\hat{\mathbf{I}} - \sum_{i=1}^{6} |i\rangle\langle i| = 0 \iff \sum_{i=1}^{6} |i\rangle\langle i| = \hat{\mathbf{I}} \tag{4.29}$$

This is a fundamental result in quantum mechanics, and holds in the case where the number of outcome is (countably) infinite as well:

Theorem 4.2.1 Completeness Relation

Given an orthonormal basis $\{|i\rangle\}$, the completeness relation holds:

$$\sum_{i} |i\rangle\langle i| = \hat{\mathbf{I}} \tag{4.30}$$

Another way to think about this is to think of a so-called **projection matrix**:

$$\hat{\mathcal{P}}_i = |i\rangle\langle i| \tag{4.31}$$

the reason for its name is made clear when we consider the action of this operator on a vector $|\omega\rangle$:

$$\hat{\mathbf{P}}_i|\omega\rangle = |i\rangle\langle i|\omega\rangle = \langle i|\omega\rangle|i\rangle \tag{4.32}$$

note that $\langle i|\omega\rangle$ is a scalar, so it can be moved around. It "projects" the vector $|\omega\rangle$ onto the vector $|i\rangle$, which is the basis vector corresponding to the outcome i. Or in other words, it isolates the component of $|i\rangle$ from $|\omega\rangle$. Using this projection operator, the completeness relation can be written as:

$$\hat{\mathbf{I}} = \sum_i \hat{\mathcal{P}}_i \tag{4.33}$$

This matrix can be written explicitly too. For example, in the case of the fair dice, we have:

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{4.34}$$

then the projection matrix for the outcome 1, $\hat{\mathcal{P}}_1$, is given by:

$$\hat{\mathcal{P}}_1 = |1\rangle\langle 1| = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.35}$$

from which we can immidiately see that the sum of all projection matrices gives the identity matrix.

Note:

Here, we just applied normal matrix multiplication rules. The product between a row vector and column vector is sometimes called the **dyadic product**. A more general products between vectors are **tensor products**, but in this particular context, they give the same result.

Now, also note that by the matrix multiplication rules,

$$\hat{X}|i\rangle = x_i|i\rangle \quad \text{for } i = 1, 2, 3, 4, 5, 6 \tag{4.36}$$

notice that on LHS, we have a matrix $\hat{X}$ acting on a vector $|i\rangle$, and on RHS, we have a scalar x_i multiplying the same vector $|i\rangle$. This means that $|i\rangle$ and $\hat{X}$ are special set of vectors and a matrix. We call these special vectors **eigenvectors** and the scalars **eigenvalues**:

Definition 4.7: Eigenvectors and Eigenvalues

When there is a matrix $\hat{M}$ and a vector $|\psi\rangle$ such that:

$$\hat{M}|\psi\rangle = \lambda|\psi\rangle \tag{4.37}$$

where λ is a scalar, we say that $|\psi\rangle$ is an **eigenvector** of the matrix $\hat{M}$ and λ is the corresponding **eigenvalue**.

and importantly, the outcome of the measurement of the random variable X is given as the eigenvalue of the matrix $\hat{X}$, x_i.

Formally, if we put $\langle i|$ to the right of both sides of the equation, we have:

$$\hat{X}|i\rangle\langle i| = x_i|i\rangle\langle i| = x_i\hat{\mathcal{P}}_i \tag{4.38}$$

by taking the summation over all i (notice that $\hat{X}$ does not depend on i), we have:

$$\sum_{i=1}^{6}\hat{X}|i\rangle\langle i| = \hat{X}\sum_{i=1}^{6}|i\rangle\langle i| = \sum_{i=1}^{6}x_i\hat{\mathcal{P}}_i = \sum_{i=1}^{6}x_i|i\rangle\langle i| \tag{4.39}$$

An important note is that we have obtained eigenvalues from a diagonal matrix. However, it is also possible to obtain diagonal matrix from eigenvalues (a process called **diagonalization**):

4.2.4 Continous Case

It may be intuitive to compare expectation value in the discrete case with the continous case:

$$\text{Discrete}: \ \langle X\rangle = \sum_i x_i \cdot \langle i|\omega\rangle^2 \tag{4.40}$$

$$\text{Continous:} \ \langle X\rangle = \int_{-\infty}^{\infty} dx\, x\cdot P(x) \tag{4.41}$$

A naive transition may be to consider that the PDF may also be a square of ”something”, let’s say $\psi(x)$, such that $P(x) = [\psi(x)]^2$, as well as $\sum_i \to \int_{-\infty}^{\infty} dx$. For simplicity, let us simply omit the integraiton bound. Since $\langle i|\omega\rangle$ was the probability of observing the outcome i, we can think of $\psi(x)$ as the probability of observing a value x:

$$\int_{-\infty}^{\infty} dx = \int dx, \quad \psi(x) := \langle x|\psi\rangle \tag{4.42}$$

and the completeness relation becomes

$$|\omega\rangle = \sum_i |i\rangle\langle i|\omega\rangle \to |\psi\rangle = \int dx\,|x\rangle\langle x|\psi\rangle \implies \int dx\,|x\rangle\langle x| = \hat{\mathbf{I}} \tag{4.43}$$

Let us also look at the orthonormality condition of the basis vectors. Remember in the discrete case, we had $\langle i|j\rangle = \delta_{ij}$, which is the Kronecker delta. This comes from the completeness relation applied onto the basis vectors:

$$|j\rangle = \sum_i |i\rangle\langle i|j\rangle \implies \langle i|j\rangle = \delta_{ij} \tag{4.44}$$

and similarly, we should think that for the continous eigenvector $|x\rangle$,

$$|x\rangle = \int dx' |x'\rangle\langle x'|x\rangle \tag{4.45}$$

if we compare this with the definition of **Dirac delta function**(refer to Eq. (5.34)), we see that

$$\langle x'|x\rangle = \delta(x' - x) \tag{4.46}$$

This is essentially the same mathematics used in the quantum mechanics, where we have three fundamental objects:

- A **ket** vector $|\psi\rangle$ containing "probabilities"
- A **matrix** $\hat{M}$ containing "outcomes"
- A **basis** that diagonalizes the matrix $\hat{M}$

4.3 Preparation for Quantum Mechanics

4.3.1 A More Formal Summary

In quantum mechanics, there is some physical requirements so that it is more than just a general probability theory. First of the is the existence of a **state vector** $|\psi\rangle$, which is a normalized vector in **Hilbert space** $\mathbf{H}$.

Note:

Explanation of Hilbert space is kind of complicated, but we have some perspectives to look at it from:

1. Linear Algebra PoV: A vector space equal to $\mathbb{C}^n$
2. Analysis PoV: A space of square integrable functions $L^2(\mathbb{R}^d)$

but essence is that it is a "nice" space where quantum mechanics can be formulated.

We have seen that a matrix and its eigenvalues can represent the "outcomes", but in quantum mechanics, we use a little more general object called an **linear operator**:

Definition 4.8: Linear Operator

A linear operator $\hat{A}$ acts on a vector to change it to another vector ($\hat{A} : \mathbf{H} \ni |\psi\rangle \mapsto \hat{A}|\psi\rangle \in \mathbf{H}$). It must satisfy linearity

$$\hat{A}(a|\psi_1\rangle + b|\psi_2\rangle) = a\hat{A}(|\psi_1\rangle) + b\hat{A}(|\psi_2\rangle) \tag{4.47}$$

for all vectors $|\psi_1\rangle, |\psi_2\rangle$ and scalars a, b.

whose **Hermitian conjugate** is defined as follows:

Definition 4.9: Hermitian Conjugate

The Hermitian conjugate (or adjoint) of a linear operator $\hat{A}$, denoted as $\hat{A}^\dagger$, is defined such that:

$$\langle\psi_1|\hat{A}\psi_2\rangle = \langle\hat{A}^\dagger\psi_1|\psi_2\rangle \tag{4.48}$$

for all vectors $|\psi_1\rangle, |\psi_2\rangle$. Or as a matrix, if

$$\hat{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{4.49}$$

then

$$\hat{A}^\dagger = \begin{pmatrix} a_{11}^* & a_{21}^* & \cdots & a_{m1}^* \\ a_{12}^* & a_{22}^* & \cdots & a_{m2}^* \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n}^* & a_{2n}^* & \cdots & a_{mn}^* \end{pmatrix} \tag{4.50}$$

where a_{ij}^* is the complex conjugate of a_{ij}.

Theorem 4.3.1 Property of Hermitian Conjugate

Given two linear operators $\hat{A}$ and $\hat{B}$, the following properties hold:

$$\left(\hat{A} + \hat{B}\right)^\dagger = \hat{A}^\dagger + \hat{B}^\dagger, \tag{4.51}$$

$$\left(a\hat{A}\right)^\dagger = a^*\hat{A}^\dagger, \tag{4.52}$$

$$\left(\hat{A}\hat{B}\right)^\dagger = \hat{B}^\dagger\hat{A}^\dagger, \tag{4.53}$$

$$\hat{\mathbf{I}}^\dagger = \hat{\mathbf{I}}, \tag{4.54}$$

$$\left(\hat{A}^\dagger\right)^\dagger = \hat{A}. \tag{4.55}$$

Definition 4.10: Inner Product

The inner product of two vectors $|\psi\rangle, |\phi\rangle \in \mathbf{H}$ is defined as:

$$\langle\phi|\psi\rangle = (|\phi\rangle)^\dagger|\psi\rangle = \begin{pmatrix}\phi_1^* & \phi_2^* & \cdots & \phi_n^*\end{pmatrix}\begin{pmatrix}\psi_1 \\ \psi_2 \\ \vdots \\ \psi_n\end{pmatrix} = \sum_i \phi_i^*\psi_i \tag{4.56}$$

where ϕ_i^* is the complex conjugate of ϕ_i.

for the operator to represent some physical quantity, its eigenvalues must be real numbers.

Definition 4.11: Hermitian Operator

A Hermitian operator $\hat{A}$ is a linear operator that satisfies:

$$\langle\psi_1|\hat{A}\psi_2\rangle = \langle\hat{A}\psi_1|\psi_2\rangle \tag{4.57}$$

where $|\hat{A}\psi_{1,2}\rangle = \hat{A}|\psi_{1,2}\rangle$. Or equivalently,

$$\langle a_i\,|\,\hat{A}\,|\,a_i\rangle = a_i \in \mathbb{R} \tag{4.58}$$

where a_i are the eigenvalues of the operator $\hat{A}$.

and a physical quantity represented by a Hermitian operator or vice versa is called an **observable**.

Note:
Sometimes you might see the word "**self-adjoint**" instead of "Hermitian", but they are essentially the same. In the context of this note and most of the physics literature, we don't distinguish them as the difference is incredibly technical. Refer to Sec. **??** for more details.

Another important type of operator is the **unitary operator**:

Definition 4.12: Unitary Operator

A unitary operator $\hat{U}$ is a linear operator that satisfies:

$$\hat{U}^{-1} = \hat{U}^\dagger \iff \hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{\mathbf{I}}, \tag{4.59}$$

where $\hat{U}^\dagger$ is the adjoint (or conjugate transpose) of the operator $\hat{U}$ and $\hat{U}^{-1}$ is the inverse of the operator $\hat{U}$. Unitary operators preserve the inner product, i.e., they preserve the norm of vectors.

4.3.2 Rules for Quantum Mechanics

Now, as always in physics, we need to find a way to connect the mathematics to the real world. One part of this is to assign a Hermitian operator to each observable, which is a physical quantity that can be measured.

Definition 4.13: Observable

An **observable** is a physical quantity that can be measured, such as position, momentum, or energy. In quantum mechanics, each observable is represented by a Hermitian operator. Its possible value is given by the eigenvalues of the operator, and corresponding vector is called the **eigenstate** of the observable.

Another part of it is to do exactly what we have seen in the previous section: interpret the inner products as probabilities.

Principle 4.2: State Vector

A **state vector** (or **state ket**) $|\psi\rangle$ is a normalized vector in the Hilbert space $\mathbf{H}$ that represents the state of a quantum system. You can think of it as the vector containing the probability distribution for all possible measurements of the system, as we have seen in the previous section.

Principle 4.3: Born's Probability Rule/ Interpretation

Assume that an observable $\hat{A}$ has eigenstates $|a\rangle$ with eigenvalues a. For a general state $|\psi\rangle \in \mathbf{H}$, the probability of measuring the observable $\hat{A}$ and obtaining the value a is given by

$$\mathbb{P}(a) = |\langle a|\psi\rangle|^2 = |\langle\psi|a\rangle|^2 . \tag{4.60}$$

Corollary 4.3.1 Normalization Condition

Any state vector must be normalized, meaning that the total probability of measuring any value of the observable must equal 1:

$$\||\psi\rangle\| = \sqrt{(\psi)\,\psi} = 1. \tag{4.61}$$

Principle 4.4: Collapse of the State

After we measure an observable $\hat{A}$ and obtain the value a, the state of the system collapses to the corresponding eigenstate $|a\rangle$ of $\hat{A}$:

$$|\psi\rangle \xrightarrow{\text{measure } a} |a\rangle. \tag{4.62}$$

Theorem 4.3.2 Rays

Two states $|\psi\rangle$ and $|\phi\rangle$ are physically indistinguishable if they differ only by a complex phase factor:

$$|\phi\rangle = e^{i\theta}|\psi\rangle. \tag{4.63}$$

In this case, we say that $|\psi\rangle$ and $|\phi\rangle$ represent the same physical state, or that they are **rays** in the Hilbert space.

4.4 Position and Momentum as Observables

4.4.1 Position Operator

As an observable, let us consider the position of a particle: $\hat{x}$. We can set up the position operator $\hat{x}$ in the Hilbert space $\mathbf{H}$ and its eigenstates $|x\rangle$, which satisfy the eigenvalue equation

$$\hat{x}|x\rangle = x|x\rangle. \tag{4.64}$$

As we have seen in continous case of Sec 4.2, the inner product $(x)\, y$ is defined as

$$\langle x|x'\rangle = \delta(x - x'), \tag{4.65}$$

and similarly to the discrete case, for a state $|\psi\rangle$, its probability density $P(x)$ is given by

$$P(x) = |\langle x|\psi\rangle|^2 = |\psi(x)|^2\,, \tag{4.66}$$

where we defined the **wavefunction**:

Definition 4.14: wavefunction

The **wavefunction** $\psi(x)$ of a quantum state $|\psi\rangle$ is a function that describes the probability amplitude of finding the particle at position x.

$$\psi(x) := \langle x|\psi\rangle \tag{4.67}$$

4.4.2 Momentum Operator

We can also consider the momentum of a particle p as an observable. Really, the same discussion applies to the momentum operator $\hat{p}$. It is defined in the Hilbert space $\mathbf{H}$ and has eigenstates $|p\rangle$ that satisfy the eigenvalue equation

$$\hat{p}|p\rangle = p|p\rangle. \tag{4.68}$$

the completeness relation reads

$$\int dp\, |p\rangle\langle p| = \hat{\mathbf{I}} \iff \langle p|p'\rangle = \delta(p-p') \tag{4.69}$$

and the wavefunction $\psi(p)$ in the momentum space is defined as

$$\psi(p) := \langle p|\psi\rangle \tag{4.70}$$

Definition 4.15: Wavefunction in Momentum Space

The wavefunction in momentum space $\psi(p)$ is a function that describes the probability amplitude of finding the particle with momentum p.

$$\psi(p) := \langle p|\psi\rangle \tag{4.71}$$

This is often called the **momentum representation** of the wavefunction.

4.4.3 Canonical Commutation Relation

We define a operation called the **commutator** of two operators $\hat{A}$ and $\hat{B}$ as follows:

Definition 4.16: Commutator

The commutator $\hat{A}\hat{B}$ of two operators $\hat{A}$ and $\hat{B}$ is defined as

$$\left[\hat{A},\, \hat{B}\right] := \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{4.72}$$

Note:

If the commutator of two operators is zero, they are said to **commute**. If the commutator is non-zero, they are said to **not commute**.

Theorem 4.4.1 Properties of Commutator

For operators $\hat{A}, \hat{B}, \hat{C} \in \mathbf{H}$ and constants $\alpha, \beta, \gamma \in \mathbb{C}$, the following properties hold:

- (Bi-)Linearity:

$$\left[\alpha\hat{A} + \beta\hat{B},\, \hat{C}\right] = \alpha\left[\hat{A},\, \hat{C}\right] + \beta\left[\hat{B},\, \hat{C}\right] \tag{4.73}$$

$$\left[\hat{A},\, \beta\hat{B} + \gamma\hat{C}\right] = \beta\left[\hat{A},\, \hat{B}\right] + \gamma\left[\hat{A},\, \hat{C}\right] \tag{4.74}$$

- Anti-symmetry:

$$\left[\hat{A},\, \hat{B}\right] = -\left[\hat{B},\, \hat{A}\right] \tag{4.75}$$

Now, remember the **Poisson Bracket** from classical mechanics, which is defined as

Definition 4.17: Poisson Bracket

The Poisson bracket of two functions f and g in phase space is defined as

$$\{f,\, g\} = \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q}, \tag{4.76}$$

where q_i and p_i are the generalized coordinates and momenta, respectively.

Theorem 4.4.2 Properties of Poisson Bracket

For functions f, g, the following properties hold:

- (Bi-)Linearity:

$$\{\alpha f + \beta g,\, h\} = \alpha\,\{f,\, h\} + \beta\,\{g,\, h\} \tag{4.77}$$
$$\{f,\, \alpha g + \beta h\} = \alpha\,\{f,\, g\} + \beta\,\{f,\, h\} \tag{4.78}$$

- Anti-symmetry:

$$\{f,\, g\} = -\,\{g,\, f\} \tag{4.79}$$

- Jacobi Identity:

$$\{f,\, \{g,\, h\}\} + \{g,\, \{h,\, f\}\} + \{h,\, \{f,\, g\}\} = 0 \tag{4.80}$$

- Coordinate and Momentum Relation:

$$\{q,\, p\} = 1, \quad \{q,\, q\} = \{p,\, p\} = 0 \tag{4.81}$$

These two objects have the same algebraic properties if we impose the canonical commutation relation onto the commutator of the position and momentum operators:

Principle 4.5: Canonical Commutation Relation (1D)

The commutator of position operator $\hat{x}$ and momentum operator $\hat{p}$ is defined to be

$$[\hat{x},\, \hat{p}] = i\hbar. \tag{4.82}$$

Theorem 4.4.3 Commutator as a derivative

Consider a function $f(x)$ that can be Taylor expanded:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \tag{4.83}$$

where $a_n = \frac{1}{n!}\frac{\partial f}{\partial x}(x_0)$ and x. Then, the commutator of the position operator $\hat{x}$ and the momentum operator $\hat{p}$ can be interpreted as a derivative:

$$[f(\hat{x}),\, \hat{p}] = i\hbar \frac{\partial f}{\partial x}(\hat{x}) \tag{4.84}$$

where $f(\hat{x})$ refers to the substitution of $\hat{x}$ into the function $f(x)$.

a: We have to prove that this holds for $f(x) = x^n$, from which our goal can be proven easily.

$$\begin{aligned}
[\hat{x}^n,\, \hat{p}] &= \hat{x}^{n-1}\hat{x}\hat{p} - \hat{p}\hat{x}^{n-1}\hat{x} && (4.85)\\
&= \hat{x}^{n-1}\,[\hat{x},\, \hat{p}] + \hat{x}^{n-1}\hat{p}\hat{x} - \hat{p}\hat{x}^{n-1}\hat{x} && (4.86)\\
&= \hat{x}^{n-1}\,[\hat{x},\, \hat{p}] + [\hat{x}^{n-1},\, \hat{p}]\,\hat{x} && (4.87)\\
&= i\hbar\hat{x}^{n-1} + [\hat{x}^{n-1},\, \hat{p}]\,\hat{x} && (4.88)\\
&= i\hbar\hat{x}^{n-1} + (i\hbar\hat{x}^{n-2} + [\hat{x}^{n-2},\, \hat{p}]\,\hat{x})\hat{x} && (4.89)\\
&= 2i\hbar\hat{x}^{n-1} + [\hat{x}^{n-2},\, \hat{p}]\,\hat{x} && (4.90)\\
&= ni\hbar\hat{x}^{n-1} + \underbrace{[\hat{x}^0,\, \hat{p}]}_{=0}\,\hat{x} = ni\hbar\hat{x}^{n-1} = i\hbar\frac{\partial(x^n)}{\partial x}(\hat{x}) && (4.91)
\end{aligned}$$

thus, if the function f is given by a Taylor series, we can write

$$[\hat{x}^n,\, \hat{p}] = i\hbar \frac{\partial f}{\partial x}(\hat{x}) \tag{4.92}$$

□

4.4.4 Heisenberg's Uncertainty Principle

Definition 4.18: Variance of an Observable

The variance of an observable $\hat{A}$ in a quantum state $|\psi\rangle$ is defined as

$$\Delta A^2 := \langle\psi|\,\hat{A}^2\,|\psi\rangle - \big(\langle\psi|\,\hat{A}\,|\psi\rangle\big)^2. \tag{4.93}$$

It measures the spread of the observable's values around its expectation value.

Theorem 4.4.4 Robertson Inequality

The Robertson inequality states that for any two observables $\hat{A}$ and $\hat{B}$, the variance of $\hat{A}$ and $\hat{B}$ satisfies

$$\Delta A^2 \Delta B^2 \geq \frac{1}{4}\left|\big[\hat{A},\, \hat{B}\big]\right|^2 \tag{4.94}$$

Corollary 4.4.1 Heisenberg's Uncertainty Principle

The Heisenberg uncertainty principle is a special case of the Robertson inequality, which states that for position $\hat{x}$ and momentum $\hat{p}$,

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \tag{4.95}$$

What this tells us is not a practical limit of measurements, but a fundamental property of quantum systems, as a result of the non-commutativity of the position and momentum operators.

This also means that if the commutator of two observables $\hat{A}$ and $\hat{B}$ is zero, they can be measured with any precision at the same time.

Definition 4.19: Simultaneous Observable

Two observables $\hat{A}$ and $\hat{B}$ are said to be **simultaneous observables**, if they commute:

$$\left[\hat{A}, \hat{B}\right] = 0. \tag{4.96}$$

4.4.5 Derivative on Vectors

It is a bit arbitrary, but let us consider the following operator:

$$\hat{U}_x(a) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-i\frac{\hat{p}}{\hbar}a\right)^n \tag{4.97}$$

This is indeed an operator (guaranteed by the completeness of Hilbert space). As a short-hand notation for the infinite sum, we can write it as an exponential:

$$\hat{U}_x(a) = e^{-i\frac{\hat{p}}{\hbar}a}. \tag{4.98}$$

Remember that momentum must be an Hermitian operator ($\hat{p}^\dagger = \hat{p}$), so we see that by taking the the Hermitian conjugate,

$$\hat{S}_p^\dagger(a) := \left(\hat{U}_x(a)\right)^\dagger = \sum_{n=0}^{\infty} \frac{1}{n!} \left(i\frac{\hat{p}}{\hbar}a\right)^n = \hat{S}_p(-a) \tag{4.99}$$

and as you might guess from the exponential form,

$$\begin{aligned} \hat{U}_x(a)\hat{U}_x(-a) &= \left(1 - i\frac{\hat{p}}{\hbar}a + \frac{1}{2!}\left(-i\frac{\hat{p}}{\hbar}a\right)^2 + \ldots\right)\left(1 + i\frac{\hat{p}}{\hbar}a + \frac{1}{2!}\left(i\frac{\hat{p}}{\hbar}a\right)^2 + \ldots\right) && (4.100) \\ &= 1\left(-i\frac{\hat{p}}{\hbar}a + i\frac{\hat{p}}{\hbar}a + \frac{1}{2!}\left(-i\frac{\hat{p}}{\hbar}a\right)^2 + \frac{1}{2!}\left(i\frac{\hat{p}}{\hbar}a\right)^2 + \ldots\right) = 1 && (4.101) \end{aligned}$$

thus we see that

$$\hat{U}_x(a)\hat{U}_x(-a) = \hat{\mathbf{I}} \iff \hat{U}_x(a) = \left(\hat{U}_x(a)\right)^\dagger = \hat{U}_x(-a). \tag{4.102}$$

which means that $\hat{U}_x(a)$ is a **unitary operator**. Note that unitary operators do not change the norm of a vector, i.e. $\left\|\hat{U}_x(a)|\psi\rangle\right\| = \||\psi\rangle\|$. Now, notice that this operator can be differentiated by the parameter a:

$$\frac{d}{da}\left(\hat{U}_x(a)\right) = -i\frac{\hat{p}}{\hbar}\hat{U}_x(a), \frac{d}{da}\left(\hat{U}_x(a)\right)^\dagger = i\frac{\hat{p}}{\hbar}\hat{U}_x(a)^\dagger \tag{4.103}$$

Note:

The momentum operator $\hat{p}$ and $\hat{U}_x(a)$ commute: since $[\hat{p}^n, \hat{p}] = 0$ and $\hat{U}_x(a)$ is essentially a linear combination of $\hat{p}^n$. Thus,

$$\left[\hat{U}_x(a), \hat{p}\right] = 0 \iff \hat{U}_x(a)\hat{p} = \hat{p}\hat{U}_x(a) \tag{4.104}$$

meaning that we can change the order of the said operators.

thus, by considering the following quantity:

$$\frac{d}{da}\left(\hat{S}_p^\dagger(a)\hat{x}\left(\hat{U}_x(a)\right)\right) = \frac{d}{da}\left(\hat{S}_p^\dagger(a)\right)\hat{x}\hat{U}_x(a) + \hat{U}_x(-a)\hat{x}\frac{d}{da}\left(\hat{U}_x(a)\right) \tag{4.105}$$

$$= \hat{S}_p^\dagger(a)\left(i\frac{\hat{p}}{\hbar}\right)\hat{x}\hat{U}_x(a) - \hat{U}_x(-a)\hat{x}\left(i\frac{\hat{p}}{\hbar}\right)\hat{U}_x(a) \tag{4.106}$$

$$= \hat{S}_p^\dagger(a)\left[i\frac{\hat{p}}{\hbar}, \hat{x}\right]\hat{U}_x(a) \tag{4.107}$$

$$= \frac{1}{i\hbar}\hat{S}_p^\dagger(a)\underbrace{[\hat{x}, \hat{p}]}_{=i\hbar}\hat{U}_x(a) \tag{4.108}$$

$$= \hat{S}_p^\dagger(a)\hat{U}_x(a) = \hat{\mathbf{I}} \tag{4.109}$$

$$\implies \quad \hat{S}_p^\dagger(a)\hat{x}\left(\hat{U}_x(a)\right) = a\hat{\mathbf{I}} + \hat{C} \tag{4.110}$$

where $\hat{C}$ is a constant operator that does not depend on a. If we substitute $a = 0$, we see that

$$\hat{U}_x(0) = \hat{S}_p^\dagger(0) = \hat{\mathbf{I}} \implies \hat{C} = \hat{x} \tag{4.111}$$

Thus, we have the following relation:

$$\hat{S}_p^\dagger(a)\hat{x}\,\hat{U}_x(a) = a\hat{\mathbf{I}} + \hat{x} \implies \quad \hat{x}\,\hat{U}_x(a) \quad = a\hat{U}_x(a) + \hat{U}_x(a)\hat{x} \tag{4.112}$$

$$\implies \quad [\hat{s}, \hat{x}] = a\hat{U}_x(a) \tag{4.113}$$

thus by applying $\hat{x}\hat{U}_x(a)$ to $|x\rangle$,

$$\hat{x}\hat{U}_x(a)|x\rangle = (\left[\hat{x}, \hat{U}_x(a)\right] + \hat{U}_x(a)\hat{x})|x\rangle \tag{4.114}$$

$$= a\hat{U}_x(a)|x\rangle + \hat{U}_x(a)\hat{x}|x\rangle \tag{4.115}$$

$$= (x + a)\hat{U}_x(a)|x\rangle \tag{4.116}$$

thus the vector $\hat{U}_x(a)|x\rangle$ behaves like $|x + a\rangle$(whose eigenvalue for $\hat{x}$ is $x + a$). and so we call $\hat{U}_x(a)$ the **shift operator**:

Definition 4.20: Shift Operator

The shift operator $\hat{U}_x(a)$ is defined as

$$\hat{U}_x(a) = e^{-i\frac{\hat{p}}{\hbar}a}, \tag{4.117}$$

and it shifts the position eigenstate $|x\rangle$ by a:

$$\hat{U}_x(a)|x\rangle = |x+a\rangle. \tag{4.118}$$

It is a unitary operator that commutes with the momentum operator $\hat{p}$.

Now, notice that when a is small, let's say $a = \epsilon$, we can expand the exponential:

$$\hat{U}_x(\epsilon) = 1 - i\frac{\hat{p}}{\hbar}\epsilon + \frac{1}{2!}\left(-i\frac{\hat{p}}{\hbar}\epsilon\right)^2 + \ldots \tag{4.119}$$

if we take ϵ small enough, the $\epsilon^2, \epsilon^3, \ldots$ terms will be much smaller than the ϵ term, so:

$$\hat{U}_x(\epsilon) \approx 1 - i\frac{\hat{p}}{\hbar}\epsilon \tag{4.120}$$

thus by applying this to a vector $|x\rangle$, we have

$$\hat{U}_x(\epsilon)|x\rangle \approx \left(1 - i\frac{\hat{p}}{\hbar}\epsilon\right)|x\rangle = |x+\epsilon\rangle \tag{4.121}$$

$$\implies \quad \frac{|x+\epsilon\rangle - |x\rangle}{\epsilon} \approx -i\frac{\hat{p}}{\hbar}|x\rangle \tag{4.122}$$

the LHS is the **derivative** of the vector $|x\rangle$ with respect to the parameter ϵ:

Definition 4.21: Derivative of a Vector

The derivative of a continous eigenvector $|x\rangle$ with respect to x is defined as

$$\frac{d}{dx}|x\rangle := \lim_{\epsilon\to 0}\frac{|x+\epsilon\rangle - |x\rangle}{\epsilon} \tag{4.123}$$

Thus, we can write the above relation as

$$\frac{\partial}{\partial x}|x\rangle = -i\frac{\hat{p}}{\hbar}|x\rangle \iff \hat{p}|x\rangle = i\hbar\frac{d}{dx}|x\rangle \tag{4.124}$$

taking the Hermitian conjugate,

$$\langle x|\hat{p} = -i\hbar\frac{\partial}{\partial x}\langle x| \tag{4.125}$$

given a state $|\psi\rangle$, this can be treated as the derivative on the wavefunction $\psi(x) = \langle x|\psi\rangle$:

$$\langle x \,|\, \hat{p} \,|\, \psi\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|\psi\rangle = -i\hbar \frac{\partial}{\partial x} \psi(x). \tag{4.126}$$

This is called the position representation of the momentum operator:

$$\hat{p}|x\rangle = i\hbar \frac{\partial}{\partial x} |x\rangle \tag{4.127}$$

because in the momentum representation $\hat{p}$ acts as a multiplication operator:

$$\hat{p}|p\rangle = p|p\rangle \implies \hat{p} = p. \tag{4.128}$$

and by the duality of commutator, $\hat{x}$ acts as a derivative operator in the momentum representation:

$$\hat{x} = i\hbar \frac{\partial}{\partial p}. \tag{4.129}$$

4.4.6 3D Case

In the 3D case, we can define the position operator $\hat{\vec{x}}$ and momentum operator $\hat{\vec{p}}$ as vectors:

$$\hat{\vec{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3), \quad \hat{\vec{p}} = (\hat{p}_1, \hat{p}_2, \hat{p}_3) \tag{4.130}$$

which has eigenvectors corresponding to each position and momentum vectors:

$$\hat{\vec{x}}|\vec{x}\rangle = \vec{x}|\vec{x}\rangle, \quad \hat{\vec{p}}|\vec{p}\rangle = \vec{p}|\vec{p}\rangle \tag{4.131}$$

The canonical commutation relation in 3D is given by

$$\left[\hat{x}_i, \, \hat{p}_j\right] = i\hbar\delta_{ij}, \quad i, j = 1, 2, 3. \tag{4.132}$$

and the momentum operator in the position representation is given by

$$\hat{\vec{p}} = -i\hbar\nabla, \quad \text{where } \nabla = \begin{pmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix} \tag{4.133}$$

Essentially, the same discussion applies to the 3D case as well.

Now, in 3D, we can define another important quantity in quantum mechanics: the **orbital angular momentum** operator, which is defined as follows:

Definition 4.22: Orbital Angular Momentum Operator

The orbital angular momentum operator $\hat{\vec{L}}$ is defined as

$$\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}} = \begin{pmatrix} \hat{L}_x \\ \hat{L}_y \\ \hat{L}_z \end{pmatrix} = \begin{pmatrix} \hat{x}_2\hat{p}_3 - \hat{x}_3\hat{p}_2 \\ \hat{x}_3\hat{p}_1 - \hat{x}_1\hat{p}_3 \\ \hat{x}_1\hat{p}_2 - \hat{x}_2\hat{p}_1 \end{pmatrix}. \tag{4.134}$$

or equivalently,

$$\hat{L}_i = \sum_{j,k=1}^{3} \epsilon_{ijk}\hat{x}_j\hat{p}_k, \quad i,j,k = 1,2,3. \tag{4.135}$$

and the total orbital angular momentum operator $\hat{\vec{L}}^2$:

Definition 4.23: Total Orbital Angular Momentum Operator

The total orbital angular momentum operator $\hat{\vec{L}}^2$ is defined as

$$\hat{\vec{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2. \tag{4.136}$$

Theorem 4.4.5 Properties of Angular Momentum Operators

$$\left[\hat{L}_i,\, \hat{L}_j\right] = i\hbar\epsilon_{ijk}\hat{L}_k, \quad i,j,k = 1,2,3 \tag{4.137}$$

$$\left[\hat{\vec{L}}^2,\, \hat{L}_i\right] = 0 \tag{4.138}$$

These properties imply two important properties:

- Any two(or all three) components of the angular momentum operator $\hat{L}_i$ do not commute, which means that they cannot be measured simultaneously with arbitrary precision.
- However, the total angular momentum operator $\hat{\vec{L}}^2$ commutes with any component of the angular momentum operator $\hat{L}_i$, which means that we can measure the total angular momentum and one component of it simultaneously with arbitrary precision.

4.5 Heisenberg Equation and Schrödinger Equation

Now, we have postulated the canonical commmutation from the Poisson brackets in classical mechanics:

$$\{x,\, p\} = 1 \longrightarrow [\hat{x},\, \hat{p}] = i\hbar \tag{4.139}$$

and so we postulate that in general, there exists the following correspondence:

$$\{A, B\} = C \longrightarrow \left[\hat{A}, \hat{B}\right] = i\hbar\hat{C} \tag{4.140}$$

One example is:

$$\{A, H\} = \frac{dA}{dt} \longrightarrow \left[\hat{A}(t), \hat{H}\right] = i\hbar\frac{d\hat{A}}{dt} \tag{4.141}$$

This is called the **Heisenberg Equation**, and at the same time, it is insinuated that what really evolves through time is the operator $\hat{A}$ (the outcomes), and the state (or probability distribution of the system) is fixed:

Principle 4.6: Heisenberg Picture and Heisenberg Equation

The point of view that the operators evolve through time, while the states are fixed is called the **Heisenberg picture**. The time evolution of an operator $\hat{A}(t)$ in the Heisenberg picture is given by the **Heisenberg Equation**:

$$\left[\hat{A}(t), \hat{H}\right] = i\hbar\frac{d\hat{A}(t)}{dt} \tag{4.142}$$

note: we assume that at $t = 0$, $\hat{A}(0) = \hat{A}$.

in contrast, we have another point of view:

Principle 4.7: Schrödinger Picture

The point of view that the states evolve through time, while the operators are fixed is called the **Schrödinger picture**. We denote the time evolved states as $|\psi(t)\rangle$,

If we remember the Theorem 4.4.3:

$$[f(\hat{x}), \hat{p}] = i\hbar\frac{\partial f}{\partial x}(\hat{x}), \quad \hat{p}|x\rangle = i\hbar\frac{\partial}{\partial x}|x\rangle. \tag{4.143}$$

and the shift operator was defined as

$$\hat{U}_x(a) = e^{-i\frac{\hat{p}}{\hbar}a} \tag{4.144}$$

we postulate then that by changing $\hat{p} \to \hat{H}$ and $\frac{\partial}{\partial x} \to \frac{d}{dt}$, we can define the time evolution operator:

Definition 4.24: Time Evolution Operator

The time evolution operator $\hat{U}(t)$ is defined as

$$\hat{U}(t) = e^{-i\frac{\hat{H}}{\hbar}t}, \tag{4.145}$$

where $\hat{H}$ is the Hamiltonian operator of the system.

and hence the time evolution of a state $|\psi(0)\rangle$ is given by

$$|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle \tag{4.146}$$

thus by taking the time derivative on both sides, we obtain the **Schrödinger Equation**:

Principle 4.8: Schrödinger Equation

The **Schrödinger Equation** is given by

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle, \tag{4.147}$$

where $\hat{H}$ is the Hamiltonian operator of the system.

Note that the Hamiltonian operator $\hat{H}$ corresponds to the Hamiltonian in the classical mechanics. Also, if we apply $\langle x|$, we can obtain the Schrödinger Equation we are familiar with:

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = \hat{H}\psi(x,t), \tag{4.148}$$

where $\psi(x,t) = \langle x|\psi(t)\rangle$ is the wavefunction in the position representation.

Coming back to the Heisenberg picture, we expect that the expectation value of $\hat{A}(t)$ and $\hat{A}$ in state $|\psi(t)\rangle$ are the same:

$$\langle\psi(0)\,|\,\hat{A}(t)\,|\,\psi(0)\rangle = \langle\psi(t)\,|\,\hat{A}\,|\,\psi(t)\rangle \tag{4.149}$$

$$= \langle\psi(0)\,|\,\hat{U}^\dagger(t)\hat{A}\hat{U}(t)\,|\,\psi(0)\rangle \tag{4.150}$$

Thus in the Heisenberg picture, the time evolution of operator $\hat{A}(t)$ is given by

$$\hat{A}(t) = \hat{U}^\dagger(t)\hat{A}\hat{U}(t). \tag{4.151}$$

4.6 Harmonic Oscillator

4.6.1 Ladder Operators and Number Operator

4.6.2 Solution

Chapter 5

Mathematical Remarks

5.1 Functional Derivative

5.1.1 Definition

Consider a quantity I defined as follows:

$$I := \int_A^B dx\, F(x) \tag{5.1}$$

Notice that I is not really a function of x, but if you had to say, it is more a "function" of F - may be $F(x) = e^x$, or $F(x) = ax^2 + bx + c$, or, etc. So, to denote the dependence of I on the function F, we write

$$I[f] := \int_A^B dx\, F(x) \tag{5.2}$$

This is called a **(linear) functional**. Now, imagine that F is a function of f, for example, $F[f] = f(x)^2$. By chain rule, a small change in F, denoted as δF, can be expressed as:

$$\delta F[f] = F[f + \delta f] - F[f] \tag{5.3}$$

$$= \frac{\partial F}{\partial f} \delta f \tag{5.4}$$

so, similarly,

$$\delta I[F] = I[F + \delta F] - I[F] \tag{5.5}$$

$$= \int_A^B dx\ \delta F[f] \tag{5.6}$$

$$= \int_A^B dx\, \frac{\partial F}{\partial f} \delta f \tag{5.7}$$

Then, the **functional derivative** of I with respect to f, $\frac{\delta I}{\delta f}$, is defined as follows:

Definition 5.1: Functional Derivative

If a function $\phi(x)$ exists, such that

$$\delta I = \int_A^B dx\, \phi(x)\, \delta f(x), \tag{5.8}$$

we say that $\phi(x)$ is the **functional derivative** of I with respect to f, and denote it as

$$\frac{\delta I}{\delta f(x)} := \phi(x) \iff \delta I := \int_B^A dx\, \frac{\delta I}{\delta f(x)}\, \delta f(x). \tag{5.9}$$

Immidiately, by comparing Eq.(5.9) and Eq.(5.7), we see the following relation:

Theorem 5.1.1 Functional-density relation

For a quantity $I[F]$ and its density $F[f(x)]$, the functional derivative satisfies the following relation:

$$I = \int_A^B dx\, F[f(x)] \quad \Longrightarrow \quad \frac{\delta I}{\delta f(x)} = \frac{\partial F[f(x)]}{\partial f(x)} \tag{5.10}$$

5.1.2 Two function case

Consider a case where I is the functional of F, which is also a functional of f and g:

$$I[F[f,g]] = \int_B^A dx\, F[f(x), g(x)] \tag{5.11}$$

Or more generally, if a function $D(x)$ satisfies the following Now, let us add some small change of f, δf:

$$I[F[f+\delta f, g]] = \int_B^A dx\, F[f(x) + \delta f(x), g(x)] \tag{5.12}$$

$$= \int_B^A dx\, F[f,g] + \frac{\partial F}{\partial f}\delta f \tag{5.13}$$

and similarly, by adding δg,

$$I[F[f, g+\delta g]] = \int_B^A dx\, F[f(x), g(x) + \delta g(x)] \tag{5.14}$$

$$= \int_B^A dx\, F[f,g] + \frac{\partial F}{\partial g}\delta g \tag{5.15}$$

Combining these two, we have

$$I[F[f+\delta f, g+\delta g]] = \int_B^A dx\, F[f(x)+\delta f(x), g(x)+\delta g(x)] \tag{5.16}$$

$$= \int_B^A dx\, F[f,g] + \frac{\partial F}{\partial f}\delta f + \frac{\partial F}{\partial g}\delta g \tag{5.17}$$

$$\implies \quad I[F[f+\delta f, g+\delta g]] - I[F[f,g]] = \int_B^A dx\, \frac{\partial F}{\partial f}\delta f + \frac{\partial F}{\partial g}\delta g \tag{5.18}$$

In this case, we just like our normal derivative, we should denote the LHS as:

$$\delta I := \int_B^A dx\, \frac{\partial F}{\partial f}\delta f + \frac{\partial F}{\partial g}\delta g \tag{5.19}$$

or alternatively,

$$\delta I := \int_B^A dx\, \frac{\delta I}{\delta f}\delta f + \frac{\delta I}{\delta g}\delta g \tag{5.20}$$

5.1.3 Euler-Lagrange Equation

For the two function case, especially consider that $g = \frac{df}{dx}$, and let us see what happens. Specifically, let us set that $\delta f(A) = \delta f(B) = 0$. Then, we have:

$$\frac{\delta I}{\delta g}\delta g = \frac{\delta I}{\delta \frac{df}{dx}}\delta \frac{df}{dx} \tag{5.21}$$

we can change the order of the derivative:

$$= \frac{\delta I}{\delta \frac{df}{dx}} \frac{d\,\delta f}{dx} \tag{5.22}$$

from the differentiation of a product, we have

$$= \frac{d}{dx}\left(\frac{\delta I}{\delta\left(\frac{df}{dx}\right)}\delta f\right) - \frac{d}{dx}\frac{\delta I}{\delta\left(\frac{df}{dx}\right)}\delta f \tag{5.23}$$

Now, subsituting this to the two function case, we get:

$$\delta I = \int_B^A dx \left[\left(\frac{\delta I}{\delta f} - \frac{d}{dx}\frac{\delta I}{\delta\left(\frac{df}{dx}\right)}\right)\delta f + \frac{d}{dx}\left(\frac{\delta I}{\delta\left(\frac{df}{dx}\right)}\delta f\right)\right] \tag{5.24}$$

the total derivative term is zero, since $\delta f(A) = \delta f(B) = 0$. Thus, we have

$$\delta I = \int_B^A dx \left(\frac{\delta I}{\delta f} - \frac{d}{dx} \frac{\delta I}{\delta\left(\frac{df}{dx}\right)} \right) \delta f \tag{5.25}$$

and since $I = \int_B^A dx\, F[f(x), g(x)]$, we can say that

$$\frac{\partial F}{\partial f} = \frac{\delta I}{\delta f}, \quad \frac{\partial F}{\partial g} = \frac{\delta I}{\delta g} \tag{5.26}$$

Then

$$\delta I = \int_B^A dx \left(\frac{\partial F}{\partial f} - \frac{d}{dx} \frac{\partial F}{\partial\left(\frac{df}{dx}\right)} \right) \delta f \tag{5.27}$$

And if we somehow want to find a minimum of I, we can set $\delta I = 0$:

$$\implies \quad \frac{\partial F}{\partial f} - \frac{d}{dx} \frac{\partial F}{\partial\left(\frac{df}{dx}\right)} = 0 \tag{5.28}$$

This is called the **Euler-Lagrange equation**.

Theorem 5.1.2 Euler-Lagrange Equation
For a functional $I\left[F(f, \frac{df}{dx})\right]$ to be stationary, ($\delta I = 0$), the **Euler-Lagrange equation** must be satisfied:

$$\delta I\left[F(f, \frac{df}{dx})\right] \iff \frac{\partial F}{\partial f} - \frac{d}{dx}\left(\frac{\partial F}{\partial\left(\frac{df}{dx}\right)} \right) = 0 \tag{5.29}$$

5.1.4 Important Property

In general, consider that the functional F is a function of $f_1(t), f_2(t), \dots, f_n(t)$:

$$F[f_1, f_2, \dots, f_n] \implies \delta F = \sum_{j=1}^{n} \frac{\delta F}{\delta f_j(t)} \delta f_j(t) \tag{5.30}$$

If $F = f_i$, we expect that

$$\frac{\delta f_i}{\delta f_i} = 1 \implies \delta f_i = \sum_{j=1}^{n} \frac{\delta f_i}{\delta f_j} \delta f_j \implies \quad \frac{\delta f_i}{\delta f_j} = \delta_{ij} \tag{5.31}$$

Similarly, consider a continous case where F is a function of $f(x)$:,

$$F[f(x,t)] \implies \delta F = \int_A^B dx' \frac{\delta F}{\delta f(x', t)} \delta f(x', t) \tag{5.32}$$

Note the distinction between the variable x and the integration variable x'. This is because x is an "index" of $f(t)$: $f_i \to f(x)$. Then, if we set $F = f(x)$, we expect that

$$\frac{\delta f(x,t)}{\delta f(x,t)} = 1 \implies \delta f(x,t) = \int_A^B dx' \, \frac{\delta f(x,t)}{\delta f(x',t)} \delta f(x',t) \tag{5.33}$$

Comparing with this with the definition of **Dirac delta function**:

Definition 5.2: Dirac Delta Function

The **Dirac delta function** $\delta(x)$ is defined as a function that satisfies the following property:

$$\int dx' \, \delta(x'-x)\varphi(x') = \varphi(x), \quad {}^\forall \varphi(x') \in C^\infty \tag{5.34}$$

we have

Theorem 5.1.3 Property of Functional Derivative

For a functional $f_i(t)$ or $f(x,t)$, the functional derivative satisfies the following property:

$$\frac{\delta f_i(t)}{\delta f_j(t)} = \delta_{ij}, \quad \text{or} \quad \frac{\delta f(x,t)}{\delta f(x',t)} = \delta(x-x') \tag{5.35}$$

5.1.5 In a n-dimensional space

In the previous discussions, we have considered the functional I as an integral on 1D space represented by x. Here, we aim to generalize the discussion to n-dimensional space, e.g. $\mathbb{R}^n$. For $\mathbb{R}^n$, let us define a functional I and functional derivative as follows:

Definition 5.3: Functional on $\mathbb{R}^n$

A quantity $I \in \mathbb{R}$ defined on $V \subset \mathbb{R}^n$ is called a **(linear) functional** if it can be expressed as follows:

$$I[f_i] := \int_{x \in V} d^n x \, F(f_i(x)) = \int_V d^n x \, F(f_i(x)), \quad i \in \mathbb{N} \tag{5.36}$$

Definition 5.4: Functional Derivative on $\mathbb{R}^n$

The **functional derivative** of a functional $I[f_i]$ with respect to $f_i(x)$ is defined as follows:

$$\frac{\delta I[f_j]}{\delta f_i(x)} := \phi_i(x) \underset{\text{def}}{\iff} \delta I = \int_V d^n x' \, \frac{\delta I[f_j]}{\delta f_i(x')} \, \delta f_i(x') \tag{5.37}$$

where $\phi_i(x)$ is a function of x.

Then, the variation of the functional δI can be expressed as:

$$\delta I = \delta \int_V d^n x' \, F(f_j(x')) = \int_V d^n x' \; \delta F(f_j(x')) = \int_V d^n x' \; \sum_j \frac{\partial F}{\partial f_j(x')} \, \delta f_j(x') \tag{5.38}$$

Now,

$$\delta f_j(x') = \int_V d^n x \, \delta^n (x - x') \delta f_j(x) \tag{5.39}$$

$$= \int_V d^n x \, \sum_i \delta_{ij} \, \delta^n (x - x') \delta f_i(x) \tag{5.40}$$

$$\implies \quad \frac{\delta f_j(x')}{\delta f_i(x)} = \delta_{ij} \, \delta^n (x - x') \tag{5.41}$$

so,

$$\frac{\delta I}{\delta f_i(x)} = \int_V d^n x' \, \frac{\partial F}{\partial f_i(x')} \delta(x - x') = \frac{\partial F}{\partial f_i(x)} \tag{5.42}$$

thus we see that the functional-density relation still holds:

Theorem 5.1.4 Functional-density relation

For a quantity $I[F]$ and its density $F[f_i(x)]$, the functional derivative satisfies the following relation:

$$I = \int_V d^n x \, F[f_j(x)] \qquad \implies \qquad \frac{\delta I[f_j]}{\delta f_i(x)} = \frac{\partial F[f_j]}{\partial f_i(x)} \tag{5.43}$$

[4]

Chapter 6 Extra (1): Proofs

6.1 Invariance of Euler-Lagrange Equation

Let us look at how the Euler-Lagrange Equation transforms under a point transformation of the generalized coordinates. The Euler-Lagrange Equation in the old coordinates q_i is given by:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0 \tag{6.1}$$

Proposition 6.1.1 Euler-Lagrange Equation

In the new coordinates Q_j, the Euler-Lagrange Equation still holds:

$$\frac{\partial L}{\partial Q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_j}\right) = 0 \tag{6.2}$$

where $L(Q_j, \dot{Q}_j, t)$ is the Lagrangian written in the new coordinates.

Under the point transformation:

$$Q_j = Q_j(q_1, q_2, \dots, q_n, t) = Q_j(q_i, t) \tag{6.3}$$

the Lagrangian L transforms as follows:

$$\tilde{L}(Q, \dot{Q}, t) = L(q(Q, t), \dot{q}(Q, \dot{Q}, t), t) \tag{6.4}$$

Let us consider how the new coordinates $\dot{Q}_j$ can be written in terms of the old coordinates $\dot{q}_i$:

$$\dot{Q}_j = \frac{dQ_j}{dt} = \frac{\partial Q_j}{\partial t} + \sum_k \frac{\partial Q_j}{\partial q_k}\dot{q}_k \implies \frac{\partial \dot{Q}_j}{\partial \dot{q}_i} = \frac{\partial Q_j}{\partial q_i} \tag{6.5}$$

using this,

$$\frac{d}{dt}\left(\frac{\partial \dot{Q}_j}{\partial \dot{q}_i}\right) = \frac{d}{dt}\left(\frac{\partial Q_j}{\partial q_i}\right) = \frac{\partial^2 Q_j}{\partial t\, \partial q_i} + \sum_k \frac{\partial^2 Q_j}{\partial q_k\, \partial q_i} = \frac{\partial^2 Q_j}{\partial q_i\, \partial t} + \sum_k \frac{\partial^2 Q_j}{\partial q_i\, \partial q_k} \tag{6.6}$$

and

$$\frac{\partial \dot{Q}_j}{\partial q_i} = \frac{\partial}{\partial q_i}\left(\frac{\partial Q_j}{\partial t} + \sum_k \frac{\partial Q_j}{\partial q_k}\right) = \frac{\partial^2 Q_j}{\partial q_i\,\partial t} + \sum_k \frac{\partial^2 Q_j}{\partial q_i\,\partial q_k} = \frac{d}{dt}\left(\frac{\partial \dot{Q}_j}{\partial \dot{q}_i}\right) \tag{6.7}$$

Now, we try to write the Euler-Lagrange Equation in the old coordinates in terms of the new coordinates:

$$\frac{\partial L(Q_j, \dot{Q}_j, t)}{\partial q_i} = \sum_j \frac{\partial L}{\partial Q_j}\frac{\partial Q_j}{\partial q_i} + \frac{\partial L}{\partial \dot{Q}_j}\frac{\partial \dot{Q}_j}{\partial q_i} + \frac{\partial L}{\partial t}\frac{\partial t}{\partial q_i} \tag{6.8}$$

Now, time t is an independent variable: a parameter that "we" set to describe the system, and crucially does not depend on the generalized coordinates q_i or the generalized velocities $\dot{q}_i$. Thus, we have $\frac{\partial t}{\partial q_i} = 0$.

$$\implies \quad \frac{\partial L(Q_j, \dot{Q}_j, t)}{\partial q_i} = \sum_j \frac{\partial L}{\partial Q_j}\frac{\partial Q_j}{\partial q_i} + \frac{\partial L}{\partial \dot{Q}_j}\frac{\partial \dot{Q}_j}{\partial q_i} \tag{6.9}$$

and

$$\frac{\partial L(Q_j, \dot{Q}_j, t)}{\partial \dot{q}_i} = \sum_j \frac{\partial L}{\partial Q_j}\frac{\partial Q_j}{\partial \dot{q}_i} + \frac{\partial L}{\partial \dot{Q}_j}\frac{\partial \dot{Q}_j}{\partial \dot{q}_i} \tag{6.10}$$

since $Q_j(q_i, t)$ is not a function of $\dot{q}_i$, we have $\frac{\partial Q_j}{\partial \dot{q}_i} = 0$.

$$\implies \quad \frac{\partial L(Q_j, \dot{Q}_j, t)}{\partial \dot{q}_i} = \sum_j \frac{\partial L}{\partial \dot{Q}_j}\frac{\partial \dot{Q}_j}{\partial \dot{q}_i} \tag{6.11}$$

$$\implies \quad \frac{d}{dt}\left(\frac{\partial \dot{Q}_j}{\partial \dot{q}_i}\right) = \sum_j \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_j}\right)\frac{\partial \dot{Q}_j}{\partial \dot{q}_i} + \frac{\partial L}{\partial \dot{Q}_j}\frac{d}{dt}\left(\frac{\partial \dot{Q}_j}{\partial \dot{q}_i}\right) \tag{6.12}$$

So, if we compare each term of the Euler-Lagrange Equation in the old coordinates,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \sum_j \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_j}\right)\frac{\partial \dot{Q}_j}{\partial \dot{q}_i} + \frac{\partial L}{\partial \dot{Q}_j}\frac{d}{dt}\left(\frac{\partial \dot{Q}_j}{\partial \dot{q}_i}\right) \tag{6.13}$$

$$\frac{\partial L}{\partial q_i} = \sum_j \frac{\partial L}{\partial Q_j}\frac{\partial Q_j}{\partial q_i} + \frac{\partial L}{\partial \dot{Q}_j}\frac{\partial \dot{Q}_j}{\partial q_i} \tag{6.14}$$

from Eq. (6.5) and Eq. (6.7), we can see that

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \sum_j \frac{\partial Q_j}{\partial q_i}\left[\frac{\partial L}{\partial Q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_j}\right)\right] = 0 \tag{6.15}$$

Now, if we remember that the (i-th) component of the product of a matrix and a vector is given by

$$(A\vec{x})_i = \sum_j A_{ij}x_j \tag{6.16}$$

So Eq. (6.15) can be rewritten as a product of a matrix and a vector:

$$\sum_j J_{ij}x_j = 0 \quad \text{where} \quad J_{ij} = \frac{\partial Q_j}{\partial q_i}, \quad x_j = \frac{\partial L}{\partial Q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_j}\right) \tag{6.17}$$

As long as the point transformation is invertible(i.e. changing from 3D Cartesian coordinates to 3D spherical coordinates), the Jacobian matrix J_{ij} is invertible, and thus the only solution to the equation above is $x_j = 0$ for all j. Thus, we have

$$\frac{\partial L}{\partial Q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_j}\right) = 0 \tag{6.18}$$

[3]

6.2 Noether's Theorem

6.2.1 Symmetry under Point Transformation

Theorem 6.2.1 Noether's Theorem under Point Transformation

Every real paramter of a continuous symmetry of the action corresponds to a conserved quantity $\mathcal{Q}$. If the continous symmetry is given by an infinitesimal transformation: $Q_i(t) = q_i(t) + \epsilon F_i(q(t), t) + \mathcal{O}(\epsilon^2)$,

$$\mathcal{Q} = \sum_i F_i(q(t), t)\frac{\partial L(q_i(t), \dot{q}_i(t), t)}{\partial \dot{q}_i} - \Lambda(q(t), t) \tag{6.19}$$

where $\Lambda(q(t), t)$ is the change in the Lagrangian under the infinitesimal transformation.

Assume that a system has a continous symmetry. Thus if we consider a very "small" transformation that is almost the identity transformation:

$$Q_i(t) = q_i(t) + \epsilon F_i(q(t), t) + \mathcal{O}(\epsilon^2) \tag{6.20}$$

by differntiating w.r.t. t, we get

$$\dot{Q}_i(t) = \dot{q}_i(t) + \epsilon \frac{d}{dt} F_i(q(t), t) + \mathcal{O}(\epsilon^2) \tag{6.21}$$

Here, since the system has a symmetry under this transformation, the change in Lagrangian must be up to $\frac{df}{dt}(q(t),t)$:

$$L(Q_i,\dot{Q}_i,t) - L(q_i,\dot{q}_i,t) = \frac{d}{dt}f(q(t),t) \tag{6.22}$$

given that the transformation is infinitesimal, we can write the change in Lagrangian as:

$$\frac{d}{dt}f(q(t),t) = \epsilon\frac{d}{dt}\Lambda(q(t),t) + \mathcal{O}(\epsilon^2) \tag{6.23}$$

$$\implies \quad L(Q_i,\dot{Q}_i,t) - L(q_i,\dot{q}_i,t) = \epsilon\frac{d}{dt}\Lambda(q(t),t) + \mathcal{O}(\epsilon^2) \tag{6.24}$$

Now, the new Lagrangian can be expanded as a Taylor series around the old Lagrangian with a small parameter ϵ:

$$L(Q_i,\dot{Q}_i,t) = L(q_i,\dot{q}_i,t) + \sum_i\left(\epsilon F_i(q,t)\frac{\partial L(q_i,\dot{q}_i,t)}{\partial q_i} + \epsilon\frac{d}{dt}F_i(q_i,t)\frac{\partial L(q_i,\dot{q}_i,t)}{\partial \dot{q}_i}\right) + \mathcal{O}(\epsilon^2) \tag{6.25}$$

So the change in Lagrangian is:

$$L(Q_i,\dot{Q}_i,t) - L(q_i,\dot{q}_i,t) = \epsilon\sum_i\left(F_i(q,t)\frac{\partial L(q_i,\dot{q}_i,t)}{\partial q_i} + \frac{d}{dt}F_i(q_i,t)\frac{\partial L(q_i,\dot{q}_i,t)}{\partial \dot{q}_i}\right) + \mathcal{O}(\epsilon^2) \tag{6.26}$$

From the Euler-Lagrange equation, we know that

$$\frac{\partial L(q_i,\dot{q}_i,t)}{\partial q_i} = \frac{d}{dt}\left(\frac{\partial L(q_i,\dot{q}_i,t)}{\partial \dot{q}_i}\right) \tag{6.27}$$

the change in Lagrangian can be written as:

$$L(Q_i,\dot{Q}_i,t) - L(q_i,\dot{q}_i,t) = \epsilon\frac{d}{dt}\sum_i\left(F_i(q,t)\frac{\partial L(q_i,\dot{q}_i,t)}{\partial \dot{q}_i}\right) + \mathcal{O}(\epsilon^2) \tag{6.28}$$

By comparing Eq. (6.24) and Eq. (6.28), we can see that

$$\frac{d}{dt}\sum_i F_i(q_i(t),t)\frac{\partial L(q_i(t),\dot{q}_i(t),t)}{\partial \dot{q}_i} = \frac{d}{dt}\Lambda(q_i(t),t) \tag{6.29}$$

$$\implies \quad \frac{d}{dt}\left[\sum_i F_i(q_i(t),t)\frac{\partial L(q_i(t),\dot{q}_i(t),t)}{\partial \dot{q}_i} - \Lambda(q_i(t),t)\right] = 0 \tag{6.30}$$

Thus, the quantity:

$$\mathcal{Q} := \sum_i\left[F_i(q(t),t)\frac{\partial L(q_i(t),\dot{q}_i(t),t)}{\partial \dot{q}_i}\right] - \Lambda(q(t),t) \tag{6.31}$$

does not change with time, that is, $\mathcal{Q}$ is conserved.

6.2.2 Symmetry under Transformation Including Time

Theorem 6.2.2 Noether's Theorem under Transformation Including Time

If the system is symmetric under the transformation includes time translation:

$$t' = t + \epsilon T(q_i(t), t) + \mathcal{O}(\epsilon^2) \tag{6.32}$$

$$Q_i(t') = q_i(t) + \epsilon F_i(q_i(t), t) + \mathcal{O}(\epsilon^2) \tag{6.33}$$

There is a conserved quantity $\mathcal{Q}$ given by:

$$\begin{aligned}\mathcal{Q} = &\sum_i \left[F_i(q_i(t), t) \frac{\partial L(q_i(t), \dot{q}_i(t), t)}{\partial \dot{q}_i} \right] - \Lambda(q_i, t) \\ &- T(q_i(t), t) \sum_i \left[\dot{q}_i(t) \frac{\partial L(q_i(t), \dot{q}_i(t), t)}{\partial \dot{q}_i} - L(q_i(t), \dot{q}_i(t), t) \right]\end{aligned} \tag{6.34}$$

Given that $t' = t + \epsilon T(q_i(t), t) + \mathcal{O}(\epsilon^2)$, the Taylor expansion of $L(Q_i(t'), \dot{Q}_i(t'), t')$ around $L(Q_i(t), \dot{Q}_i(t), t)$ is:

$$L(Q_i(t'), \dot{Q}_i(t'), t') = L(Q_i(t), \dot{Q}_i(t), t) + \epsilon T(q_i(t), t) \frac{d}{dt} L(Q_i(t), \dot{Q}_i(t), t) + \mathcal{O}(\epsilon^2) \tag{6.35}$$

and

$$\frac{dt'}{dt} = 1 + \epsilon \frac{d}{dt} T(q_i(t), t) + \mathcal{O}(\epsilon^2) \tag{6.36}$$

here, remember the definition of symmetry:

$$\frac{dt'}{dt} L(Q_i(t'), \dot{Q}_i(t'), t') = L(q_i, \dot{q}_i, t) + \frac{d}{dt} f(q_i, t) \tag{6.37}$$

thus, for the system to be symmetric under the transformation including time, we must have:

$$L(q_i, \dot{q}_i, t) + \frac{d}{dt} f(q_i(t), t) \tag{6.38}$$

$$= \left(1 + \epsilon \frac{d}{dt} T(q_i(t), t) + \mathcal{O}(\epsilon^2)\right) \tag{6.39}$$

$$\times \left(L(Q_i(t), \dot{Q}_i(t), t) + \epsilon T(q_i(t), t) \frac{d}{dt} L(Q_i(t), \dot{Q}_i(t), t) + \mathcal{O}(\epsilon^2) \right)$$

$$= L(Q_i, \dot{Q}_i, t) + \epsilon \frac{d}{dt} \Big(T(q_i(t), t) L(Q_i(t), \dot{Q}_i(t), t) \Big) + \mathcal{O}(\epsilon^2) \tag{6.40}$$

Since the difference between $L(Q_i, \dot{Q}_i, t)$ and $L(q_i, \dot{q}_i, t)$ is $\mathcal{O}(\epsilon)$:

$$L(Q_i, \dot{Q}_i, t) = L(q_i, \dot{q}_i, t) + \mathcal{O}(\epsilon) \tag{6.41}$$

we can replace $L(Q_i(t), \dot{Q}_i(t), t)$ with $L(q_i, \dot{q}_i, t)$ in the above equation as the difference is of order $\mathcal{O}(\epsilon^2)$:

$$L(q_i, \dot{q}_i, t) + \frac{d}{dt} f(q_i(t), t) = L(Q_i(t), \dot{Q}_i(t), t) + \epsilon \frac{d}{dt}\Big(T(q_i(t), t) L(q_i(t), \dot{q}_i(t), t)\Big) + \mathcal{O}(\epsilon^2) \tag{6.42}$$

and if we write the change in Lagrangian as:

$$\frac{d}{dt} f(q_i(t), t) = \epsilon \frac{d}{dt} \Lambda(q_i(t), t) + \mathcal{O}(\epsilon^2) \tag{6.43}$$

we get:

$$L(Q_i(t), \dot{Q}_i(t), t) - L(q_i(t), \dot{q}_i(t), t) = -\epsilon \frac{d}{dt}\Big(T(q_i(t), t) L(q_i(t), \dot{q}_i(t), t) - \Lambda(q_i(t), t)\Big) + \mathcal{O}(\epsilon^2) \tag{6.44}$$

Now, given that $t' = t + \epsilon T(q_i(t), t) + \mathcal{O}(\epsilon^2)$, the coordinates and velocities transform as follows:

$$Q_i(t) = \underbrace{q_i(t) + \epsilon F_i(q_i(t), t)}_{=Q_(t')} - \underbrace{\epsilon \frac{dQ_i(t)}{dt} T(q_i(t), t)}_{\text{cancel time transformation}} + \mathcal{O}(\epsilon^2) \tag{6.45}$$

since $Q_i(t) - q_i(t) \approx \mathcal{O}\epsilon$, we can write

$$\dot{Q}_i(t) = q_i(t) + \epsilon\Big(F_i(q_i(t), t) - T(q_i(t), t)\dot{q}_i(t)\Big) + \mathcal{O}(\epsilon^2) \tag{6.46}$$

$$\Longrightarrow \quad \dot{Q}_i(t) = \dot{q}_i(t) + \epsilon \frac{d}{dt}\Big(F_i(q_i(t), t) - T(q_i(t), t)\dot{q}_i(t)\Big) + \mathcal{O}(\epsilon^2) \tag{6.47}$$

by the same discussion in Eq.(6.24) to (6.28), we can write the change in Lagrangian as:

$$L(Q_i(t), \dot{Q}_i(t), t) - L(q_i(t), \dot{q}_i(t), t) \tag{6.48}$$

$$= \epsilon \frac{d}{dt} \sum_i \left[\Big(F_i(q_i(t), t) - T(q_i(t), t)\dot{q}_i(t)\Big) \frac{\partial L(q_i(t), \dot{q}_i(t), t)}{\partial \dot{q}_i}\right] + \mathcal{O}(\epsilon^2) \tag{6.49}$$

by comparing Eq. (6.44) and Eq. (6.49), we can see that

$$-\epsilon \frac{d}{dt}\Big(T(q_i(t), t) L(q_i(t), \dot{q}_i(t), t) - \Lambda(q_i(t), t)\Big) \tag{6.50}$$

$$= \epsilon \frac{d}{dt} \sum_i \left[\Big(F_i(q_i(t), t) - T(q_i(t), t)\dot{q}_i(t)\Big) \frac{\partial L(q_i(t), \dot{q}_i(t), t)}{\partial \dot{q}_i}\right] \tag{6.51}$$

which gives the conserved quantity $\mathcal{Q}$ as:

$$\begin{aligned} \mathcal{Q} = \sum_i \left[F_i(q_i(t), t) \frac{\partial L(q_i(t), \dot{q}_i(t), t)}{\partial \dot{q}_i}\right] - \Lambda(q_i, t) \\ - T(q_i(t), t) \sum_i \left[\dot{q}_i(t) \frac{\partial L(q_i(t), \dot{q}_i(t), t)}{\partial \dot{q}_i} - L(q_i(t), \dot{q}_i(t), t)\right] \end{aligned} \tag{6.52}$$

Chapter 7 Extra (2): Miscellaneous

7.1 Schrödinger Equation - Original Formulation

7.1.1 Historical Background

In 1905, Albert Einstein found the photoelectric effect, suggesting that light has a particle like property, such as discrete energy [5]. Following this work, Arthur Compton in 1923 discovered the Compton effect, which is the scattering of X-rays by electrons, further supporting the particle nature of light and confirming the discrete momentum of photons [6]. Specifically, energy and momentum are each related to the frequency and wavelength of light, respectively, as follows:

Proposition 7.1.1 Planck-Einstein relation

For a photon with frequency ν and wave length λ, the energy E and momentum p are given by

$$E = h\nu, \quad p = \frac{h}{\lambda} = \frac{h\nu}{c} = \frac{E}{c}, \tag{7.1}$$

In 1925, Louis de Brogile posulated that other particles (or any matter) also have a wave-like property, and the energy-frequency/ momentum-wavelength relation is given by the Planck-Einstein relation Eq. (7.1) [7].

Then in 1913, Niels Bohr proposed a model of the hydrogen atom, which describes the electron as a particle orbiting the nucleus in discrete energy levels [8]. The implication of this model is that electrons have a fixed, discrete(quantized) angular momentum $\vec{L} := \vec{r} \times \vec{p}$. For a particle orbiting in a circular orbit, the angular momentum is given by

$$\vec{L} = rp = \frac{hr}{\lambda} \tag{7.2}$$

For the electron wave to be continous around the orbit of radius r, the wavelength must be an integer multiple of the circumference of the orbit:

$$\lambda = \frac{2\pi r}{n}, \quad n = 1, 2, 3, \ldots \implies \vec{L} = n\frac{h}{2\pi} := n\hbar \tag{7.3}$$

7.1.2 Schrödinger Equation

Now, a classical wave obeys a wave equation:

$$\frac{1}{v^2}\frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi = 0 \tag{7.4}$$

but does the matter wave also obey wave equation?

Well, not a classical wave equation, but a quantum wave should follow a quantum wave equation, which must satisfy a few conditions that match with the physics as we know it:

- Planck-Einstein relation: $E = \hbar\omega$, $\vec{p} = \hbar\vec{k}$

- Energy equation: $E = T + V$
- For $V = V_0$, the solution must be a plane wave: $\psi(\vec{x}, t) = Ae^{i(\vec{k}\cdot\vec{x}-\omega t)}$(i.e. constant momentum and constant energy).

The energy equation is given by

$$E = T + V = \frac{p^2}{2m} + V_0 \iff \hbar\omega = \frac{\hbar^2\vec{k}^2}{2m} + V \tag{7.5}$$

If we multiply both sides by $\psi(\vec{x}, t)$, we get

$$\hbar\omega\psi(\vec{x}, t) = \frac{\hbar^2}{2m}\vec{k}^2\psi(\vec{x}, t) + V\psi(\vec{x}, t) \tag{7.6}$$

For a plane wave solution,

$$\frac{\partial}{\partial t}\psi(\vec{x}, t) = -i\omega\psi(\vec{x}, t) \qquad \implies i\hbar\frac{\partial}{\partial t}\psi(\vec{x}, t) = \hbar\omega\psi(\vec{x}, t) \tag{7.7}$$

$$\nabla^2\psi(\vec{x}, t) = -\vec{k}^2\psi(\vec{x}, t) \quad \implies -\frac{\hbar^2}{2m}\nabla^2\psi(\vec{x}, t) = \frac{\hbar^2\vec{k}^2}{2m}\psi(\vec{x}, t) \tag{7.8}$$

Thus, we can write the wave equation as

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{x}, t) = -\frac{\hbar^2}{2m}\nabla^2\psi(\vec{x}, t) + V_0\psi(\vec{x}, t) \tag{7.9}$$

And we postulate that for any potential $V(\vec{x})$, the wave equation is given by

Principle 7.1: Schrödinger Equation

The Schrödinger equation is given by

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{x}, t) = -\frac{\hbar^2}{2m}\nabla^2\psi(\vec{x}, t) + V(\vec{x})\psi(\vec{x}, t) \tag{7.10}$$

where $\psi(\vec{x}, t)$ is the wave function, $V(\vec{x})$ is the potential, and m is the mass of the particle.

7.2 Derivative by Derivative

7.2.1 An Interesting Problem

Question 1: Derivative by derivative

What would the following operator mean?

$$\frac{d}{d\frac{d}{dx}} \tag{7.11}$$

Let us define a differential operator D:

Definition 7.1: Mysterious Operator

The operator D is defined as the derivative of the derivative with respect to x:

$$D := \frac{d}{dx} \tag{7.12}$$

Now let us actually apply D to some functions:

Example 7.2.1 (A power function)
Let us apply $\frac{d}{dx}$ to a power function $f(x) = x^n$, $n \in \mathbb{Z}^+$:

$$Df(x) = \frac{d}{dx}x^n = nx^{n-1} = \frac{n}{x}x^n \implies D = \frac{n}{x} \tag{7.13}$$

Then,

$$\frac{d}{dD}f(x) = \frac{d}{dD}\left(\frac{n}{D}\right)^n = n^n \cdot (-n) \cdot D^{-n-1} = -\left(\frac{n}{D}\right)^{n+1} = -x^{n+1} \tag{7.14}$$

Example 7.2.2 (Exponential Function)
Let us apply $\frac{d}{dx}$ to the exponential function $f(x) = e^{\alpha x}$:

$$\frac{df}{dx} = \frac{d}{dx}e^{\alpha x} = \alpha f(x) \implies \frac{d}{dx} = \alpha \tag{7.15}$$

$$\implies \quad Df(x) = \frac{de^{\alpha x}}{d\alpha} = xe^{\alpha x} \tag{7.16}$$

An important property of this operator is in the commutator with the x operator:

Definition 7.2: x operator

The x operator is defined as the operator that maps a function f to the function $x \cdot f$:

$$x : f \mapsto x \cdot f \tag{7.17}$$

Definition 7.3: Commutator

For two operators $A : f \mapsto Af$ and $B : f \mapsto Bf$, the commutator is defined as:

$$[A,\, B] := AB - BA \tag{7.18}$$

Then the commutator of the x operator and the D operator is given by:

$$[x,\, D] = xD - Dx \tag{7.19}$$

applying this to a function $f(x)$, we have:

$$\begin{aligned} [x,\, D]\, f(x) &= xDf(x) - D(xf(x)) && (7.20)\\ &= xDf(x) - Dxf(x) - xDf(x) && (7.21)\\ &= -Df(x) = -1 \cdot f(x) && (7.22)\\ \implies \quad [x,\, D] &= -1 && (7.23) \end{aligned}$$

and notice since both x and D are just linear operators, if we define $x = \frac{d}{dD}$, then we have:

$$\begin{aligned} [x,\, D]\, f(D) &= \frac{d}{dD} Df(D) - D\frac{d}{dD} f(D) && (7.24)\\ &= f(D) + D\frac{df}{dD} - D\frac{df}{dD} && (7.25)\\ &= f(D) && (7.26) \end{aligned}$$

so in general,

$$\left[\omega,\, \frac{d}{d\omega}\right] = -1 \tag{7.27}$$

Note the bi-linearity of the commutator:

Theorem 7.2.1 Bi-linearity of Commutator

For $\alpha, \beta \in \mathbb{C}$ and operators A, B, the commutator satisfies:

$$[\alpha A,\, \beta B] = \alpha\beta\, [A,\, B] \tag{7.28}$$

This immidiately leads to:

Principle 7.2: Canonical Commutation Relation

For the x operator and the D operator, the commutation relation is given by:

$$[x,\, -i\hbar D] = i\hbar = \left[x,\, -i\hbar\frac{\partial}{\partial x}\right] = \left[i\hbar\frac{\partial}{\partial p},\, p\right] \tag{7.29}$$

Let us remark on other properties:

Theorem 7.2.2 Commutator of General Function

For a general function $f(x)$, the commutator with the D operator is given by:

$$[D,\, f(x)] = \frac{df(x)}{dx} \tag{7.30}$$

Proof: Consider functions $f(x)$ and $g(x)$, then:

$$\begin{aligned} [D,\, f(x)]\, g(x) &= Df(x)g(x) - f(x)Dg(x) && (7.31)\\ &= \frac{df(x)}{dx}g(x) + f(x)\frac{dg(x)}{dx} - f(x)\frac{dg(x)}{dx} && (7.32)\\ &= \frac{df(x)}{dx}g(x) && (7.33)\\ \implies [D,\, f(x)] &= \frac{df(x)}{dx} && (7.34) \end{aligned}$$

□

Theorem 7.2.3 Commutator of General Operator

For a general operator F that is a function of D, the commutator with the x operator is given by:

$$[F(D),\, x] = \frac{dF(D)}{dD} \tag{7.35}$$

Proof: Assume that F is a function of D, and we can write it as a Taylor series:

$$F(D) = \sum_{n=0}^{\infty} a_n D^n \tag{7.36}$$

Then,

$$
\begin{aligned}
[F, x] &= Fx - xF && (7.37)\\
&= \sum_{n=0}^{\infty} a_n D^n x - x \sum_{n=0}^{\infty} a_n D^n && (7.38)\\
&= \sum_{n=0}^{\infty} a_n (D^n x - x D^n) && (7.39)\\
&= \sum_{n=0}^{\infty} a_n n D^{n-1} = \frac{dF}{dD} && (7.40)
\end{aligned}
$$

□

Finally, let us comment on the relationship with the **shift operator**:

Definition 7.4: Shift Operator

The **shift operator** S_a is defined as the operator that shifts a function by a constant a:

$$S_a[f](x) := f(x+a) \tag{7.41}$$

The commutator of the shift operator and the x operator is given by:

$$
\begin{aligned}
[S_a, x]\, f(x) &= S_a[x f(x)] - x S_a[f(x)] && (7.42)\\
&= (x+a) f(x+a) - x f(x+a) && (7.43)\\
&= a f(x+a) = a S_a[f(x)] && (7.44)
\end{aligned}
$$

Thus,

Theorem 7.2.4 Commutator of Shift Operator and x

The commutator of the shift operator S_a and the x operator is given by:

$$[S_a, x] = a S_a \implies S_a = e^{aD} \tag{7.45}$$

7.3 Variational Principle in Fluid Mechanics

7.3.1 Overview: Cauchy's Equation of Motion

The equation of motion of a point mass is given by Newton's second law:

$$m\frac{d^2\vec{r}}{dt^2} = \vec{F} \tag{7.46}$$

The equation of motion of a fluid (in general) is given by Cauchy's equation of motion:

$$\int_V \rho\, dV \frac{D\vec{u}}{Dt} = \oint_S dS\, \sigma\vec{n} + \int_V \rho\, dV \vec{\tilde{F}} \tag{7.47}$$

where

- V: Volume of the fluid
- S: Surface of V
- $\rho(\vec{r}, t)$: Density of the fluid at position $\vec{r}$ and time t
- $\vec{u}(\vec{r}, t)$: Velocity of the fluid at position $\vec{r}$ and time t
- $\vec{n}$: Normal vector of the surface S pointing outward
- $\sigma(\vec{r}, t)$: Stress tensor of the fluid
- $\vec{\tilde{F}}$: External force per unit volume acting on the infinitisimal volume dV of fluid

the stress tensor shows how the forces act on any surface within a fluid.

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \tag{7.48}$$

The Newtonian derivation of Cauchy's equation of motion is based on the conservation of the momentum of a fluid element:

$$\begin{aligned} (\frac{\partial p}{\partial t} + \Delta p \text{ flux per time}) &= (\Delta p \text{ from the surface force}) \\ &\quad + (\Delta p \text{ from the volume force}) \end{aligned} \tag{7.49}$$

$$\implies \quad \rho\, dV \frac{D\vec{u}}{Dt} = \nabla \cdot (\sigma\vec{n}) + \rho\, dV \vec{\tilde{F}} \tag{7.50}$$

[9]

7.3.2 Lagrange Derivative

Imagine that you want to measure the height of the water in a river at a certain point and time. Let us assume that you put a very light, small box with some measuring equipment on the water surface at some point $\vec{r}_i$ at time t_i. Then we put another box at another point $\vec{r}_f$ at the same time t_i, but we fix this box so that it does not move.

Since the water surface is not static, box 1 will be carried by the water flow. Let us denote the position of the box 1 at time t as $\vec{\xi}(\vec{r}_i, t; t_i)$, where the box 1 is placed at the point $\vec{r}_i$ at time t_i.

Let us denote the height of the water surface at a point $\vec{r}$ at time t by $X(\vec{r}, t)$. Then, we can write the height of the water surface at the positions of box 1 and 2 at time t as

$$X(\vec{\xi}, t) = X_1(t) \tag{7.51}$$
$$X(\vec{r}_f, t) = X_2(t) \tag{7.52}$$

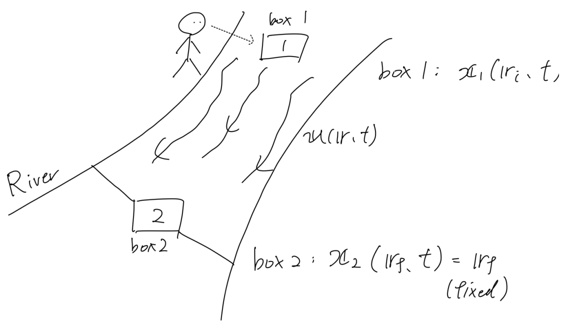

Figure 7.1: Boxes on a river

Finally, assume that the box 1 will reach $\vec{r}_f$ at time t_f.

Now, at initial time t_i, the change in height of the water surface is given by:

$$\begin{aligned} X(\vec{r}_f, t_i) - X(\vec{r}_i, t_i) &= \nabla X(\vec{r}_i, t_i) \cdot (\vec{r}_f - \vec{r}_i) && (7.53) \\ &= \nabla X(\vec{r}_i, t_i) \cdot \Delta\vec{r} && (7.54) \end{aligned}$$

assuming that $\Delta\vec{r} := \vec{r}_f - \vec{r}_i$ is small enough. On the other hand, for box 2, the change in height of the water surface in a time interval $\Delta t = t_f - t_i$ is given by:

$$X(\vec{r}_f, t_f) - X(\vec{r}_f, t_i) = \frac{\partial X(\vec{r}_f, t_i)}{\partial t}\Delta t \tag{7.55}$$

Thus, the total change of X measured by box 1, during Δt, moving at the velocity $\vec{u}$ is given by

$$\begin{aligned} X(\vec{r}_f, t_f) - X(\vec{r}_i, t_i) &= X(\vec{r}_f, t_f) \underbrace{- X(\vec{r}_f, t_i) + X(\vec{r}_f, t_i)}_{=0} - X(\vec{r}_i, t_i) && (7.56) \\ &= \frac{\partial X}{\partial t}\Delta t + \nabla X(\vec{r}_i, t_i) \cdot \Delta\vec{r} && (7.57) \end{aligned}$$

Now, since $\vec{r}_i = \vec{\xi}(\vec{r}_i, t_i)$ and $\vec{r}_f = \vec{\xi}(\vec{r}_i, t_f)$, we can rewrite the equation above:

$$X(\vec{\xi}(\vec{r}_i, t_f), t_f) - X(\vec{\xi}, t_i) = \frac{\partial X}{\partial t}\Delta t + \nabla X(\vec{\xi}, t_i) \cdot \Delta\vec{\xi} \tag{7.58}$$

where $\Delta\vec{\xi} = \vec{\xi}(\vec{r}_i, t_f) - \vec{\xi}(\vec{r}_i, t_i)$.

Using $\Delta t = t_f - t_i$, we can rewrite the equation above as:

$$\frac{X(\vec{\xi} + \Delta\vec{\xi}, t_i + \Delta t) - X(\vec{\xi}, t_i)}{\Delta t} = \frac{\partial X}{\partial t} + \nabla X(\vec{\xi}, t_i) \cdot \frac{\Delta\vec{\xi}}{\Delta t} \tag{7.59}$$

$$\tag{7.60}$$

Since the box 1 is a part of the fluid, the time derivative of $\vec{\xi}$ is the velocity of the fluid at the position $\vec{\xi}$:

$$\frac{\partial\vec{\xi}(\vec{r}_i, t)}{\partial t} = \vec{u}(\vec{\xi}, t) \tag{7.61}$$

Note:

In this derivative, we fix the starting position $\vec{r}_i$, we only care about the time evolution from the position $\vec{r}_i$. This is different from the total derivative, which takes into account a change in the starting position $\vec{r}_i$.

thus, by taking $\Delta t \to 0 (\implies \Delta x \to 0)$, we can rewrite the equation above as:

$$\frac{DX(\vec{\xi}, t_i)}{Dt} := \frac{\partial X(\vec{\xi}, t_i)}{\partial t} + \vec{u}(\vec{\xi}, t_i) \cdot \nabla X(\vec{\xi}, t_i) \tag{7.62}$$

where $\frac{D}{Dt}$ indicates that we are tracking the box 1 and its measurement of X. This is called the **Lagrange derivative**:

Definition 7.5: Lagrange Derivative

The Lagrange description in Eulerian perspective is given as the **Lagrange derivative**:

$$\frac{DX}{Dt} := \frac{\partial X}{\partial t} + \vec{u} \cdot \nabla X \tag{7.63}$$

which tracks the change of a physical quantity X with the flow.

Now, define a new quantity called the ”position function” $\vec{x}(\vec{r}, t)$:

$$\vec{x}(\vec{r}, t) := \vec{r} \tag{7.64}$$

If we measure this position function at the box 1 at $t = t_i$,

$$\vec{x}(\vec{r}_i, t_i) = \vec{r}_i \tag{7.65}$$

and notice that tracking the position of box 1 is equivalent to observing the time-evolved position $\xi(\vec{r}_i, t_i)$:

$$\frac{\partial\vec{\xi}}{\partial t} = \frac{D\vec{x}}{Dt} = \vec{u}(\vec{x}, t) \tag{7.66}$$

7.3.3 Derivation from Action Integral

In Newtonian mechanics, for a particle of mass m, the action integral S is given by:

$$S = \int dt\, L = \int dt\, m\mathcal{L}, \quad \mathcal{L} := \frac{L}{m} = \left[\frac{1}{2}\dot{\vec{x}}^2 - \tilde{V}(\vec{x})\right] \tag{7.67}$$

where $\mathcal{L}$ is the Lagrangian density, and

$$\tilde{V}(\vec{x}) = \frac{V(\vec{x})}{m} \tag{7.68}$$

by assuming a similar form of the Lagrangian density for a fluid, we can write the action integral for a fluid as:

$$S = \int dt\, L = \int dt \int_V \rho(\vec{x},t)\, dV(\vec{x},t)\, \mathcal{L}\left(\vec{x}, \frac{D\vec{x}}{Dt}\right) \tag{7.69}$$

where the Lagrangian density is

$$\mathcal{L}\left(\vec{x}(\vec{r},t), \frac{D\vec{x}(\vec{r},t)}{Dt}\right) = \frac{1}{2}\left(\frac{D\vec{x}(\vec{r},t)}{Dt}\right)^2 - \tilde{V}(\vec{x}(\vec{r},t)) + \frac{1}{\rho(\vec{r},t)}\nabla\cdot\sigma(\vec{r},t)\vec{x}(\vec{r},t) \tag{7.70}$$

Then the Euler-Lagrange equation for this Lagrangian density is given by:

$$\frac{\partial\mathcal{L}}{\partial\vec{x}} - \frac{D}{Dt}\left(\frac{\partial\mathcal{L}}{\partial(\frac{D\vec{x}}{Dt})}\right) = 0 \tag{7.71}$$

calculating the partial derivative gives:

$$\frac{\partial\mathcal{L}}{\partial\vec{x}} = \frac{1}{\rho}\nabla\cdot\sigma - \nabla\tilde{V}(\vec{x}) = \frac{1}{\rho}\nabla\cdot\sigma + \vec{\tilde{F}} \tag{7.72}$$

$$\frac{\partial\mathcal{L}}{\partial(\frac{D\vec{x}}{Dt})} = \frac{D\vec{x}}{Dt} = \vec{u}(\vec{x},t) \tag{7.73}$$

hence the Euler-Lagrange equation becomes:

$$\frac{D\vec{u}}{Dt} = \frac{1}{\rho}\nabla\cdot\sigma + \vec{\tilde{F}} \tag{7.74}$$

by integrating over the volume V, we get

$$\int_V \rho\, dV\frac{D\vec{u}}{Dt} = \int_V dV\nabla\cdot\sigma + \int_V \rho\, dV\vec{\tilde{F}} \tag{7.75}$$

$$\Longrightarrow \int_V \rho\, dV\frac{D\vec{u}}{Dt} = \int_V ds\,\vec{n}\cdot\sigma + \int_V \rho\, dV\vec{\tilde{F}} \tag{7.76}$$

7.3.4 Hamilton Formalism

Similarly to the point mass case, we should define the momentum density $\vec{\pi}(\vec{r},t)$ as the derivative of the Lagrangian density with respect to the velocity:

$$\vec{\pi}(\vec{r},t) = \frac{\partial \mathcal{L}}{\partial(\frac{D\vec{x}}{Dt})} \tag{7.77}$$

then the Hamiltonian density should be defined by the Legendre transformation:

$$\mathcal{H}(\vec{\pi},\vec{x}) = \vec{\pi} \cdot \frac{D\vec{x}}{Dt} - \mathcal{L}\left(\vec{x}, \frac{D\vec{x}}{Dt}\right) \tag{7.78}$$

$$= \frac{1}{2}\left(\frac{D\vec{x}}{Dt}\right)^2 + \tilde{V}(\vec{x}) - \frac{1}{\rho}\nabla \cdot \sigma\vec{x} \tag{7.79}$$

[10]

7.4 Physical Constants

All values are taken from **The NIST Reference on Constants, Units, and Uncertainty** [11].

Table 7.1: Physical Constants

Symbol	Constant	Value	Units
c	Speed of light in vacuum	$2.997\,924\,58 \times 10^{8}$	$\mathrm{m\,s^{-1}}$
e	Elementary charge	$-1.602\,176\,634 \times 10^{-19}$	C
G	Gravitational constant	$6.674\,30 \times 10^{-11}$	$\mathrm{N\,m^2\,kg^{-2}}$
h	Planck constant	$6.626\,070\,15 \times 10^{-34}$	$\mathrm{J\,s}$
$\hbar := \frac{h}{2\pi}$	Reduced Planck constant	$1.054\,571\,817 \times 10^{-34}$	$\mathrm{J\,s}$

Chapter 8

Textbooks

8.1 Physics Textbooks

8.1.1 General Textbooks

- **University Physics with Modern Physics** by Hugh Young and Roger Freedman [12]
- **Feynman Lectures on Physics** by Richard P. Feynman [13]

8.1.2 Analytical Mechanics

- **Classical Mechanics** by Goldstein, Poole, and Safko [14]
- **Mechanics** by Landau and Lifshitz [15]

8.1.3 Quantum Mechanics

- **Quantum Mechanics** by Leonard Schiff [16]
- **Introduction to Quantum Mechanics** by David Griffiths [17]
- **Modern Quantum Mechanics** by J. J. Sakurai and Jim Napolitano [18]

8.1.4 Particle Physics

- **Introduction to Elementary Particles** by David Griffiths [19]
- **An Introduction to Quantum Field Theory** by Michael Peskin and Daniel Schroeder [20]

8.2 Math Textbooks

Bibliography

[1] 前野昌弘. “最小作用の原理はどこからくるか？. ”[Online]. Available: http://irobutsu.a.la9.jp/fromRyukyu/wiki/index.php?%BA%C7%BE%AE%BA%EE%CD%D1%A4%CE%B8%B6%CD%FD%A4%CF%A4%C9%A4%B3%A4%AB%A4%E9%A4%AF%A4%EB%A4%AB%A1%A9.

[2] 広江克彦. “ルジャンドル変換. ”[Online]. Available: https://eman-physics.net/analytic/legendre.html.

[3] 渡辺悠樹, 解析力学. 共立出版, 2024, isbn: 9784320036314.

[4] 広江克彦. “汎関数微分. ”[Online]. Available: https://eman-physics.net/analytic/functional.html.

[5] A. Einstein, “The collected papers of albert einstein,” in J. Stachel, Ed. Princeton University Press, 1990, vol. 2, ch. On a Heuristic Point of View Concerning the Production and Transformation of Light.

[6] A. H. Compton, “Physical review letters,” *A Quantum Theory of the Scattering of X-rays by Light Elements*, vol. 21, no. 483, 1923.

[7] L. de Broglie, “Recherches sur la théorie des quanta,” Ph.D. dissertation, Paris, 1925.

[8] N. Bohr, “On the constitution of atoms and molecules,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 26, no. 151, 1913.

[9] 広江克彦. “流体の運動方程式. ”[Online]. Available: https://eman-physics.net/fluid/eq_of_motion.html.

[10] 鈴木康夫. “流体力学における最小作用の原理（提案）. ”[Online]. Available: https://hooktail.sub.jp/fluid/PrincipleofLeastActionforFluidMechanics/.

[11] National Institute of Standards and Technology, *Fundamental physical constants*, 2024. [Online]. Available: https://physics.nist.gov/cuu/Constants/index.html.

[12] H. D. Young, R. A. Freedman, T. Sandin, and A. L. Ford, *University physics*. Addison-Wesley Reading, MA, 1996, vol. 9.

[13] R. P. Feynman, R. B. Leighton, and M. Sands. “The feynman lectures on physics. ”[Online]. Available: https://www.feynmanlectures.caltech.edu.

[14] H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*. Pearson Education, 2002, isbn: 9788131758915.

[15] L. D. Landau and E. M. Lifshitz, *Mechanics*. Butterworth-Heinemann, 1976, isbn: 978-0750628969.

[16] L. Schiff, *Quantum Mechanics*. McGraw-Hill, 1968, isbn: 978-0070552876.

[17] D. J. Griffiths, *Introduction to quantum mechanics*. Cambridge University Press, 2018, isbn: 978-1107189638.

[18] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, 2nd. Cambridge University Press, 2017, isbn: 978-1-108-42241-3.

[19] D. J. Griffiths, *Introduction to Elementary Particles*, Second edition. Wiley-VCH, 2008, isbn: 9783527406012.

[20] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Frontiers in Physics), First edition. Westview Press; CRC Press, 2018, isbn: 9780201503975.