Imperial College London

NOTES

IMPERIAL COLLEGE LONDON

DEPARTMENT OF PHYSICS

Advanced Classical Physics

Author:
Chen Huang

Email:
chen.huang23@imperial.ac.uk

Date: May 17, 2024

Contents

1 Lagrangian Mechanics

1.1 Action principle

1.1.1 Fermat's principle

The idea of the 'principle of least action' has its origin in Fermat's principle in optics, according to which light follows the **shortest optical path**, i.e., the path of shortest time to reach its destination.

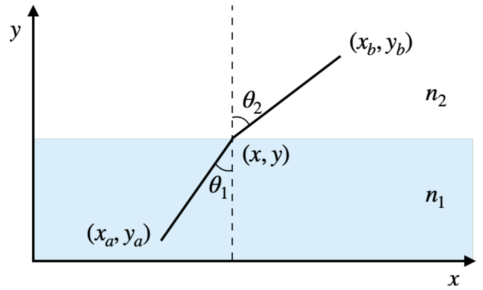

Figure 1: Snell's law of refraction for light passing through media of different indices of refraction.

According to Figure 1, consider a light ray from point (x_a, y_a) to (x_b, y_b), the optical path length follows

$$T(x) = \frac{n_1}{c}\sqrt{(x_a - x)^2 + (y_a - y)^2} + \frac{n_2}{c}\sqrt{(x_b - x)^2 + (y_b - y)^2}. \tag{1}$$

According to Fermat's principle, we need to find the minimum of this quantity.

$$\begin{aligned}\frac{\partial T(x)}{\partial x} &= \frac{n_1}{c}\frac{x - x_a}{\sqrt{(x_a - x)^2 + (y_a - y)^2}} + \frac{n_2}{c}\frac{x - x_b}{\sqrt{(x_b - x)^2 + (y_b - y)^2}} \\ &= -\frac{n_1}{c}\sin\theta_1 + \frac{n_2}{c}\sin\theta_2 = 0,\end{aligned} \tag{2}$$

which shows the **Snell's law**:

$$\boxed{\frac{\sin\theta_2}{\sin\theta_1} = \frac{n_1}{n_2}.} \tag{3}$$

What happens if the refractive index is a function of space $n(x, y)$? The Fermat's action becomes

$$T = \int \mathrm{d}T = \int \frac{n(x,y)}{c}\sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2}. \tag{4}$$

Suppose we parameterise the trajectory of the particle by a monotonic parameter λ:

$$\boldsymbol{r}(\lambda) = (x(\lambda), y(\lambda)), \tag{5}$$

such that

$$x(\lambda_a) = x_a, \quad x(\lambda_b) = x_b, \quad y(\lambda_a) = y_a, \quad y(\lambda_b) = y_b. \tag{6}$$

Then the Fermat's action is

$$T = \int_{\lambda_a}^{\lambda_b} \mathrm{d}\lambda \frac{n(x(\lambda), y(\lambda))}{c}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\lambda}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\lambda}\right)^2}. \tag{7}$$

1.1.2 Euler-Lagrange equations

In configuration space (the space of coordinates and velocities), the action S is defined as an integral over a function $L(x, \dot{x}, t)$ known as the Lagrangian, as

$$S[x(t)] = \int_{t_a}^{t_b} \mathrm{d}t L(x(t), \dot{x}(t), t). \tag{8}$$

For standard conservative systems the Lagrangian is simply the difference of the kinetic energy T and the potential energy V, i.e.,

$$L = T - V. \tag{9}$$

The actual physical trajectory is the function x that minimises the action subject to the boundary conditions $x(t_a) = x_a$ and $x(t_b) = x_b$. This perturbation changes the action by an amount δS given by

$$\delta S[x(t)] = \int_{t_a}^{t_b} \mathrm{d}t \left[\frac{\partial L}{\partial x} \delta x(t) + \frac{\partial L}{\partial \dot{x}} \delta \dot{x}(t) \right]. \tag{10}$$

For the second term, we may integrate by parts,

$$\int_{t_a}^{t_b} \mathrm{d}t \frac{\partial L}{\partial \dot{x}} \delta \dot{x}(t) = \left[\frac{\partial L}{\partial \dot{x}} \delta x(t) \right]_{t_a}^{t_b} - \int_{t_a}^{t_b} \mathrm{d}t \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{x}} \right) \delta x(t) = - \int_{t_a}^{t_b} \mathrm{d}t \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{x}} \right) \delta x(t). \tag{11}$$

Substituting this result into the expression (10) for δS, we have

$$\delta S = \int_{t_a}^{t_b} \mathrm{d}t \left[\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{x}} \right) \right] \delta x(t). \tag{12}$$

In order that S be stationary, i.e., $\delta S = 0$, we require

$$\boxed{\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{x}} \right) = 0.} \tag{13}$$

This is known as the **Euler-Lagrange equation**, and it is the equation of motion in the Lagrangian formulation of mechanics.

1.1.3 Back to Fermat

In the gauge $\lambda = y$, the total time was given to

$$T = \int_{y_a}^{y_b} \mathrm{d}y \frac{n(x(y), y)}{c} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}y} \right)^2 + 1}. \tag{14}$$

The corresponding Lagrangian is

$$L(x(y), \dot{x}(y), y) = \frac{n(x(y), y)}{c} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}y} \right)^2 + 1}. \tag{15}$$

The Euler-Lagrange equation for this problem is

$$\frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{\partial L}{\partial \frac{\mathrm{d}x}{\mathrm{d}y}}\right) = \frac{\partial L}{\partial x}, \tag{16}$$

which gives

$$\frac{\mathrm{d}}{\mathrm{d}y}\left[n(x,y)\frac{\frac{\mathrm{d}x}{\mathrm{d}y}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^2+1}}\right] = \frac{\partial n(x,y)}{\partial x}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^2+1}. \tag{17}$$

It simplifies a lot in the case where the refractive index is just a function of y, i.e., $\frac{\partial n}{\partial x} = 0$. Then

$$\frac{\mathrm{d}}{\mathrm{d}y}\left[n(y)\frac{\frac{\mathrm{d}x}{\mathrm{d}y}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^2+1}}\right] = 0, \tag{18}$$

which is easily solved as

$$n(y)\frac{\frac{\mathrm{d}x}{\mathrm{d}y}}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^2+1}} = n(y)\frac{\mathrm{d}x}{\sqrt{(\mathrm{d}x)^2+(\mathrm{d}y)^2}} = n(y)\sin(\theta(y)) = A, \tag{19}$$

where A a constant.

1.2 Generalized coordinates

We are free to choose whichever set of variables we want to parameterise the state of the system. The variables are called **generalised coordinates** and usually denoted by q_i. The action of Lagrangian $L(q_i, \dot{q}_i, t)$ can be expressed as

$$S = \int \mathrm{d}t L(q_i, \dot{q}_i, t), \tag{20}$$

and

$$\delta S = \sum_{i=1}^{N}\int \mathrm{d}t\left[\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\right] = 0. \tag{21}$$

Each coordinate satisfies the corresponding Euler-Lagrange equation

$$\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0, \quad \forall i = 1, \ldots, N. \tag{22}$$

1.3 Conservation laws

1.3.1 Momentum conservation

For any generalised coordinate q_i, we define the generalised momentum p_i by

$$\boxed{p_i = \frac{\partial L}{\partial \dot{q}_i}}. \tag{23}$$

The Euler-Lagrange equation gives

$$\frac{\partial L}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\mathrm{d}p_i}{\mathrm{d}t}, \tag{24}$$

which shows whenever the Lagrangian L does not depend *explicitly* on q_i, the corresponding generalised momentum p_i is conserved:

$$\boxed{\frac{\partial L}{\partial q_i} = 0 \quad \Leftrightarrow \quad p_i \text{ is conserved.}} \tag{25}$$

1.3.2 Energy (or Hamiltonian) conservation

First, we consider the total time derivative of the Lagrangian:

$$\begin{aligned} \frac{\mathrm{d}L}{\mathrm{d}t} &= \sum_i \frac{\partial L}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i + \frac{\partial L}{\partial t} \\ &= \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i + \frac{\partial L}{\partial t} \\ &= \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\dot{q}_i\right) + \frac{\partial L}{\partial t}. \end{aligned} \tag{26}$$

Rearranging this equation, we have

$$\frac{\partial L}{\partial t} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L\right) = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i p_i\dot{q}_i - L\right) = -\frac{\mathrm{d}H}{\mathrm{d}t}. \tag{27}$$

This gives that whenever the Lagrangian does not depend *explicitly* on t, the Hamiltonian is conserved

$$\boxed{\frac{\partial L}{\partial t} = 0 \quad \Leftrightarrow \quad H = \sum_i p_i\dot{q}_i - L \text{ is conserved.}} \tag{28}$$

1.4 Constraints and number of degrees of freedom

1.4.1 Lagrange multipliers

Starting with a Lagrangian L and a constraint function $f(q_i, \dot{q}_i, t)$, we define a new Lagrangian $\tilde{L}$:

$$\tilde{L}(q_i, \dot{q}_i, \lambda, t) = L(q_i, \dot{q}_i, t) + \lambda f(q_i, \dot{q}_i, t). \tag{29}$$

Figure 2: A helter skelter.

The Euler-Lagrange equation for λ is given by

$$\frac{\delta \tilde{S}}{\delta \lambda} = \frac{\partial \tilde{L}}{\partial \lambda} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \tilde{L}}{\partial \dot{\lambda}}\right) = f(q_i, \dot{q}_i, t), \tag{30}$$

and therefore it imposes the constraint $f = 0$ independently of what λ actually is.

1.4.2 Example: helter skelter

A child of mass m slides down a helter skelter (Figure 2), the Lagrangian is given by

$$L = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2\right) - mgz, \tag{31}$$

where

$$z = h - \alpha\theta, \qquad r = \beta\theta. \tag{32}$$

α and β are positive constants and h is the height of the helter skelter. We take the angle θ to go from 0 to infinity as the trajectory winds around multiple times. The constraints are

$$C_1 = z - h + \alpha\theta = 0, \qquad C_2 = r - \beta\theta = 0. \tag{33}$$

We have two approaches to solve the constrained system. (1) The first one is to solve the constraints and then substitute them back into the action, thereby reducing the action. (2) The second one involves using multipliers and working on an extended configuration space $(q_i, \dot{q}_i, t, \lambda)$ to solve the Euler-Lagrange equation.

(1) **Reduced Lagrangian**
Substitute the constraints into the Lagrangian:

$$L(\theta, t) = \frac{1}{2}m\left(\beta^2\dot{\theta}^2 + \beta^2\theta^2\dot{\theta}^2 + \alpha^2\dot{\theta}^2\right) - mg(h - \alpha\theta). \tag{34}$$

Then we can write he Euler-Lagrange equation for θ

$$m\frac{\mathrm{d}}{\mathrm{d}t}\left[\left(\alpha^2 + \beta^2\right)\dot{\theta} + \beta^2\theta^2\dot{\theta}\right] = m\beta^2\theta\dot{\theta}^2 + mg\alpha \tag{35}$$

(2) **'Extended' Lagrangian**
The 'extended' Lagrangian is

$$\tilde{L} = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2\right) - mgz + \lambda_1(z - h + \alpha\theta) + \lambda_2(r - \beta\theta). \tag{36}$$

The Euler-Lagrange equations are:

$$\text{for } r, \quad mr\dot{\theta}^2 - m\ddot{r} = -\lambda_2, \tag{37}$$

$$\text{for } \theta, \quad m\frac{\mathrm{d}}{\mathrm{d}t}\left(r^2\dot{\theta}\right) = \alpha\lambda_1 - \beta\lambda_2, \tag{38}$$

$$\text{for } z, \quad m\ddot{z} + mg = \lambda_1, \tag{39}$$

$$\text{for } \lambda_1, \quad z = h - \alpha\theta, \tag{40}$$

$$\text{for } \lambda_2, \quad r = \beta\theta. \tag{41}$$

After Rearranging the equations, we get the equivalent result:

$$m\frac{\mathrm{d}}{\mathrm{d}t}\left[\left(\alpha^2 + \beta^2\right) + \beta^2\theta^2\right]\dot{\theta} = m\beta^2\theta\dot{\theta}^2 + mg\alpha. \tag{42}$$

1.5 Normal Modes

1.5.1 Kinetic matrix

Now we focus on the Euler-Lagrange equation again. Using the chain rule, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \sum_j \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j}\dot{q}_j + \sum_j \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}\ddot{q}_j + \frac{\partial^2 L}{\partial \dot{q}_i \partial t} = \frac{\partial L}{\partial q_i}. \tag{43}$$

We can rewrite the expression in this form

$$\boxed{\sum_j \mathcal{Z}^{ij}\ddot{q}_j + \mathcal{F}^i = 0,} \tag{44}$$

where the **kinetic matrix** $\mathcal{Z}^{ij}$ and the vector $\mathcal{F}^i$ are both functions of the coordinates:

$$\mathcal{Z}^{ij}(q_1, \cdots, q_N, \dot{q}_1, \cdots, \dot{q}_N, t) = \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} = \mathcal{Z}^{ji}, \tag{45}$$

$$\mathcal{F}^i(q_1, \cdots, q_N, \dot{q}_1, \cdots, \dot{q}_N, t) = \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j}\dot{q}_j + \frac{\partial^2 L}{\partial \dot{q}_i \partial t} - \frac{\partial L}{\partial q_i}. \tag{46}$$

So long as the symmetric matrix $\mathcal{Z}$ is non-degenerate, i.e. $\det(\mathcal{Z}) \neq 0$,

$$\ddot{q}_i = -(\mathcal{Z})^{-1}_{ij}\mathcal{F}^j. \tag{47}$$

There are cases however where the matrix $\mathcal{Z}$ is 'degenerate' and $\det(\mathcal{Z}) = 0$ means that not all coordinates q_i are independent.

We consider the situation where particles can move around their equilibrium positions, which means that the kinetic energy is a **quadratic homogeneous function** of the generalised velocities. We can then write it as

$$T = \frac{1}{2}\sum_{ij} a_{ij}(q_1, \cdots, q_N)\dot{q}_i\dot{q}_j = \frac{1}{2}\dot{\boldsymbol{q}}^T A\dot{\boldsymbol{q}}. \tag{48}$$

Since the Lagrangian is given by

$$L = T - V(q_1, \cdots, q_N, t), \tag{49}$$

where the potential does not depend on the velocities $\dot{q}_i$, we have

$$\begin{aligned}\mathcal{Z}^{ij} =&\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_j} = \frac{\partial^2 T}{\partial\dot{q}_i\partial\dot{q}_j} = \frac{\partial^2}{\partial\dot{q}_i\partial\dot{q}_j}\left(\frac{1}{2}a_{\alpha\beta}\dot{q}_\alpha\dot{q}_\beta\right)\\ =&\frac{\partial}{\partial\dot{q}_j}\left(\frac{1}{2}a_{\alpha\beta}\frac{\partial\dot{q}_\alpha}{\partial\dot{q}_i}\dot{q}_\beta + \frac{1}{2}a_{\alpha\beta}\dot{q}_\alpha\frac{\partial\dot{q}_\beta}{\partial\dot{q}_i}\right) = \frac{\partial}{\partial\dot{q}_j}\left(\frac{1}{2}a_{\alpha\beta}\delta_{i\alpha}\dot{q}_\beta + \frac{1}{2}a_{\alpha\beta}\dot{q}_\alpha\delta_{i\beta}\right)\\ =&\frac{\partial}{\partial\dot{q}_j}\left(\frac{1}{2}a_{i\beta}\dot{q}_\beta + \frac{1}{2}a_{\alpha i}\dot{q}_\alpha\right) = \frac{1}{2}a_{ij}\dot{q}_j + \frac{1}{2}a_{ji}\dot{q}_j = a_{ij}.\end{aligned} \tag{50}$$

The kinetic matrix is then given directly by the coefficients a_{ij}.

1.5.2 Equilibrium points

From the calculation in above, we know that

$$\frac{\partial L}{\partial\dot{q}_i} = a_{ij}\dot{q}_j. \tag{51}$$

Then the Euler-Lagrangian equation becomes

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial\dot{q}_i}\right) - \frac{\partial L}{\partial q_i} =&\frac{\mathrm{d}}{\mathrm{d}t}\left(a_{ij}\dot{q}_j\right) - \frac{\partial T}{\partial q_i} + \frac{\partial V}{\partial q_i}\\ =&\frac{\mathrm{d}}{\mathrm{d}t}\left(a_{ij}\dot{q}_j\right) - \frac{\partial}{\partial q_i}\left(\frac{1}{2}a_{jk}\dot{q}_j\dot{q}_k\right) + \frac{\partial V}{\partial q_i}\\ =&a_{ij}\ddot{q}_j - \frac{1}{2}\frac{\partial a_{jk}}{\partial q_i}\dot{q}_j\dot{q}_k + \frac{\partial V}{\partial q_i} = 0.\end{aligned} \tag{52}$$

In the equilibrium point, $\dot{q}_i$ is a constant, so we require

$$\left.\frac{\partial V}{\partial q_i}\right|_{q_i=q_{0i}} = 0, \qquad \forall i = 1, 2, \cdots, N. \tag{53}$$

1.5.3 Small oscillations

We can write the coordinates as $q_i(t) = q_{0i} + \delta q_i(t)$. The potential energy in the form of Taylor expansion:

$$\begin{aligned}V(\boldsymbol{q}_0 + \delta\boldsymbol{q}) =&V(\boldsymbol{q}_0) + \left.\frac{\partial V}{\partial q_i}\right|_{\boldsymbol{q}_0}\delta q_i + \frac{1}{2}\left.\frac{\partial^2 V}{\partial q_i\partial q_j}\right|_{\boldsymbol{q}_0}\delta q_i\delta q_j + \mathcal{O}(\delta q^3)\\ =&V(\boldsymbol{q}_0) + \frac{1}{2}\left.\frac{\partial^2 V}{\partial q_i\partial q_j}\right|_{\boldsymbol{q}_0}\delta q_i\delta q_j.\end{aligned} \tag{54}$$

The action is given by

$$\begin{aligned}
S &= \int \mathrm{d}t \left[\frac{1}{2}a_{ij}(\boldsymbol{q}_0)\frac{\mathrm{d}}{\mathrm{d}t}(q_{0i}+\delta q_i)\frac{\mathrm{d}}{\mathrm{d}t}(q_{0j}+\delta q_j) - V(\boldsymbol{q}_0+\delta\boldsymbol{q})\right] \\
&= \int \mathrm{d}t \left[\frac{1}{2}a_{ij}(\boldsymbol{q}_0)\delta\dot{q}_i\delta\dot{q}_j - \left(V(\boldsymbol{q}_0) + \left.\frac{\partial V}{\partial q_i}\right|_{\boldsymbol{q}_0}\delta q_i + \frac{1}{2}\left.\frac{\partial^2 V}{\partial q_i \partial q_j}\right|_{\boldsymbol{q}_0}\delta q_i \delta q_j + \mathcal{O}(\delta q^3)\right)\right] \\
&= \int \mathrm{d}t \left[\frac{1}{2}a_{ij}(\boldsymbol{q}_0)\delta\dot{q}_i\delta\dot{q}_j - \frac{1}{2}b_{ij}(\boldsymbol{q}_0)\delta q_i\delta q_j\right] - \int \mathrm{d}t V(\boldsymbol{q}_0),
\end{aligned} \tag{55}$$

where

$$b_{ij} = \frac{\partial^2 V}{\partial q_i \partial q_j}. \tag{56}$$

The second term is just a constant, so for small fluctuations all that matters is the quadratic part, which we refer to as the action for the quadratic fluctuations

$$\begin{aligned}
S_{(2)} &= \int \mathrm{d}t \left(\frac{1}{2}a_{ij}\delta\dot{q}_i\delta\dot{q}_j - \frac{1}{2}b_{ij}\delta q_i \delta q_j\right) \\
&= \int \mathrm{d}t \left(\frac{1}{2}\delta\dot{\boldsymbol{q}}^T\mathbf{A}\delta\dot{\boldsymbol{q}} - \frac{1}{2}\delta\boldsymbol{q}^T\mathbf{B}\delta\boldsymbol{q}\right)
\end{aligned} \tag{57}$$

where the matrix $\mathbf{A}$ and $\mathbf{B}$ have the components $\mathbf{A}_{ij} = a_{ij}$ and $\mathbf{B}_{ij} = b_{ij}$.

The Euler-Lagrange equation for the fluctuations is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{(2)}}{\partial \delta\dot{q}_i}\right) = \frac{\partial L_{(2)}}{\partial \delta q_i}, \tag{58}$$

which is

$$a_{ij}\delta\ddot{q}_j = -b_{ij}\delta q_j \qquad \text{or} \qquad \mathbf{A}\delta\ddot{\boldsymbol{q}} = -\mathbf{B}\delta\boldsymbol{q}. \tag{59}$$

If $\det(\mathbf{A}) \neq 0$, then

$$\delta\ddot{\boldsymbol{q}} = -\mathbf{K}\delta\boldsymbol{q}, \tag{60}$$

with

$$\mathbf{K} = \mathbf{A}^{-1}\mathbf{B}. \tag{61}$$

1.5.4 Gram-Schmidt diagonalization

Since the matrix $\mathbf{A}$ is symmetric (we may assume this from the outset), we can always diagonalize it using a $N \times N$ orthogonal matrix $\mathbf{O}$ ($\mathbf{O}^T\mathbf{O} = \mathbb{1}$) so that

$$\mathbf{A}_D = \mathbf{O}^T\mathbf{A}\mathbf{O}, \tag{62}$$

with $\mathbf{A}_D$ diagonal. The kinetic term

$$\begin{aligned}
T =& \frac{1}{2}\delta\dot{\boldsymbol{q}}^T\mathbf{A}\delta\dot{\boldsymbol{q}} = \frac{1}{2}\delta\dot{\boldsymbol{q}}^T\mathbf{O}\mathbf{A}_D\mathbf{O}^T\delta\dot{\boldsymbol{q}} = \frac{1}{2}\delta\dot{\tilde{\boldsymbol{q}}}^T\mathbf{A}_D\delta\dot{\tilde{\boldsymbol{q}}} \\
=& \frac{1}{2}\delta\dot{\tilde{q}}_i(a_D)_{ij}\delta\dot{\tilde{q}}_j = \frac{1}{2}(a_D)_{11}\delta\dot{\tilde{q}}_1^2 + \frac{1}{2}(a_D)_{22}\delta\dot{\tilde{q}}_2^2 + \cdots + \frac{1}{2}(a_D)_{NN}\delta\dot{\tilde{q}}_N^2,
\end{aligned} \tag{63}$$

where

$$\delta\tilde{\boldsymbol{q}} = \mathbf{O}^T \delta \boldsymbol{q}. \tag{64}$$

To put the kinetic term in diagonal and *normalized* form, we require

$$\delta q_i = \sqrt{(a_D)_{ii}^{-1}} \delta Q_i. \tag{65}$$

This is the statement that there is a diagonal matrix $\mathbf{W}$ whose matrix elements are

$$\mathbf{W}_{ij} = \sqrt{(a_D)_{ii}^{-1}} \delta_{ij}, \tag{66}$$

such that

$$\delta\tilde{\boldsymbol{q}} = \mathbf{W} \delta \boldsymbol{Q}, \tag{67}$$

for which

$$\mathbf{W}^T \mathbf{A}_D \mathbf{W} = \mathbb{1}. \tag{68}$$

These two operations amount to saying that there is a matrix $\mathbf{S} = \mathbf{OW}$ such that

$$\mathbf{S}^T \mathbf{A}_D \mathbf{S} = \mathbb{1}. \tag{69}$$

the action for the fluctuations becomes

$$\begin{aligned} S_{(2)} &= \int \mathrm{d}t \left(\frac{1}{2} \delta \dot{\boldsymbol{Q}}^T \mathbf{S}^T \mathbf{A} \mathbf{S} \delta \dot{\boldsymbol{Q}} - \frac{1}{2} \delta \dot{\boldsymbol{Q}}^T \mathbf{S}^T \mathbf{B} \mathbf{S} \delta \dot{\boldsymbol{Q}} \right) \\ &= \int \mathrm{d}t \left(\frac{1}{2} \delta \dot{\boldsymbol{Q}}^T \delta \dot{\boldsymbol{Q}} - \frac{1}{2} \delta \dot{\boldsymbol{Q}}^T \mathbf{k} \delta \dot{\boldsymbol{Q}} \right), \end{aligned} \tag{70}$$

where

$$\mathbf{k} = \mathbf{S}^T \mathbf{B} \mathbf{S}. \tag{71}$$

This is the canonically normalized form for the fluctuations.

1.5.5 Cholesky decomposition

The Cholesky decomposition of a *real Hermitian positive*-definite[1] matrix $\mathbf{A}$, is a decomposition of the form

$$\mathbf{A} = \mathbf{L} \mathbf{L}^T, \tag{72}$$

where $\mathbf{L}$ is a left triangular matrix with real and positive diagonal entries

$$\mathbf{L} = \begin{pmatrix} \# & 0 & 0 & 0 \\ \# & \# & 0 & 0 \\ \# & \# & \# & 0 \\ \# & \# & \# & \# \end{pmatrix}. \tag{73}$$

So the kinetic term can be written as

$$T = \frac{1}{2} \delta \dot{\boldsymbol{q}} \mathbf{A} \delta \dot{\boldsymbol{q}} = \frac{1}{2} \delta \dot{\boldsymbol{q}} \mathbf{L} \mathbf{L}^T \delta \dot{\boldsymbol{q}} = \frac{1}{2} \delta \dot{\boldsymbol{Q}}^T \delta \dot{\boldsymbol{Q}}, \tag{74}$$

[1] 'Positive' means the eigenvalues of $\mathbf{A}$ are all positive.

where

$$\delta \boldsymbol{Q} = \mathbf{L}^T \delta \boldsymbol{q}. \tag{75}$$

The action can be written in

$$\begin{aligned} S &= \int \mathrm{d}t \left(\frac{1}{2} \delta \dot{\boldsymbol{q}} \mathbf{A} \delta \dot{\boldsymbol{q}} - \frac{1}{2} \delta \boldsymbol{q}^T \mathbf{B} \delta \boldsymbol{q} \right) \\ &= \int \mathrm{d}t \left(\frac{1}{2} \delta \dot{\boldsymbol{Q}} \delta \dot{\boldsymbol{Q}} - \frac{1}{2} \delta \boldsymbol{Q}^T \mathbf{k} \delta \boldsymbol{Q} \right) \\ &= \int \mathrm{d}t \left(\frac{1}{2} \delta \dot{Q}_i \delta \dot{Q}_i - \frac{1}{2} k_{ij} \delta Q_i \delta Q_j \right), \end{aligned} \tag{76}$$

where

$$\mathbf{k} = \mathbf{L}^{-1} \mathbf{B} (\mathbf{L}^T)^{-1}. \tag{77}$$

The Euler-Lagrange equation for Q_i is given by

$$\delta \ddot{Q}_i = -k_{ij} \delta Q \qquad \text{or} \qquad \delta \ddot{\boldsymbol{Q}} = -\mathbf{k} \delta \boldsymbol{Q}. \tag{78}$$

We can look for solutions of the form

$$\delta \boldsymbol{Q} = \mathrm{e}^{i\omega t} \delta \boldsymbol{Q}_\omega, \tag{79}$$

then we have the eigenfunction equation with the matrix $\mathbf{k}$ and eigenvalue ω^2

$$\mathbf{k} \delta \boldsymbol{Q}_\omega = \omega^2 \delta \boldsymbol{Q}_\omega, \tag{80}$$

which means that each normal coordinate $\boldsymbol{Q}_{\omega_\alpha}$ oscillates independently of all others with its own normal frequency ω_α^2.

1.5.6 Example: double pendulum

Consider a double pendulum, with a second pendulum hanging from the first as depicted in Figure 3. The kinetic energy of the system is

$$\begin{aligned} T =& \frac{1}{2} M \dot{\boldsymbol{R}}^2 + \frac{1}{2} m (\dot{\boldsymbol{R}} + \dot{\boldsymbol{r}})^2 \\ =& \frac{1}{2} M \dot{\boldsymbol{R}}^2 + \frac{1}{2} m \left(\dot{\boldsymbol{R}}^2 + \dot{\boldsymbol{r}}^2 + 2 \dot{\boldsymbol{R}} \dot{\boldsymbol{r}} \right) \\ =& \frac{1}{2} M R^2 \dot{\theta}^2 + \frac{1}{2} m R^2 \dot{\theta}^2 + \frac{1}{2} m r^2 \dot{\phi}^2 + m R r \dot{\theta} \dot{\phi} \cos(\phi - \theta). \end{aligned} \tag{81}$$

The potential energy is

$$\begin{aligned} V =& M g (-R \cos\theta) + m g (-R \cos\theta - r \cos\phi) \\ =& -(M + m) g R \cos\theta - m g r \cos\phi. \end{aligned} \tag{82}$$

The equilibrium solution is $\theta_0 = \phi_0 = 0$, so θ and ϕ can be expressed by

$$\theta(t) = \delta\theta(t), \qquad \phi(t) = \delta\phi(t) \tag{83}$$

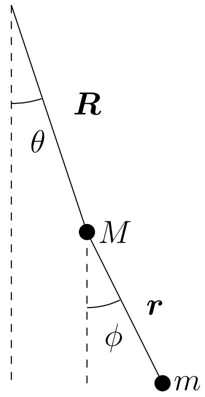

Figure 3: Double pendulum

then we have

$$\begin{aligned} L =&\frac{1}{2}(M+m)R^2\delta\dot{\theta}^2 + \frac{1}{2}mr^2\delta\dot{\phi}^2 + mRr\delta\dot{\theta}\delta\dot{\phi} - \frac{1}{2}(M+m)gR\delta^2\theta - \frac{1}{2}mgr\delta\phi^2 \\ =&\frac{1}{2}MR^2\delta\dot{\theta}^2 + \frac{1}{2}m\left(R\delta\dot{\theta} + r\delta\dot{\phi}\right)^2 - \frac{1}{2}(M+m)gR\delta^2\theta - \frac{1}{2}mgr\delta\phi^2. \end{aligned} \tag{84}$$

We define that

$$\delta Q_1 = \sqrt{M}R\delta\theta, \qquad \delta Q_2 = \sqrt{m}(R\delta\theta + r\delta\phi), \tag{85}$$

so the Lagrangian becomes

$$L_{(2)} =\frac{1}{2}\delta\dot{Q}_1^2 + \frac{1}{2}\delta\dot{Q}_2^2 - \frac{1}{2}\frac{(M+m)g}{MR}\delta Q_1^2 - \frac{1}{2}\frac{mg}{Mr}(\delta Q_1 - \delta Q_2)^2, \tag{86}$$

and the potential energy is given by

$$V = -\frac{1}{2}\delta\boldsymbol{Q}^T\mathbf{k}\delta\boldsymbol{Q} = -\frac{1}{2}\begin{pmatrix}\delta Q_1 & \delta Q_2\end{pmatrix}\mathbf{k}\begin{pmatrix}\delta Q_1 \\ \delta Q_2\end{pmatrix}, \tag{87}$$

where

$$\mathbf{k} = \begin{pmatrix} \frac{(M+m)g}{MR} + \frac{mg}{Mr} & -\frac{mg}{Mr} \\ -\frac{mg}{Mr} & \frac{mg}{Mr} \end{pmatrix} = \frac{mg}{Mr}\begin{pmatrix} \frac{(M+m)r}{mR} + 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{88}$$

To simplify $\mathbf{k}$, we choose $R = r$ and $M = m$, then we have

$$\mathbf{k} = \frac{g}{R}\begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}. \tag{89}$$

The eigenvalues ω_1^2 and ω_2^2 satisfies

$$\det(\mathbf{k} - \omega^2\mathbb{1}) = \frac{g}{R}\left[(3-\omega^2)(1-\omega^2) - 1\right] = 0. \tag{90}$$

Solve this, and we get

$$\omega_{1,2}^2 = (2 \pm \sqrt{2})\frac{g}{R} > 0. \tag{91}$$

The associated eigenvectors are $\boldsymbol{Q}_\pm = (1 \pm \sqrt{2}, -1)$. Go back to the original coordinates θ, ϕ, we know that

$$Q_1 \sim \theta, \qquad Q_2 \sim \theta + \phi. \tag{92}$$

So the normal modes are

$$(\theta, \phi)_\pm = (1 \pm \sqrt{2}, -(2 \pm \sqrt{2})) \sim (1, \pm\sqrt{2}) \tag{93}$$

1.5.7 Summary

The Lagrangian describing the dynamics of a small oscillation system, subject to a slight variation $\boldsymbol{q}$, is given by

$$L = \frac{1}{2}\dot{\boldsymbol{q}}^T \cdot \dot{\boldsymbol{q}} - \frac{1}{2}\boldsymbol{q}^T \boldsymbol{k}\boldsymbol{q}. \tag{94}$$

Here, $\boldsymbol{k}$ represents the stiffness matrix. By solving the characteristic equation

$$\det(\boldsymbol{k} - \lambda \mathbb{1}) = 0, \tag{95}$$

we determine the eigenvalues, denoted by λ. The system is stable at a stationary point if these eigenvalues are positive ($\lambda > 0$), and the square of the **normal frequencies** ω^2 corresponds to these eigenvalues, i.e., $\omega^2 = \lambda$. The associated **normal modes** are represented by the eigenvectors $\boldsymbol{q}$.

1.6 Symmetries and conservation laws

1.6.1 Noether's theorem

Let us parameterize our action by arbitrary coordinates q_i

$$S = \int \mathrm{d}t L(q, \dot{q}). \tag{96}$$

Let λ denote the parameter describing the continuous transformation. As this is a global symmetry, this parameter is constant. Then under an infinitesimal transformation we assume the coordinates transform as

$$\delta q_i = F_i(q, \dot{q})\delta\lambda. \tag{97}$$

Then the Lagrangian transforms as

$$\delta L = \frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}}\delta \dot{q} = \left(\frac{\partial L}{\partial q_i}F_i + \frac{\partial L}{\partial \dot{q}_i}\dot{F}_i\right)\delta\lambda. \tag{98}$$

If this is a symmetry, then this must be a total derivative, then

$$\frac{\partial L}{\partial q_i}F_i + \frac{\partial L}{\partial \dot{q}_i}\dot{F}_i = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}F_i\right) = \frac{\mathrm{d}A}{\mathrm{d}t}, \quad \delta L = \frac{\mathrm{d}A}{\mathrm{d}t}\delta\lambda. \tag{99}$$

Noether then performed a clever trick. She make λ a function of time:

$$\delta q_i = F_i(q_i, \dot{q}_i)\delta\lambda(t). \tag{100}$$

then the Lagrangian transforms as

$$\begin{aligned} \delta L(q_i, \dot{q}_i, t) =& \frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i \\ =& \left(\frac{\partial L}{\partial q_i}F_i + \frac{\partial L}{\partial \dot{q}_i}\dot{F}_i\right)\delta\lambda + \frac{\partial L}{\partial \dot{q}_i}F_i\frac{\mathrm{d}\delta\lambda}{\mathrm{d}t} = \frac{\mathrm{d}A}{\mathrm{d}t}\delta\lambda + \frac{\partial L}{\partial \dot{q}_i}F_i\frac{\mathrm{d}\delta\lambda}{\mathrm{d}t}, \end{aligned} \tag{101}$$

and the integration

$$\delta S = \int \mathrm{d}t \delta L = \int \mathrm{d}t \left(\frac{\mathrm{d}A}{\mathrm{d}t}\delta\lambda + \frac{\partial L}{\partial \dot{q}_i} F_i \frac{\mathrm{d}\delta\lambda}{\mathrm{d}t} \right) = \int \mathrm{d}t \delta\lambda \frac{\mathrm{d}}{\mathrm{d}t}\left(A - \frac{\partial L}{\partial \dot{q}_i} F_i \right). \tag{102}$$

Define the **Noether charge** as

$$\boxed{G = \frac{\partial L}{\partial \dot{q}_i} F_i - A,} \tag{103}$$

then we have

$$\frac{\mathrm{d}G}{\mathrm{d}t} = 0, \quad \text{(up to the boundary term)}, \tag{104}$$

which is known as the **Noether's theorem**: *For every continuous symmetry, there exist a conserved quantity G conserved in time.*

Example
Consider the action

$$S = \int \mathrm{d}t \frac{1}{2} m \dot{q}^2 \tag{105}$$

and the transform $q' = q + \lambda$, *i.e.*, $\delta q = \delta\lambda$ with $F = 1$ and $A = 0$.

$$G = \frac{\partial L}{\partial \dot{q}} F - A = \frac{\partial L}{\partial \dot{q}} = p \tag{106}$$

which means the corresponding generalized momentum component p to be a constant of motion.

1.6.2 Hamiltonian as the Noether charge for time translations

Consider a system which is time translation invariant meaning that the action is invariant (un to boundary terms) under the symmetry $t \to t + \delta t$. Such a transformation induces a change of coordinates

$$\delta q_i(t) = q_i(t + \delta t) - q_i(t) = \dot{q}_i \delta t. \tag{107}$$

Similarly the Lagrangian transforms as

$$\delta L = L(q(t + \delta t), \dot{q}(t + \delta t), t + \delta t) - L(q, \dot{q}, t) = \frac{\mathrm{d}L}{\mathrm{d}t}\delta t, \tag{108}$$

from which we infer

$$A = L. \tag{109}$$

Thus the conserved charge implied by Noether's theorem associated with the symmetry of time translation invariance is

$$G = \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L = p_i \dot{q}_i - L = H \tag{110}$$

which we recognize to be the Hamiltonian.

2 Hamiltonian Mechanics

2.1 Hamiltonian formulation

In Lagrangian, we use coordinates and velocities $(q, \dot{q})$ to parameterise the system, while in Hamiltonian, we use coordinates and momenta (q, p) to parameterise the system.

Lagrangian		Hamiltonian
$L = L(q, \dot{q}, t)$	$\longleftrightarrow$	$H = H(q, p, t)$
	$p = \dfrac{\partial L}{\partial \dot{q}}$	

2.1.1 From Lagrangian to Hamiltonian

The total derivative of Lagrangian can be expressed as

$$\begin{aligned} \mathrm{d}L(q, \dot{q}, t) =& \frac{\partial L}{\partial q_i}\mathrm{d}q_i + \frac{\partial L}{\partial \dot{q}_i}\mathrm{d}\dot{q}_i + \frac{\partial L}{\partial t}\mathrm{d}t \\ =& \frac{\partial L}{\partial q_i}\mathrm{d}q_i + p_i\mathrm{d}\dot{q}_i + \frac{\partial L}{\partial t}\mathrm{d}t. \end{aligned} \tag{111}$$

And we find that

$$\mathrm{d}(L - p_i\dot{q}_i) = \frac{\partial L}{\partial q_i}\mathrm{d}q_i - \dot{q}_i\mathrm{d}p_i + \frac{\partial L}{\partial t}\mathrm{d}t. \tag{112}$$

Therefore, we can define the **Hamiltonian function** H by Legendre transformation

$$\boxed{H = \sum_i p_i\dot{q}_i - L.} \tag{113}$$

2.1.2 Hamilton's equations

The total derivative of Hamiltonian can be expressed as

$$\begin{aligned} \mathrm{d}H =& -\frac{\partial L}{\partial q_i}\mathrm{d}q_i + \dot{q}_i\mathrm{d}p_i - \frac{\partial L}{\partial t}\mathrm{d}t \\ =& \frac{\partial H}{\partial q_i}\mathrm{d}q_i + \frac{\partial H}{\partial p_i}\mathrm{d}p_i + \frac{\partial H}{\partial t}\mathrm{d}t. \end{aligned} \tag{114}$$

Comparing the two expressions, we have

$$\boxed{\frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i} = -\dot{p}_i, \quad \frac{\partial H}{\partial p_i} = \dot{q}_i, \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}.} \tag{115}$$

These results are known as **Hamilton's equations**.

2.1.3 Conservation of the Hamiltonian

We can also calculate the time derivative of the Hamiltonian

$$\begin{aligned} \frac{\mathrm{d}H}{\mathrm{d}t} =& \frac{\partial H}{\partial t} + \sum_i \left(\frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i \right) \\ =& \frac{\partial H}{\partial t} + \sum_i \left(\frac{\partial H}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q_i} \right) = \frac{\partial H}{\partial t}, \end{aligned} \tag{116}$$

which we summarize below

$$\boxed{\frac{\partial H}{\partial t} = \frac{\mathrm{d}H}{\mathrm{d}t}.} \tag{117}$$

We can conclude that if the Hamiltonian has no *explicit* time dependence (*i.e.* $\partial H/\partial t = 0$), then the Hamiltonian is conserved.

2.2 Hamilton's principle of the least action

The phase space action is

$$S = \int \mathrm{d}t \left(\sum_i p_i \frac{\mathrm{d}q_i}{\mathrm{d}t} - H(q,p,t) \right). \tag{118}$$

Let us now perform the variation of this action in which we regard $q_i(t)$ and $p_i(t)$ as independent functions

$$\begin{aligned} \delta S =& \int_{t_1}^{t_2} \mathrm{d}t \sum_i \left(\delta p_i \frac{\mathrm{d}q_i}{\mathrm{d}t} + p_i \frac{\mathrm{d}\delta q_i}{\mathrm{d}t} - \frac{\partial H}{\partial q_i}\delta q_i - \frac{\partial H}{\partial p_i}\delta p_i \right) \\ =& \sum_i p_i \delta q_i \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \mathrm{d}t \sum_i \left[\delta p_i \left(\frac{\mathrm{d}q_i}{\mathrm{d}t} - \frac{\partial H}{\partial p_i} \right) + \left(-\frac{\mathrm{d}p_i}{\mathrm{d}t} - \frac{\partial H}{\partial q_i} \right) \delta q_i \right]. \end{aligned} \tag{119}$$

Thus demanding that $\delta S = 0$ for variations which vanish as the boundary $\delta q_i(t_1) = \delta q_i(t_2) = 0$, which implies Hamilton's equations:

$$\boxed{\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}.} \tag{120}$$

2.3 Poisson brackets

The total time derivative of $A(q,p,t)$ is

$$\begin{aligned} \frac{\mathrm{d}A}{\mathrm{d}t} =& \frac{\partial A}{\partial t} + \sum_i \left(\frac{\partial A}{\partial q_i}\dot{q}_i + \frac{\partial A}{\partial p_i}\dot{p}_i \right) \\ =& \frac{\partial A}{\partial t} + \sum_i \left(\frac{\partial A}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial H}{\partial q_i} \right) \\ =& \frac{\partial A}{\partial t} + \{A, H\}, \end{aligned} \tag{121}$$

where the quantity

$$\{A, H\} = \sum_i \left(\frac{\partial A}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial H}{\partial q_i} \right), \tag{122}$$

is known as **Poisson bracket**. The Poisson brackets for coordinates and momenta are

$$\{q_i, q_j\} = \sum_k \left(\frac{\partial q_i}{\partial q_k}\frac{\partial q_j}{\partial p_k} - \frac{\partial q_i}{\partial p_k}\frac{\partial q_j}{\partial q_k} \right) = 0, \tag{123}$$

$$\{p_i, p_j\} = \sum_k \left(\frac{\partial p_i}{\partial q_k}\frac{\partial p_j}{\partial p_k} - \frac{\partial p_i}{\partial p_k}\frac{\partial p_j}{\partial q_k} \right) = 0, \tag{124}$$

$$\{q_i, p_j\} = \sum_k \left(\frac{\partial q_i}{\partial q_k}\frac{\partial p_j}{\partial p_k} - \frac{\partial q_i}{\partial p_k}\frac{\partial p_j}{\partial q_k} \right) = \delta_{ij}. \tag{125}$$

We can also write Hamilton's equations in terms of Poisson brackets as

$$\dot{p}_i = \{p_i, H\}, \quad \dot{q}_i = \{q_i, H\}. \tag{126}$$

2.3.1 Properties of Poisson brackets

1. Anti-symmetric

$$\{A, B\} = -\{B, A\}. \tag{127}$$

2. Satisfying the Leibniz/product rule

$$\{A, BC\} = \{A, B\}C + B\{A, C\}. \tag{128}$$

3. Jacobi identity

$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0. \tag{129}$$

2.4 Canonical transformations

The canonical transformation means a transformation which *preserve* Poisson brackets. In other words, the transformation

$$q \to Q(q, p), \quad p \to P(q, p), \tag{130}$$

take the same Poisson brackets, i.e.,

$$\{f, g\}_{Q,P} = \{f, g\}_{q,p}, \tag{131}$$

is known as the **canonical transformation**. In fact, it turns out that it is enough to check that the Poisson brackets for coordinates and momenta are unchanged under the transformation, i.e.,

$$\{Q_i, Q_j\}_{q,p} = \{P_i, P_j\}_{q,p} = 0, \quad \{Q_i, P_j\}_{q,p} = \delta_{ij}, \tag{132}$$

or the opposite one

$$\{q_i, q_j\}_{Q,P} = \{p_i, p_j\}_{Q,P} = 0, \quad \{q_i, p_j\}_{Q,P} = \delta_{ij}. \tag{133}$$

Any set of variables that satisfy these conditions are called **canonical conjugates**.

2.4.1 Infinitesimal canonical transformations

The general form of an infinitesimal canonical transformation is

$$\delta q_i = \{q_i, G(q,p,t)\}\delta\lambda, \quad \delta p_i = \{p_i, G(q,p,t)\}\delta\lambda. \tag{134}$$

We define the phase space coordinates

$$\rho_I = \begin{cases} q_I & I = 1, \cdots, N, \\ p_{I-N} & I = N+1, \cdots, 2N, \end{cases} \tag{135}$$

then the form of the Poisson brackets are

$$\{\rho_I, \rho_J\} = \Omega_{IJ} = \begin{cases} 0 & 1 \le I \le N, 1 \le J \le N, \\ \delta_{I,J-N} & 1 \le I \le N, N+1 \le J \le 2N, \\ -\delta_{(I-N)J} & N+1 \le I \le 2N, 1 \le J \le N, \\ 0 & N+1 \le I \le 2N, N+1 \le J \le 2N. \end{cases} \tag{136}$$

Consider a infinitesimal change of coordinates $\rho_I \to \rho_I + \delta\rho_O$. The Poisson bracket will change as

$$\delta\{\rho_I, \rho_J\} = \delta\Omega_{IJ} = \{\rho_I, \delta\rho_J\} + \{\delta\rho_I, \rho_J\} = 0. \tag{137}$$

According to the Jacobi identity $\{A, \{B, C\} + \{B, \{C, A\}\} + \{C, \{A, B\}\}$, we have

$$\{\rho_I, \{\rho_J, G\}\} + \{\rho_J, \{G, \rho_I\}\} + \{G, \underbrace{\{\rho_I, \rho_J\}}_{=\Omega_{IJ}}\} = 0. \tag{138}$$

Thus

$$\{\rho_I, \{\rho_J, G\}\} + \{\rho_J, \{G, \rho_I\}\} = 0, \quad \forall G(q,p,t). \tag{139}$$

2.5 Symmetries and conservation laws

2.5.1 Noether's theorem

In Lagrange frame, the infinitesimal form of the continuous global symmetry transformation is

$$\delta q_i = F_i(q, \dot{q}, t)\delta\lambda. \tag{140}$$

In Hamilton frame, we assume the existence of a generating function $G(q,p)$ under which the coordinates and momenta change as

$$\delta q_i = \{q_i, G(q,p)\}\delta\lambda = \frac{\partial G}{\partial p_i}\delta\lambda, \quad \delta p_i = \{p_i, G(q,p)\}\delta\lambda = -\frac{\partial G}{\partial q_i}\delta\lambda, \tag{141}$$

where $\delta\lambda$ is an infinitesimal parameter. If the Hamiltonian does not change under such a transformation, then the equations of motion are clearly *invariant* and so we have a symmetry. Explicitly, under this transformation the Hamiltonian changes as

$$\delta H = \frac{\mathrm{d}H}{\mathrm{d}\lambda}\delta\lambda, \tag{142}$$

where

$$\frac{\mathrm{d}H}{\mathrm{d}\lambda} = \sum_i \left(\frac{\partial H}{\partial q_i}\frac{\partial q_i}{\partial \lambda} + \frac{\partial H}{\partial p_i}\frac{\partial p_i}{\partial \lambda} \right) = \sum_i \left(\frac{\partial H}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial G}{\partial q_i} \right) = \{H, G\}. \tag{143}$$

This means that the Hamiltonian is invariant under the transformation if $\{H, G\} = 0$. But if this is the case, then G is conserved because then

$$\frac{\mathrm{d}G}{\mathrm{d}t} = \{G, H\} = 0. \tag{144}$$

This is of course none other than Noether's theorem: *If the Hamiltonian is invariant under a continuous transformation, then the generator G of the transformation is a conserved charge.*

2.5.2 Noether charge as the generator

In phase space, the infinitessimal form of the continuous gloval symmetry transformation transforms both the coordiniates and momenta

$$\delta q_i = F_i(q,p)\delta\lambda, \quad \delta p_i = G_i(q,p)\delta\lambda. \tag{145}$$

The corresponding generator in phase space expressed as

$$G = \sum_i \frac{\partial L}{\partial \dot{q}_i} F_i - A = \sum_i p_i F_i(p,q) - A. \tag{146}$$

2.5.3 Generator of rotations in two dimensions

Consider a two-dimensional problem with Cartesian coordinates (x, y)

$$H = \frac{p_x^2 + p_y^2}{2m} + V(\sqrt{x^2+y^2}). \tag{147}$$

Consider an infinitesimal rotation $\delta\theta$, the coordinates and momenta transform as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\delta\theta & -\sin\delta\theta \\ \sin\delta\theta & \cos\delta\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y\delta\theta \\ y + x\delta\theta \end{pmatrix} = \begin{pmatrix} x + \delta x \\ y + \delta y \end{pmatrix}, \tag{148}$$

$$\begin{pmatrix} p_x' \\ p_y' \end{pmatrix} = \begin{pmatrix} \cos\delta\theta & -\sin\delta\theta \\ \sin\delta\theta & \cos\delta\theta \end{pmatrix} \begin{pmatrix} p_x \\ p_y \end{pmatrix} = \begin{pmatrix} p_x - p_y\delta\theta \\ p_y + p_x\delta\theta \end{pmatrix} = \begin{pmatrix} p_x + \delta p_x \\ y + \delta p_y \end{pmatrix} \tag{149}$$

So we have the changes

$$\delta x = -y\delta\theta, \quad \delta y = x\delta\theta, \quad \delta p_x = -p_y\delta\theta, \quad \delta p_y = p_x\delta\theta. \tag{150}$$

Since L is invariant, i.e., $\delta L = 0$, the generator is easily seen to be

$$G = p_x F_x + p_y F_y = p_x(-y) + p_y x = x p_y - y p_x = L_z. \tag{151}$$

To check explicitly

$$\delta x = \frac{\partial G}{\partial p_x}\delta\lambda = -y\delta\lambda, \qquad \delta y = \frac{\partial G}{\partial p_y}\delta\lambda = x\delta\lambda, \tag{152}$$

$$\delta p_x = -\frac{\partial G}{\partial x}\delta\lambda = -p_y\delta\lambda, \quad \delta p_y = -\frac{\partial G}{\partial y}\delta\lambda = p_x\delta\lambda. \tag{153}$$

2.5.4 Generator of time translations

Noether charge associated with time translations is the Hamiltonian, so we choose $G = H$ as the generator.

$$\delta q_i = \{q_i, H\}\delta t = \dot{q}_i \delta t, \quad \delta p_i = \{p_i, H\}\delta t = \dot{p}_i \delta t. \tag{154}$$

2.6 Hamilton-Jacobi formulation

2.6.1 Schwinger's formulation of Hamilton's principle

If we evaluate the action

$$S = \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda \left[\sum_i p_i \frac{\mathrm{d}q_i}{\mathrm{d}\lambda} - H(q,p,t)\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right], \tag{155}$$

with the variation $q_i \to q_i + \delta q_i$, $p_i \to p_i + \delta p_i$ and $t \to t + \delta t$. Follows by the chain rule that

$$\begin{aligned}
\delta S &= \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda \sum_i \left[\delta p_i \frac{\mathrm{d}q_i}{\mathrm{d}\lambda} + p_i \frac{\mathrm{d}\delta q_i}{\mathrm{d}\lambda} - H\frac{\mathrm{d}\delta t}{\mathrm{d}\lambda} - \left(\frac{\partial H}{\partial q_i}\delta q_i + \frac{\partial H}{\partial p_i}\delta p_i + \frac{\partial H}{\partial t}\delta t\right)\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right] \\
&= \left(\sum_i p_i \delta q_i\right)\Bigg|_{\lambda_1}^{\lambda_2} - (H\delta t)\big|_{\lambda_1}^{\lambda_2} + \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda \\
&\quad \times \sum_i \left[\delta p_i \frac{\mathrm{d}q_i}{\mathrm{d}\lambda} - \delta q_i \frac{\mathrm{d}p_i}{\mathrm{d}\lambda} + \delta t \frac{\mathrm{d}H}{\mathrm{d}\lambda} - \left(\frac{\partial H}{\partial q_i}\delta q_i + \frac{\partial H}{\partial p_i}\delta p_i + \frac{\partial H}{\partial t}\delta t\right)\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right] \\
&= \left(\sum_i p_i \delta q_i\right)\Bigg|_{\lambda_1}^{\lambda_2} - (H\delta t)\big|_{\lambda_1}^{\lambda_2} + \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda \\
&\quad \times \sum_i \left[\delta p_i \left(\frac{\mathrm{d}q_i}{\mathrm{d}\lambda} - \frac{\partial H}{\partial p_i}\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right) + \delta q_i \left(-\frac{\mathrm{d}p_i}{\mathrm{d}\lambda} - \frac{\partial H}{\partial q_i}\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right) + \delta t \left(\frac{\mathrm{d}H}{\mathrm{d}\lambda} - \frac{\partial H}{\partial t}\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right)\right].
\end{aligned} \tag{156}$$

If we set $\lambda = t$ then

$$\begin{aligned}
\delta S &= \left(\sum_i p_i(t_2)\delta q_i(t_2) - H(t_2)\delta t_2\right) - \left(\sum_i p_i(t_1)\delta q_i(t_1) - H(t_1)\delta t_1\right) \\
&\quad + \int_{t_1}^{t_2} \mathrm{d}\lambda \sum_i \left[\delta p_i \left(\frac{\mathrm{d}q_i}{\mathrm{d}t} - \frac{\partial H}{\partial p_i}\right) + \delta q_i \left(-\frac{\mathrm{d}p_i}{\mathrm{d}t} - \frac{\partial H}{\partial q_i}\right) + \delta t \left(\frac{\mathrm{d}H}{\mathrm{d}t} - \frac{\partial H}{\partial t}\right)\right].
\end{aligned} \tag{157}$$

Again demanding that the variation only comes from the boundary terms we find in addition to Hamilton equations

$$\frac{\mathrm{d}H(q,p,t)}{\mathrm{d}t} = \frac{\partial H(q,p,t)}{\partial t}. \tag{158}$$

Thus on the solutions of the equations of motion we have the boundary variations

$$\delta S = \left(\sum_i p_i(t_2)\delta q_i(t_2) - H(t_2)\delta t_2\right) - \left(\sum_i p_i(t_1)\delta q_i(t_1) - H(t_1)\delta t_1\right). \tag{159}$$

2.6.2 Hamilton-Jacobi equation

If we evaluate the action on a solution of its equations of motion, the answer will depend on the initial and final coordinates $q_1 = q(t_1)$, $q_2 = q(t_2)$ and times t_1 and t_2. The total derivative of $S = S(q_2, t_2; q_1, t_1)$ is

$$\mathrm{d}S = \sum_i \frac{\partial S}{\partial q_i(t_2)}\mathrm{d}q_i(t_2) + \frac{\partial S}{\partial t_2}\mathrm{d}t_2 + \sum_i \frac{\partial S}{\partial q_i(t_1)}\mathrm{d}q_i(t_1) + \frac{\partial S}{\partial t_2}\mathrm{d}t_1. \tag{160}$$

Comparing the equations we see that

$$p_i(t_2) = \frac{\partial S}{\partial q_i(t_2)}, \quad H(t_2) = -\frac{\partial S}{\partial t_2}, \quad p_i(t_1) = -\frac{\partial S}{\partial q_i(t_1)}, \quad H(t_1) = \frac{\partial S}{\partial t_1}. \tag{161}$$

For example, given a Hamiltonian for a non-relativistic particle in one dimension

$$H = \frac{p^2}{2m} + V(x,t), \tag{162}$$

using the above equations at time $t = t_2$ we have

$$\boxed{-\frac{\partial S}{\partial t} = \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + V(x,t).} \tag{163}$$

This is known as the **Hamilton-Jacobi equation**.

2.6.3 Constants of motion

The variation of the action

$$\delta S = \left(\sum_i p_i(t_2)\delta q_i(t_2) - H(t_2)\delta t_2\right) - \left(\sum_i p_i(t_1)\delta q_i(t_1) - H(t_1)\delta t_1\right). \tag{164}$$

Consider a system in which the Hamiltonian is independent of t

$$\frac{\partial H}{\partial t} = 0 \quad \Rightarrow \quad \frac{\mathrm{d}H}{\mathrm{d}t} = 0,\ H(t_1) = H(t_2) = E. \tag{165}$$

Now we define the new action via the Legendre transform with respect to the initial data

$$\tilde{S} = S - Et_1, \tag{166}$$

then

$$\begin{aligned}\delta\tilde{S} =& \delta S - t_1\delta E - E\delta t_1 \\ =& \sum_i p_i(t_2)\delta q_i(t_2) - E\delta t_2 - \sum_i p_i(t_1)\delta q_i(t_1) + E\delta t_1 - t_1\delta E - E\delta t_1 \\ =& \sum_i p_i(t_2)\delta q_i(t_2) - E\delta t_2 - \sum_i p_i(t_1)\delta q_i(t_1) - t_1\delta E.\end{aligned} \tag{167}$$

From which we know $S = S(q_2, t_2; q_1; E)$. We thus conclude that

$$\frac{\delta \tilde{S}}{\delta E} = -t_1. \tag{168}$$

Example
The simplest example is a free particle in one dimension

$$H = \frac{p^2}{2m} = E, \quad q = x. \tag{169}$$

Since the energy is conserved, we will perform a Legendre transformation with respect to the initial time

$$\tilde{S} = S - Et_1, \tag{170}$$

and we have

$$\frac{\partial \tilde{S}}{\partial x_2} = p, \qquad \frac{\partial \tilde{S}}{\partial t_2} = -E \tag{171}$$

whose solutions are

$$\tilde{S} = -Et_2 + W(x_2, x_1). \tag{172}$$

The Hamilton-Jacobi equation at the final time $t_2 = t$, $x(t_2) = x$ is then

$$E = \frac{1}{2m}\left(\frac{\partial \tilde{S}}{\partial x_2}\right)^2 = \frac{1}{2m}\left(\frac{\partial W}{\partial x_2}\right)^2. \tag{173}$$

The equation is easy to solve as

$$W = \pm\sqrt{2mE}(x_2 - x_1). \tag{174}$$

Hence

$$\tilde{S} = -Et_2 \pm \sqrt{2mE}(x_2 - x_1) \tag{175}$$

So the action is

$$\begin{aligned} S =& - E(t_2 - t_1) \pm \sqrt{2mE}(x_2 - x_1) \\ =& - \frac{p^2}{2m}(t_2 - t_1) + p(x_2 - x_1). \end{aligned} \tag{176}$$

2.6.4 Perturbation theory

Consider a potential of the form

$$V = V_0 + gV_1, \tag{177}$$

where $|gV_1| \ll |V_0|$. The Hamilton-Jacobi equation for V_0

$$-\frac{\partial S_0}{\partial t} = \frac{1}{2m}\left(\frac{\partial S_0}{\partial x}\right)^2 + V_0(x). \tag{178}$$

Now if we write $S = S_0 + gS'$ and substitute into the full Hamilton-Jacobi equation

$$\frac{\partial S'}{\partial t} + \frac{1}{m}\frac{\partial S_0}{\partial x}\frac{\partial S'}{\partial x} = -\frac{g}{2m}\left(\frac{\partial S'}{\partial x}\right)^2 - V_1(x) \approx -V_1(x). \tag{179}$$

We can develop a perturbative expansion

$$S = S_0 + gS' = S_0 + gS_1 + g^2 S_2 + g^3 S_3 + \cdots = \sum_{n=0}^{\infty} g^n S_n. \tag{180}$$

Example
Consider the simple case where $V_0 = 0$ so that S_0 is the action for a free particle

$$S_0(x, t; x_0, t_0) = -\frac{p^2}{2m}(t - t_0) + p(x - x_0). \tag{181}$$

So we have

$$\frac{\partial S_0}{\partial x} = p, \tag{182}$$

and the equation for first order perturbation S_1 is

$$\frac{\partial S_1}{\partial t} + \frac{p}{m}\frac{\partial S_1}{\partial x} = -V_1(x). \tag{183}$$

Perform the change of variables $x = X + \frac{pt}{m}$,

$$\begin{aligned} \mathrm{d}S_1 &= \left(\frac{\partial S_1}{\partial x}\right)_t \mathrm{d}x + \left(\frac{\partial S_1}{\partial t}\right)_x \mathrm{d}t \\ &= \left(\frac{\partial S_1}{\partial X}\right)_t \left(\mathrm{d}X + \frac{p}{m}\mathrm{d}t\right) + \left(\frac{\partial S_1}{\partial t}\right)_X \mathrm{d}t \\ &= \left(\frac{\partial S_1}{\partial X}\right)_t \mathrm{d}X + \left[\frac{p}{m}\left(\frac{\partial S_1}{\partial X}\right)_t + \left(\frac{\partial S_1}{\partial t}\right)_X\right]\mathrm{d}t \\ &= \left(\frac{\partial S_1}{\partial X}\right)_t \mathrm{d}X - V_1(X)\mathrm{d}t. \end{aligned} \tag{184}$$

So

$$\left(\frac{\partial S_1}{\partial t}\right)_X = -V_1\left(x + \frac{pt}{m}, t\right) \quad \Rightarrow \quad S_1 = -\int_{t_0}^{t} \mathrm{d}t V_1\left(x + \frac{pt}{m}, t\right). \tag{185}$$

The next order perturbation is then

$$\frac{\partial S_2}{\partial t} = -\frac{1}{2m}\left(\frac{\partial S_1}{\partial X}\right)^2, \tag{186}$$

whose solution is

$$S_2 = -\frac{1}{2m}\int_{t_0}^{t} \mathrm{d}t' \left(\int_{t_0}^{t'} \mathrm{d}t' V_1(X + pt'/m, t')\right)^2. \tag{187}$$

Together, the solution of the H-J equation to second order in perturbation is

$$\begin{aligned} S = & -\frac{p^2}{2m}(t-t_0) + p(x-x_0) - g\int_{t_0}^{t} \mathrm{d}t V_1\left(x + \frac{pt}{m}, t\right) \\ & - \frac{g^2}{2m}\int_{t_0}^{t} \mathrm{d}t' \left(\int_{t_0}^{t'} \mathrm{d}t' V_1(X + pt'/m, t')\right)^2 + \ldots . \end{aligned} \tag{188}$$

2.7 Constraints

The action in phase space

$$S = \int \mathrm{d}t \left[\sum_i p_i \frac{\mathrm{d}q}{\mathrm{d}t} - H(q,p)\right], \tag{189}$$

include a set of constraints $C_\alpha(q,p) = 0$ with Lagrangian multiples λ_α, and the action becomes

$$S = \int \mathrm{d}t \left[\sum_i p_i \frac{\mathrm{d}q}{\mathrm{d}t} - H(q,p) - \sum_\alpha \lambda_\alpha C_\alpha\right]. \tag{190}$$

Now we define the effective Hamiltonian

$$H^*(q,p) = H(q,p) + \sum_\alpha \lambda_\alpha C_\alpha. \tag{191}$$

The equations of motion of H^* become

$$\boxed{\dot{q}_i = \{q_i, H^*\} = \frac{\partial H^*}{\partial p_i}, \quad \dot{p}_i = \{p_i, H^*\} = -\frac{\partial H^*}{\partial q_i}.} \tag{192}$$

Consider a system with one constraint $C_1(q,p) = 0$, then

$$S = \int \mathrm{d}t \left(\sum_i p_i \dot{q}_i - H - \lambda_1 C_1\right), \tag{193}$$

and

$$\dot{C}_1 = \{C_1, H^*\} = \{C_1, H + \lambda_1 C_1\} = \{C_1, H\} = C_2. \tag{194}$$

The secondary constraint C_2 not automatically equals to 0. We can write the secondary effective Hamiltonian as

$$H^{**} = H + \lambda_1 C_1 + \lambda_2 C_2. \tag{195}$$

and

$$\dot{C}_1 = \{C_1, H\} + \lambda_1\{C_1, C_1\} + \lambda_2\{C_1, C_2\} = C_2 + \lambda_2\{C_1, C_2\}, \tag{196}$$

$$\dot{C}_2 = \{C_2, H\} + \lambda_1\{C_2, C_1\} + \lambda_2\{C_2, C_2\} = \{C_2, H\} + \lambda_1\{C_2, C_1\}. \tag{197}$$

If $\{C_1, C_2\} = 0$ and $\{C_2, H\} \neq 0$, then we write the tertiary constraint $C_3 = \{C_2, H\}$ and repeat the process. If $\{C_1, C_2\} \neq 0$ and $\{C_2, H\} \neq 0$, then we have

$$\lambda_1 = \frac{\{C_2, H\}}{\{C_1, C_2\}}, \quad \lambda_2 = 0. \tag{198}$$

Example

Consider the 2D harmonic oscillator

$$H = \frac{1}{2}p_x^2 + \frac{1}{2}p_y^2 + \frac{1}{2}\omega^2 x^2 + \frac{1}{2}\omega^2 y^2, \tag{199}$$

with constrain $C_1 = p_y - \frac{1}{2}\beta x^2 = 0$.

$$C_2 = \{C_1, H\} = -\frac{\partial C_1}{\partial p_y}\frac{\partial H}{\partial y} + \frac{\partial C_1}{\partial x}\frac{\partial H}{\partial p_x} = -\omega^2 y - \beta x p_x. \tag{200}$$

The secondary constraint C_2 not automatically equals to 0.

$$\{C_2, H\} = \frac{\partial C_2}{\partial y}\frac{\partial H}{\partial p_y} + \frac{\partial C_2}{\partial x}\frac{\partial H}{\partial p_x} - \frac{\partial C_2}{\partial p_x}\frac{\partial H}{\partial x} = -\omega^2 p_y - \beta p_x^2 + \beta\omega^2 x^2, \tag{201}$$

$$\{C_1, C_2\} = \beta^2 x^2 + \omega^2 > 0 \tag{202}$$

So we have

$$\lambda_1 = \frac{-\omega^2 p_y - \beta p_x^2 + \beta\omega^2 x^2}{\beta^2 x^2 + \omega^2}, \quad \lambda_2 = 0. \tag{203}$$

2.7.1 Variations at fixed energy

We choose to focus on paths which have a fixed energy $H = E$, and the constrained action is

$$S = \int_{t_1}^{t_2} \mathrm{d}t \left[\boldsymbol{p}\cdot\dot{\boldsymbol{r}} - \frac{\boldsymbol{p}^2}{2m} - \lambda\left(\frac{\boldsymbol{p}^2}{2m} + V(\boldsymbol{r}) - E\right)\right], \tag{204}$$

with

$$\dot{\boldsymbol{r}} = \{\boldsymbol{r}, H^*\} = -\frac{\partial H^*}{\partial \boldsymbol{p}} = (1+\lambda)\frac{\boldsymbol{p}}{m} \quad \Rightarrow \quad \boldsymbol{p} = \frac{m}{1+\lambda}\dot{\boldsymbol{r}}, \tag{205}$$

and substituting back in gives

$$S = \int_{t_1}^{t_2} \mathrm{d}t \left[\frac{m}{2(1+\lambda)}\dot{\boldsymbol{r}}^2 - (1+\lambda)V(\boldsymbol{r}) + \lambda E\right]. \tag{206}$$

The equation for λ is

$$\frac{m}{2(1+\lambda)^2}\dot{\boldsymbol{r}}^2 = E - V \quad \Rightarrow \quad 1 + \lambda = \sqrt{\frac{m\dot{\boldsymbol{r}}^2}{2(E-V)}}. \tag{207}$$

Substituting back in we obtain

$$\begin{aligned} S = &\int_{t_1}^{t_2} \mathrm{d}t(-E) + \int_{t_1}^{t_2} \mathrm{d}t\sqrt{2m(E-V)\dot{\boldsymbol{r}}^2} \\ = &- E(t_2 - t_1) + \int_{t_1}^{t_2} \mathrm{d}t\sqrt{2m(E-V)\dot{\boldsymbol{r}}^2}. \end{aligned} \tag{208}$$

2.8 Dynamics in phase space

2.8.1 Phase portaits

For example if the system is described by

$$\frac{\mathrm{d}^3 x}{\mathrm{d}t^3} = G\left(x, \frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}^2 x}{\mathrm{d}t^2}, t\right) = G(x, y, z, t), \tag{209}$$

which introduce variables

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ z \\ G(x, y, z, t). \end{pmatrix} \tag{210}$$

We can recover form

$$\dot{\boldsymbol{x}} = \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{F}(\boldsymbol{x}), \tag{211}$$

by grouping $\boldsymbol{x} = (x, y, z)$ and $\boldsymbol{F} = (y, z, G)$.

2.8.2 First order systems

Example

The population growth follow the *logistic* equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = kx - \sigma x^2, \tag{212}$$

which is equivalent to

$$\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{1}{x(k - \sigma x)} = \frac{1}{k}\left(\frac{1}{x} + \frac{\sigma}{k - \sigma x}\right). \tag{213}$$

We can get t by integrating x from x_0 to x

$$t = \frac{1}{k}\left(\ln x - \ln(k - \sigma x)\right)\Big|_{x_0}^{x} = \frac{1}{k}\ln\left(\frac{x}{k - \sigma x}\right)\Bigg|_{x_0}^{x} = \frac{1}{k}\ln\left(\frac{x}{x_0}\frac{k - \sigma x_0}{k - \sigma x}\right). \tag{214}$$

Therefore, the solution is

$$x(t) = \frac{k x_0}{\sigma x_0 + (k - \sigma x_0)\exp(-kt)}, \tag{215}$$

for initial condition $x(0) = x_0 > 0$. For the *critical point* $\dot{x}_c = 0$, we have two solutions $x_c = 0$ and $x_c = k/\sigma$. Let $x(t) = x_c + \epsilon\delta x(t)$

$$\begin{aligned} \dot{x} =& k(x_c + \epsilon\delta x(t)) - \sigma(x_c + \epsilon\delta x(t))^2 \\ =& \left(kx_c - \sigma x_c^2\right) + \epsilon(k\delta x - 2\sigma x_c\delta x) - \mathcal{O}(\epsilon^2) \\ =& \epsilon(k - 2\sigma x_c)\delta x. \end{aligned} \tag{216}$$

So we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(\delta x) = (k - 2\sigma x_c)\delta x. \tag{217}$$

Now we look at linear stability of each critical point:

(1) $x_c = 0 \quad \Rightarrow \quad \delta\dot{x} = k\delta x \quad \Rightarrow \quad \delta x(t) = \mathrm{e}^{kt}\delta x(0)$ (repeller)

(2) $x_c = k/\sigma \quad \Rightarrow \quad \delta\dot{x} = -k\delta x \quad \Rightarrow \quad \delta x(t) = \mathrm{e}^{-kt}\delta x(0)$ (attractor)

We summarise the procedure for solving the first order system $\dot{x} = F(x)$:

(1) Find critical point $F(x_c) = 0$.

(2) Perturb around $x = x_c + \epsilon\delta x$, and

$$\delta\dot{x} = F'(x_c)\delta x. \tag{218}$$

(3)
- $\Re[F'(x_c)] > 0 \to$ unstable (repeller)
- $\Re[F'(x_c)] < 0 \to$ stable (attractor)
- $\Re[F'(x_c)] = 0$ and $\Im[F'(x_c)] \neq 0 \to$ oscillatory (libration)

2.8.3 Second order systems (2-d phase space)

For the second order system, a critical point $\boldsymbol{x}_c = \{x_c, y_c\}$ is a point at which both F and G vanish $F(x_c, y_c) = G(x_c, y_c) = 0$.

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} F(x, y) \\ G(x, y). \end{pmatrix} \tag{219}$$

Everywhere except at the critical point, the slope of a trajectory in the phase plane is given by

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{G(x, y)}{F(x, y)}. \tag{220}$$

Example
Consider the second order differential equation

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -\frac{1}{m}\frac{\mathrm{d}V}{\mathrm{d}x}, \quad y = \dot{x}, \tag{221}$$

which can expressed as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -\frac{1}{m}\frac{\mathrm{d}V}{\mathrm{d}x}. \end{pmatrix} \tag{222}$$

At the critical point

$$y_c = 0, \qquad \left.\frac{\mathrm{d}V}{\mathrm{d}x}\right|_{(x_c,y_c)} = 0. \tag{223}$$

Consider small displacements from critical points (x_c, y_c)

$$x = x_c + \delta x, \qquad y = 0 + \delta y, \tag{224}$$

we have

$$\delta \dot{x} = \delta y \tag{225}$$

$$\begin{aligned} \delta \dot{y} =& -\frac{1}{m}\frac{\mathrm{d}V(x_c+\delta x)}{\mathrm{d}x} \\ =& -\frac{1}{m}\left[\frac{\mathrm{d}V(x_c)}{\mathrm{d}x} + \frac{\mathrm{d}^2V(x_c)}{\mathrm{d}x^2}\delta x + \mathcal{O}((\delta x)^2)\right] = -\frac{1}{m}\frac{\mathrm{d}^2V(x_c)}{\mathrm{d}x^2}\delta x. \end{aligned} \tag{226}$$

Therefore,

(1) $\delta\ddot{x} = \delta\dot{y} = -\frac{1}{m}\frac{\mathrm{d}^2V(x_c)}{\mathrm{d}x^2}\delta x = -\omega^2\delta x \quad \Rightarrow \quad \delta x = A\mathrm{e}^{\pm i\omega t}$ (libration)

(2) $\delta\ddot{x} = \delta\dot{y} = -\frac{1}{m}\frac{\mathrm{d}^2V(x_c)}{\mathrm{d}x^2}\delta x = \Omega^2\delta x \quad \Rightarrow \quad \delta x = A\mathrm{e}^{\pm \Omega t}$ (unstable)

We summarize the procedure for 2-d phase space

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} F(x,y) \\ G(x,y) \end{pmatrix}. \tag{227}$$

(1) Find critical points $F(x_c, y_c) = G(x_c, y_c) = 0$.

(2) Perturb around critical points $x = x_c + \delta x$ and $y = y_c + \delta y$

$$\delta\dot{x} = \delta x \left.\frac{\partial F}{\partial x}\right|_{\boldsymbol{x}_c} + \delta y \left.\frac{\partial F}{\partial y}\right|_{\boldsymbol{x}_c} + \mathcal{O}\left((\delta \boldsymbol{x})^2\right), \tag{228}$$

$$\delta\dot{y} = \delta x \left.\frac{\partial G}{\partial x}\right|_{\boldsymbol{x}_c} + \delta y \left.\frac{\partial G}{\partial y}\right|_{\boldsymbol{x}_c} + \mathcal{O}\left((\delta \boldsymbol{x})^2\right), \tag{229}$$

which can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \left.\begin{pmatrix} \partial_x F & \partial_y F \\ \partial_x G & \partial_y G \end{pmatrix}\right|_{\boldsymbol{x}_c} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \mathbf{M}\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \lambda \begin{pmatrix} \delta x \\ \delta y \end{pmatrix}, \tag{230}$$

where λ is the eigenvalue of $\mathbf{M}$. So we have

$$\begin{pmatrix} \delta x(t) \\ \delta y(t) \end{pmatrix} = \mathrm{e}^{\lambda t}\begin{pmatrix} \delta x(0) \\ \delta y(0) \end{pmatrix}. \tag{231}$$

(3)
- $\Re[\lambda] > 0 \rightarrow$ unstable
- $\Re[\lambda] \leq 0 \rightarrow$ stable
- $\Re[\lambda] = 0$ and $\Im[\lambda] \neq 0 \rightarrow$ libration

3 Rigid Bodies

3.1 Many-body systems

Let us consider a system consisting N particles, which we may for instance label by an integer $a = 1, \dots, N$. We define the **total mass** of the whole system

$$M = \sum_a m_a, \tag{232}$$

and the position of **center of mass**

$$\boldsymbol{R} = \frac{\sum_a m_a \boldsymbol{r}_a}{M}. \tag{233}$$

The **total momentum** of the system

$$\boldsymbol{P} = \sum_a \boldsymbol{p}_a = \sum_a m_a \dot{\boldsymbol{r}}_a = M\dot{\boldsymbol{R}}. \tag{234}$$

The **total angular momentum** of the system

$$\boldsymbol{L} = \sum_a m_a \boldsymbol{r}_a \times \dot{\boldsymbol{r}}_a. \tag{235}$$

The position $\boldsymbol{r}_a$ can be expressed as $\boldsymbol{r}_a = \boldsymbol{R} + \boldsymbol{r}_a^*$ and $\sum_a m_a \boldsymbol{r}_a^* = 0$. The total angular momentum

$$\begin{aligned} \boldsymbol{L} &= \sum_a m_a(\boldsymbol{R} + \boldsymbol{r}_a^*) \times (\dot{\boldsymbol{R}} + \dot{\boldsymbol{r}}_a^*) \\ &= \sum_a m_a \boldsymbol{R} \times \dot{\boldsymbol{R}} + \sum_a m_a \boldsymbol{R} \times \dot{\boldsymbol{r}}_a^* + \sum_a m_a \boldsymbol{r}_a^* \times \dot{\boldsymbol{R}} + \sum_a m_a \boldsymbol{r}_a^* \times \dot{\boldsymbol{r}}_a^* \\ &= \sum_a m_a \boldsymbol{R} \times \dot{\boldsymbol{R}} + \sum_a m_a \boldsymbol{r}_a^* \times \dot{\boldsymbol{r}}_a^* = \boldsymbol{L}_{\text{com}} + \boldsymbol{L}^*, \end{aligned} \tag{236}$$

where

$$\boldsymbol{L}^* = \sum_a m_a \boldsymbol{r}_a^* \times \dot{\boldsymbol{r}}_a^*, \tag{237}$$

is called the angular momentum about the center of mass.

3.2 Rotation about a fixed axis

Suppose the axis of rotation along the z axis, then the z component of the angular momentum is

$$L_z = \sum_a m_a r_a^2 \omega = I\omega, \tag{238}$$

where $I = \sum_a m_a r_a^2$ is the **moment of inertia** about the axis. The kinetic energy in terms of I and ω is

$$T = \sum_a \frac{1}{2} m_a \left(r_a \dot{\theta}_a \right)^2 = \frac{1}{2} I\omega^2. \tag{239}$$

3.3 Action for a rotating rigid body

The kinetic energy of the center of mass

$$\begin{aligned} T &= \sum_a \frac{1}{2} m_a \dot{\boldsymbol{r}}_a^2 = \sum_a \frac{1}{2} m_a (\dot{\boldsymbol{R}} + \dot{\boldsymbol{r}}^*)^2 \\ &= \sum_a \frac{1}{2} m_a \dot{\boldsymbol{R}}^2 + \sum_a m_a \dot{\boldsymbol{R}} \cdot \dot{\boldsymbol{r}}_a^* + \sum_a \frac{1}{2} m_a \dot{\boldsymbol{r}}_a^{*2} \\ &= \underbrace{\frac{1}{2} M \dot{\boldsymbol{R}}^2}_{\text{com kinetic energy}} + \underbrace{\sum_a \frac{1}{2} m_a \dot{\boldsymbol{r}}_a^{*2}}_{\text{rotational kinetic energy}} . \end{aligned} \tag{240}$$

Since the body is assumed rigid, all it can do is rotate relative to its center of mass motion. Let $\boldsymbol{\omega}$ denote the angular velocity of the rotation, then we have

$$\dot{\boldsymbol{r}}_a^* = \boldsymbol{\omega} \times \boldsymbol{r}_a^*, \tag{241}$$

this preserves all the scalar products

$$\frac{\mathrm{d}}{\mathrm{d}t}(|\boldsymbol{r}_a^* - \boldsymbol{r}_b^*|^2) = 2(\boldsymbol{r}_a^* - \boldsymbol{r}_b^*) \cdot (\dot{\boldsymbol{r}}_a^* - \dot{\boldsymbol{r}}_b^*) = 2(\boldsymbol{r}_a^* - \boldsymbol{r}_b^*) \cdot (\boldsymbol{\omega} \times (\boldsymbol{r}_a^* - \boldsymbol{r}_b^*)) = 0. \tag{242}$$

Putting this together, the action is

$$S = \int \mathrm{d}t \left[\frac{1}{2} M \dot{\boldsymbol{R}}^2 + \frac{1}{2} \sum_a m_a (\boldsymbol{\omega} \times \boldsymbol{r}_a^*)^2 \right]. \tag{243}$$

3.3.1 Rotation around a pivot

If we assume that the pivot point is taken at $\boldsymbol{r} = 0$, then we have the constraints

$$\dot{\boldsymbol{R}} = \boldsymbol{\omega} \times \boldsymbol{R}. \tag{244}$$

When this is true we have

$$\dot{\boldsymbol{r}}_a = \boldsymbol{\omega} \times \boldsymbol{r}_a, \tag{245}$$

and to the action can be evaluated as

$$S = \int \mathrm{d}t \sum_a \frac{1}{2} m_a (\boldsymbol{\omega} \times \boldsymbol{r}_a)^2. \tag{246}$$

We stress this is only the correct form if the pivot is taken at the origin.

3.3.2 Switch vector to index notation

We can use some mathematical tricks (Levi-Civita symbol, Levi-Civita identity, Einstein summation convention) to do with our work

$$(\boldsymbol{A} \times \boldsymbol{B})_i = \varepsilon_{ijk} A_j B_k, \tag{247}$$

$$\varepsilon_{ijk}\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}. \tag{248}$$

Then, we can calculate:

$$\begin{aligned}(\boldsymbol{\omega} \times \boldsymbol{r}^*)^2 =& \left(\varepsilon_{ijk}\omega_j r_k^*\right)\left(\varepsilon_{ilm}\omega_l r_m^*\right) = \left(\varepsilon_{ijk}\varepsilon_{ilm}\right)\omega_j r_k^* \omega_l r_m^* \\ =& \omega_j\omega_l\left(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}\right) r_k^* r_m^* = \omega_j\omega_l\left(\delta_{jl} r_k^{*2} - r_l^* r_j^*\right) \\ =& \omega_i\omega_j\left(\delta_{ij}\boldsymbol{r}^{*2} - r_i^* r_j^*\right).\end{aligned} \tag{249}$$

Now we define **the momenta of inertia**

$$I_{ij} = \sum_a m_a\left(\delta_{ij}\boldsymbol{r}_a^2 - r_{ai}r_{aj}\right), \tag{250}$$

where $\boldsymbol{r}_a = \boldsymbol{R} + \boldsymbol{r}_a^*$. The equivalent quantity defined relative to the center of mass

$$I_{ij}^* = \sum_a m_a\left(\delta_{ij}\boldsymbol{r}_a^{*2} - r_{ai}^* r_{aj}^*\right). \tag{251}$$

With this notation, we can rewrite the action S

$$S = \int \mathrm{d}t\left(\frac{1}{2}M\dot{\boldsymbol{R}}^2 + \frac{1}{2}\omega_i I_{ij}\omega_j\right) = \int \mathrm{d}t\left(\frac{1}{2}M\dot{\boldsymbol{R}}^2 + \frac{1}{2}\boldsymbol{\omega}^T\boldsymbol{I}\boldsymbol{\omega}\right). \tag{252}$$

3.4 Spin angular momentum

The spin angular momentum

$$\boldsymbol{L} = \sum_a m_a\boldsymbol{r}_a \times \dot{\boldsymbol{r}}_a = M\boldsymbol{R} \times \dot{\boldsymbol{R}} + \boldsymbol{L}^*, \tag{253}$$

then we have index notation

$$L_i = \sum_a m_a\epsilon_{ijk}r_{aj}\dot{r}_{ak} = \sum_a m_a\epsilon_{ijk}r_{aj}\epsilon_{klm}\omega_l r_{am} = I_{ij}\omega_j. \tag{254}$$

The angular momentum about the center of mass

$$\boldsymbol{L}^* = \sum_a m_a\boldsymbol{r}_a^* \times \dot{\boldsymbol{r}}_a^*. \tag{255}$$

Similarly, we have

$$L_i^* = I_{ij}^*\omega_j, \tag{256}$$

and we also find that

$$p_{\theta_i} = \frac{\partial L}{\partial\dot{\theta}_i} = \frac{\partial L}{\partial\omega_i} = L_i^*. \tag{257}$$

3.4.1 Component/vector notation

The angular momentum in vector notation

$$\boldsymbol{L} = \sum_a m_a \begin{pmatrix} y_a^2 + z_a^2 & -x_a y_a & -x_a z_a \\ -x_a y_a & x_a^2 + y_a^2 & -y_a z_a \\ -x_a z_a & -y_a z_a & x_a^2 + y_a^2 \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \boldsymbol{I} \cdot \boldsymbol{\omega}, \tag{258}$$

where

$$\boldsymbol{I} = \sum_a m_a \begin{pmatrix} y_a^2 + z_a^2 & -x_a y_a & -x_a z_a \\ -x_a y_a & x_a^2 + y_a^2 & -y_a z_a \\ -x_a z_a & -y_a z_a & x_a^2 + y_a^2 \end{pmatrix}. \tag{259}$$

3.5 Euler angles

3.5.1 Rotations in 3 dimensions

A rotation in the $x - y$ plane is given by

$$\boldsymbol{M}_z(\theta_z) = \begin{pmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{260}$$

A rotation in the $y - z$ plane is given by

$$\boldsymbol{M}_x(\theta_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{pmatrix}. \tag{261}$$

A rotation in the $x - z$ plane is given by

$$\boldsymbol{M}_y(\theta_y) = \begin{pmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{pmatrix}. \tag{262}$$

For instance under $\boldsymbol{M}_y$ the vector transforms as

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathrm{M}_y(\theta_y) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x\cos\theta_y + z\sin\theta_y \\ y \\ -x\sin\theta_y + z\cos\theta_y \end{pmatrix}. \tag{263}$$

Each matrix is orthogonal meaning $\boldsymbol{M}^T = \boldsymbol{M}^{-1}$.

3.5.2 Parameterzing general rotations

Euler noticed that a general rotation can be written as a rotation with axis z then y then z again

$$\boxed{\boldsymbol{M}(\psi, \theta, \varphi) = \boldsymbol{M}_z(\varphi)\boldsymbol{M}_y(\theta)\boldsymbol{M}_z(\psi).} \tag{264}$$

These angles are known as **Euler angles.** Thus

$$\boldsymbol{r}^* = \boldsymbol{M}(\psi,\theta,\varphi)\boldsymbol{u} = \boldsymbol{M}_z(\varphi)\boldsymbol{M}_y(\theta)\boldsymbol{M}_z(\psi)\boldsymbol{u}, \tag{265}$$

where $\boldsymbol{u}$ is assumed independent of time as this is a rigid body then we have

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{r}^*}{\mathrm{d}t} = \frac{\boldsymbol{M}_z(\varphi)}{\mathrm{d}t}\boldsymbol{M}_y(\theta)\boldsymbol{M}_z(\psi)\boldsymbol{u} + \boldsymbol{M}_z(\varphi)\frac{\mathrm{d}\boldsymbol{M}_y(\theta)}{\mathrm{d}t}\boldsymbol{M}_z(\psi)\boldsymbol{u} \\ + \boldsymbol{M}_z(\varphi)\boldsymbol{M}_y(\theta)\frac{\mathrm{d}\boldsymbol{M}_z(\psi)}{\mathrm{d}t}\boldsymbol{u}.\end{aligned} \tag{266}$$

Because of the relation

$$\boldsymbol{u} = \boldsymbol{M}^{-1}\boldsymbol{r}^* = \boldsymbol{M}^T\boldsymbol{r}^* = \boldsymbol{M}_z^T(\psi)\boldsymbol{M}_y^T(\theta)\boldsymbol{M}_z^T(\varphi)\boldsymbol{r}^*, \tag{267}$$

we can write

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{r}^*}{\mathrm{d}t} =& \frac{\boldsymbol{M}_z(\varphi)}{\mathrm{d}t}\boldsymbol{M}_z^T(\varphi)\boldsymbol{r}^* + \boldsymbol{M}_z(\varphi)\frac{\mathrm{d}\boldsymbol{M}_y(\theta)}{\mathrm{d}t}\boldsymbol{M}_y^T(\theta)\boldsymbol{M}_z^T(\varphi)\boldsymbol{r}^* \\ &+ \boldsymbol{M}_z(\varphi)\boldsymbol{M}_y(\theta)\frac{\mathrm{d}\boldsymbol{M}_z(\psi)}{\mathrm{d}t}\boldsymbol{M}_z^T(\psi)\boldsymbol{M}_y^T(\theta)\boldsymbol{M}_z^T(\varphi)\boldsymbol{r}^* = \boldsymbol{\Omega}\boldsymbol{r}^*,\end{aligned} \tag{268}$$

where

$$\begin{aligned}\boldsymbol{\Omega} = \frac{\boldsymbol{M}_z(\varphi)}{\mathrm{d}t}\boldsymbol{M}_z^T(\varphi) + \boldsymbol{M}_z(\varphi)\left(\frac{\mathrm{d}\boldsymbol{M}_y(\theta)}{\mathrm{d}t}\boldsymbol{M}_y^T(\theta)\right)\boldsymbol{M}_z^T(\varphi) \\ + \boldsymbol{M}_z(\varphi)\boldsymbol{M}_y(\theta)\left(\frac{\mathrm{d}\boldsymbol{M}_z(\psi)}{\mathrm{d}t}\boldsymbol{M}_z^T(\psi)\right)\boldsymbol{M}_y^T(\theta)\boldsymbol{M}_z^T(\varphi).\end{aligned} \tag{269}$$

We notice that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{M}\boldsymbol{M}^T) = \frac{\mathrm{d}\boldsymbol{M}}{\mathrm{d}t}\boldsymbol{M}^T + \boldsymbol{M}\frac{\mathrm{d}\boldsymbol{M}^T}{\mathrm{d}t} = 0, \tag{270}$$

$$\left(\frac{\mathrm{d}\boldsymbol{M}}{\mathrm{d}t}\boldsymbol{M}^T\right)^T = \boldsymbol{M}\frac{\mathrm{d}\boldsymbol{M}^T}{\mathrm{d}t} = -\left(\frac{\mathrm{d}\boldsymbol{M}}{\mathrm{d}t}\boldsymbol{M}^T\right). \tag{271}$$

So $\boldsymbol{\Omega}$ is an anti-symmetric matrix $\boldsymbol{\Omega}^T = -\boldsymbol{\Omega}$. We can write the equations above using the Levi-Civita symbol in index notation as

$$\frac{\mathrm{d}r_i^*}{\mathrm{d}t} = \Omega_{ij}r_j^* = (\boldsymbol{\omega}\times\boldsymbol{r}^*)_i = \epsilon_{ijk}\omega_j r_k^* = -\epsilon_{ijk}r_j^*\omega_k. \tag{272}$$

So $\boldsymbol{\Omega}$ can be expressed as

$$\Omega_{ij} = -\sum_{k=1}^{3}\epsilon_{ijk}\omega_k. \tag{273}$$

Noe we calculate $\boldsymbol{\Omega}$, from which derivative the angular velocity vector $\boldsymbol{\omega}$:

$$\begin{aligned}\boldsymbol{\omega} =& \dot{\varphi}\hat{\boldsymbol{k}} + \dot{\theta}\boldsymbol{M}_z(\varphi)\hat{\boldsymbol{j}} + \dot{\psi}\boldsymbol{M}_z(\psi)\boldsymbol{M}_y(\theta)\hat{\boldsymbol{k}} \\ =& (\dot{\psi}\cos\theta\sin\varphi - \dot{\theta}\sin\varphi)\hat{\boldsymbol{i}} + (\dot{\psi}\sin\theta\cos\varphi + \dot{\theta}\cos\varphi)\hat{\boldsymbol{j}} + (\dot{\varphi} + \dot{\psi}\cos\theta)\hat{\boldsymbol{k}}.\end{aligned} \tag{274}$$

The new axis are given by

$$\boldsymbol{e}_1 = \boldsymbol{M}(\psi,\theta,\varphi)\hat{\boldsymbol{i}} = \begin{pmatrix} \cos\psi\cos\varphi\cos\theta - \sin\psi\sin\varphi \\ \cos\psi\sin\varphi\cos\theta + \sin\psi\cos\varphi \\ -\cos\psi\sin\theta \end{pmatrix}, \tag{275}$$

$$\boldsymbol{e}_2 = \boldsymbol{M}(\psi,\theta,\varphi)\hat{\boldsymbol{j}} = \begin{pmatrix} -\sin\psi\cos\varphi\cos\theta - \cos\psi\sin\varphi \\ \cos\psi\cos\varphi - \sin\psi\sin\varphi\cos\theta \\ \sin\psi\sin\theta \end{pmatrix}, \tag{276}$$

$$\boldsymbol{e}_3 = \boldsymbol{M}(\psi,\theta,\varphi)\hat{\boldsymbol{k}} = \begin{pmatrix} \cos\varphi\sin\theta \\ \sin\varphi\sin\theta \\ \cos\theta \end{pmatrix}. \tag{277}$$

It is useful to write the angular velocity in the new basis

$$\begin{aligned} \boldsymbol{\omega} =& \sum_\alpha \omega_\alpha \hat{\boldsymbol{e}}_\alpha \\ =& (-\cos\psi\sin\theta\dot{\varphi} + \sin\psi\dot{\theta})\hat{\boldsymbol{e}}_1 + (\cos\psi\dot{\theta} + \sin\psi\sin\theta\dot{\varphi})\hat{\boldsymbol{e}}_2 + (\dot{\psi} + \cos\theta\dot{\varphi})\hat{\boldsymbol{e}}_3. \end{aligned} \tag{278}$$

It is important to remember that the new basis is a rotating (and hence not inertial!!!) one with basis vectors rotating with the body itself as so we have

$$\frac{\mathrm{d}\hat{\boldsymbol{e}}_\alpha}{\mathrm{d}t} = \boldsymbol{\omega} \times \hat{\boldsymbol{e}}_\alpha. \tag{279}$$

3.6 Principal axes of inertia

3.6.1 Principal axies of intertia as reference frame

The angular momentum

$$\boldsymbol{L} = \sum_\alpha I_\alpha \omega_\alpha \hat{\boldsymbol{e}}_\alpha. \tag{280}$$

The rotational kinetic energy for a body with pivot at the origin can be expressed as

$$\begin{aligned} T = \frac{1}{2}\sum_\alpha \omega_\alpha I_\alpha \omega_\alpha =& \frac{1}{2} I_1(-\cos\psi\sin\theta\dot{\varphi} + \sin\psi\dot{\theta})^2 \\ &+ \frac{1}{2} I_2(\cos\psi\dot{\theta} + \sin\psi\sin\theta\dot{\varphi})^2 + \frac{1}{2} I_3(\dot{\psi} + \cos\theta\dot{\varphi})^2. \end{aligned} \tag{281}$$

Generally, the principal axes are not equal to each other, i.e., $I_1 \neq I_2 \neq I_3$. But sometimes there are symmetries in the system. For example $I_1 = I_2$, then

$$T = \frac{1}{2} I_1 \left(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2\right) + \frac{1}{2} I_3 \left(\dot{\psi} + \cos\theta\dot{\varphi}\right)^2. \tag{282}$$

3.7 Rotation about a principal axis

As the principal axes are fixed in the body, they form a rotating frame of reference. Now we want to distinguish between the *inertial* and *rotating* frames. The rate of change of the angular momentum in the inertial frame is

$$\left.\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t}\right|_{\mathrm{I}} = \sum_a \boldsymbol{r}_a \times \boldsymbol{F}_a = \boldsymbol{\tau}, \tag{283}$$

which related to the rotating frame

$$\left.\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t}\right|_{I} = \left.\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t}\right|_{\mathrm{R}} + \boldsymbol{\omega} \times \boldsymbol{L}. \tag{284}$$

We calculate

$$\begin{aligned}\left.\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t}\right|_{I} &= \sum_\alpha \dot{\omega}_\alpha I_\alpha \hat{\boldsymbol{e}}_\alpha + \sum_\beta \omega_\beta I_\beta \frac{\mathrm{d}\hat{\boldsymbol{e}}_\beta}{\mathrm{d}t} = \sum_\alpha \dot{\omega}_\alpha I_\alpha \hat{\boldsymbol{e}}_\alpha + \sum_\beta \omega_\beta I_\beta \boldsymbol{\omega} \times \hat{\boldsymbol{e}}_\beta \\ &= \sum_\alpha \dot{\omega}_\alpha I_\alpha \hat{\boldsymbol{e}}_\alpha + \sum_{\beta\gamma} \omega_\beta I_\beta \omega_\gamma \hat{\boldsymbol{e}}_\gamma \times \hat{\boldsymbol{e}}_\beta = \sum_\alpha \dot{\omega}_\alpha I_\alpha \hat{\boldsymbol{e}}_\alpha + \sum_{\alpha\beta\gamma} \omega_\beta I_\beta \omega_\gamma \varepsilon_{\beta\alpha\gamma} \hat{\boldsymbol{e}}_\alpha.\end{aligned} \tag{285}$$

So we have

$$I_\alpha \dot{\omega}_\alpha + \sum_{\beta\gamma} \omega_\beta \omega_\gamma I_\beta \varepsilon_{\beta\alpha\gamma} = \tau_\alpha, \tag{286}$$

or written explicitly for each component this gives

$$I_1 \dot{\omega}_1 + (I_3 - I_2)\omega_2\omega_3 = \tau_1, \tag{287}$$
$$I_2 \dot{\omega}_2 + (I_1 - I_3)\omega_3\omega_1 = \tau_2, \tag{288}$$
$$I_3 \dot{\omega}_3 + (I_2 - I_1)\omega_1\omega_2 = \tau_3. \tag{289}$$

Example

Now we think about a specific system with $\omega_3 = \omega_0$ and $\omega_1 = \omega_2 = 0$. We can write the solution in a combination of its background solution and a small variation:

$$\omega_1 = \delta\omega_1, \quad \omega_2 = \delta\omega_2, \quad \omega_3 = \omega_0 + \delta\omega_3. \tag{290}$$

Then

$$I_1 \delta\dot{\omega}_1 + (I_3 - I_2)\delta\omega_2\omega_0 = 0, \tag{291}$$
$$I_2 \delta\dot{\omega}_2 + (I_1 - I_3)\omega_0\delta\omega_1 = 0, \tag{292}$$
$$I_3 \delta\dot{\omega}_3 = 0. \tag{293}$$

Solving this and we get

$$\frac{\mathrm{d}^2\delta\omega_{1,2}}{\mathrm{d}t^2} = -\frac{(I_3 - I_2)(I_3 - I_1)}{I_1 I_2}\delta\omega_{1,2}. \tag{294}$$

The system is stable if $(I_3 - I_2)(I_3 - I_1) > 0$.

4 Classical Field Theory

4.1 Continuous Systems

In addition to mechanical systems consisting of a finite number of degrees of freedom, one is often also interested in continuous systems:

$$q_i(t) \;\rightarrow\; \phi(\boldsymbol{x}, t).$$

4.1.1 One-dimensional example: string

The simplest example of the classical field theory is a string:

$$y = \phi(x, t). \tag{295}$$

We assume that in equilibrium the string is stretched to length l_0 and has tension k. The mass per unit length is μ. The kinetic energy and the potential energy are

$$T = \int_0^{l_0} \mathrm{d}x \frac{1}{2}\mu \left(\frac{\partial \phi}{\partial t}\right)^2, \quad V = k(l - l_0). \tag{296}$$

The length of the infinitesimal segment of length is

$$\mathrm{d}l = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}\phi)^2} = \mathrm{d}x\sqrt{1 + \left(\frac{\partial \phi}{\partial x}\right)^2} = \mathrm{d}x\left[1 + \frac{1}{2}\left(\frac{\partial \phi}{\partial x}\right)^2 + \dots\right], \tag{297}$$

and therefore the potential energy is

$$V = \int_0^{l_0} \mathrm{d}x \frac{1}{2} k \left(\frac{\partial \phi}{\partial x}\right)^2. \tag{298}$$

We now write the whole Lagrangian

$$L = T - V = \int_0^{l_0} \mathrm{d}x \left[\frac{1}{2}\mu \left(\frac{\partial \phi}{\partial t}\right)^2 - \frac{1}{2}k\left(\frac{\partial \phi}{\partial x}\right)^2\right]. \tag{299}$$

The integrand is called the Lagrangian density and denoted by $\mathcal{L}$

$$\mathcal{L}\left(\phi, \frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial x}\right) = \frac{1}{2}\mu \left(\frac{\partial \phi}{\partial t}\right)^2 - \frac{1}{2}k\left(\frac{\partial \phi}{\partial x}\right)^2. \tag{300}$$

The action is

$$S = \int \mathrm{d}t \int_0^{l_0} \mathrm{d}x \mathcal{L}\left(\phi, \frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial x}\right) \tag{301}$$

Variation of the action is

$$\begin{aligned} \delta S &= \int \mathrm{d}t \int_0^{l_0} \mathrm{d}x \left(\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial t)}\frac{\partial(\delta\phi)}{\partial t} + \frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial x)}\frac{\partial(\delta\phi)}{\partial x}\right) \\ &= \int \mathrm{d}t \int_0^{l_0} \mathrm{d}x \left[\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial t)}\right) - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial x)}\right)\right]\delta\phi, \end{aligned} \tag{302}$$

up to the boundary conditions. This leads to the Euler-Lagrange equation

$$\boxed{\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial t)}\right) - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial x)}\right) = 0.} \tag{303}$$

So the equation of motion for the string is

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = \frac{\partial^2\phi}{\partial x^2}, \tag{304}$$

where $c = \sqrt{k/\mu}$.

4.2 Klein-Gordon scalar field

Now we generalize the Lagrangian formalism to quantities ϕ that are functions of time and all three dimensions of space, $\phi = \phi(t, x, y, z)$. We can use the notation

$$\partial_\mu = \frac{\phi}{\partial x^\mu} = \left(\frac{\phi}{\partial x^0}, \frac{\phi}{\partial x^1}, \frac{\phi}{\partial x^2}, \frac{\phi}{\partial x^3}\right) = (\partial_t, \partial_x, \partial_y, \partial_z). \tag{305}$$

The general Lagrangian density $\mathcal{L}$ will now be a function of

$$\mathcal{L} = \mathcal{L}(\phi, \partial_\mu, x^\mu). \tag{306}$$

The resulting Euler-Lagrange equation is then

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right) = 0. \tag{307}$$

For the example of the Klein-Gordon scalar field defined as $(c = 1)$

$$\mathcal{L} = \frac{1}{2}(\partial_t\phi)^2 - \frac{1}{2}(\boldsymbol{\nabla\phi})^2 - V(\phi) = -\frac{1}{2}\eta^{\nu\omega}\partial_\nu\phi\partial_\omega\phi - V(\phi), \tag{308}$$

where $\eta^{00} = -1, \eta^{0i} = 0, \eta^{ij} = \delta_{ij}, (i, j = 1, 2, 3)$, and

$$\frac{\partial(\partial_\nu\phi)}{\partial(\partial_\mu\phi)} = \delta^\nu_\mu, \tag{309}$$

where the spacetime Kronecker delta is defined so that $\delta^0_0 = 1, \delta^i_0\delta^0_i = 0, \delta^j_i = \delta_{ij}$. Then

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} = -\eta^{\mu\nu}\partial_\nu\phi. \tag{310}$$

Thus the Euler-Lagrange equations are

$$-\frac{\partial V}{\partial \phi} - \partial_\mu(-\eta^{\mu\nu}\partial_\nu\phi) = 0. \tag{311}$$

The resulting equation is known as the **Klein-Gordon equation**

$$\Box\phi = \frac{\partial V}{\partial \phi}, \tag{312}$$

which is the equation for a wave in four spacetime dimensions and we have defined the d'Alembertian operator $\Box$ as

$$\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu = -\partial_t^2 + \nabla^2 = -\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{313}$$

4.3 Hamiltonian density for a continuous system

For a continuous system described by a Lagrangian density $\mathcal{L}(\phi, \partial_\mu \phi)$ for a field ϕ , we define the conjugate momentum associated to ϕ as

$$\pi = \frac{\partial \mathcal{L}}{\partial(\partial_t \phi)}. \tag{314}$$

The Hamiltonian is given by $H = \int \mathrm{d}^3 x \mathcal{H}$, where the Hamiltonian density $\mathcal{H}$ is then defined via Legendre transform as

$$\mathcal{H} = \frac{\partial \phi}{\partial t}\pi - \mathcal{L}(\phi, \partial_\mu \phi). \tag{315}$$

The phase space form of the action is then

$$S = \int \mathrm{d}t \int \mathrm{d}^3 x \left(\frac{\partial \phi}{\partial t}\pi - \mathcal{H}(\phi, \pi, \partial_i \phi) \right). \tag{316}$$

To obtain Hamilton equation we perform the usual variations treating ϕ and π as independent

$$\begin{aligned}
\delta S &= \int \mathrm{d}t \int \mathrm{d}^3 x \left[(\partial_t \delta\phi)\pi + (\partial_t \phi)\delta\pi - \frac{\partial \mathcal{H}}{\partial \phi}\delta\phi - \frac{\partial \mathcal{H}}{\partial(\partial_i \phi)}\partial_i \delta\phi - \frac{\partial \mathcal{H}}{\partial \pi}\delta\pi \right] \\
&= \int \mathrm{d}t \int \mathrm{d}^3 x \left[-(\partial_t \pi)\delta\phi + (\delta_t \phi)\delta\pi - \frac{\partial \mathcal{H}}{\partial \phi}\delta\phi + \partial_i \left(\frac{\partial \mathcal{H}}{\partial(\partial_i \phi)} \right)\delta\phi - \frac{\partial \mathcal{H}}{\partial \pi}\delta\pi \right] \\
&= \int \mathrm{d}t \int \mathrm{d}^3 x \left[-\left(\partial_t \pi + \frac{\partial \mathcal{H}}{\partial \phi} - \partial_i \left(\frac{\partial \mathcal{H}}{\partial(\partial_i \phi)} \right) \right)\delta\phi + \left(\partial_t \phi - \frac{\partial \mathcal{H}}{\partial \pi} \right)\delta\pi \right] \\
&= \int \mathrm{d}t \int \mathrm{d}^3 x \left[-\left(\partial_t \pi + \frac{\delta H}{\delta \phi} \right)\delta\phi + \left(\partial_t \phi - \frac{\delta H}{\delta \pi} \right)\delta\pi \right],
\end{aligned} \tag{317}$$

where

$$\frac{\delta H}{\delta X} = \frac{\partial \mathcal{H}}{\partial X} - \partial_t \left(\frac{\partial \mathcal{H}}{\partial(\partial_i X)} \right). \tag{318}$$

The Hamilton equations for a continuous system are then simply given by

$$\partial_t \phi(x^\mu) = \frac{\delta H}{\delta \pi(x^\mu)}, \quad \partial_t \pi(x^\mu) = -\frac{\delta H}{\delta \phi(x^\mu)}. \tag{319}$$

For the example of the Klein-Gordon scalar field

$$\mathcal{H} = \frac{1}{2}\pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V(\phi), \tag{320}$$

the Hamilton equations are

$$\dot{\phi} = \frac{\delta H}{\delta \pi} = \pi, \quad \dot{\pi} = -\frac{\delta H}{\delta \phi} = \boldsymbol{\nabla}^2 \phi - V'(\phi), \tag{321}$$

which is equivalent to the Klein-Gordon equation derived in the Lagrangian formalism

$$\ddot{\phi} = \dot{\pi} = \boldsymbol{\nabla}^2 \phi - V'(\phi) \quad \Rightarrow \quad \Box \phi = V'(\phi). \tag{322}$$

4.4 Noether's theorem in field theory

Consider a set of fields $\phi_I(x)$, and suppose that the action is invariant under the global transformation

$$\delta\phi_I = F_I\delta\lambda, \tag{323}$$

where $F_I = F_I(\phi, \partial_t\phi)$. The Lagrangian density transform as a total derivative

$$\delta\mathcal{L} = \partial_\mu A^\mu \delta\lambda, \tag{324}$$

and so explicitly (Einstein summation over I)

$$\frac{\partial\mathcal{L}}{\partial\phi_I}F_I + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_I)}\partial_\mu F_I = \partial_\mu A^\mu. \tag{325}$$

Now we let $\delta\lambda$ as a function of all 4 spacetime coordinates x, then the full variation of the action

$$\begin{aligned}\delta S &= \int \mathrm{d}^4x \left(\partial_\mu A^\mu\delta\lambda + F_I\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_I)}\partial_\mu\delta\lambda\right)\\ &= \int \mathrm{d}^4x\partial_\mu\left(A^\mu - F_I\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_I)}\right)\delta\lambda\\ &= \int \mathrm{d}^4x(-\partial_\mu J^\mu)\delta\lambda,\end{aligned} \tag{326}$$

where we have defined the current

$$J^\mu = F_I\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_I)} - A^\mu, \tag{327}$$

and have

$$\partial_\mu J^\mu = 0. \tag{328}$$

The Noether charge we previous obtained is just the naive charge associated with this current

$$G = \int \mathrm{d}^3\boldsymbol{x} J^0(x), \tag{329}$$

which is conserved since

$$\frac{\mathrm{d}G}{\mathrm{d}t} = \int \mathrm{d}^3\boldsymbol{x}\partial_0 J^0(x) = -\int \mathrm{d}^3\boldsymbol{x}\partial_i J^i(x) = 0. \tag{330}$$

4.5 Stress-energy-momentum tensors

The relevant global symmetry is translation invariance, of both space and time coordinates $x^\mu \to x^\mu + a^\mu$. Under this symmetry a scalar field transforms as

$$\delta\phi = \partial_\nu\phi\delta a^\nu = F_I\delta a^\nu, \quad F_I = \partial_\nu\phi. \tag{331}$$

The Lagrangian density transforms the same way

$$\delta\mathcal{L} = \partial_\nu\mathcal{L}\delta a^\nu = \partial_\mu A^\mu{}_\nu\delta a^\nu, \quad A^\mu{}_\nu = \delta^\mu{}_\nu\mathcal{L}. \tag{332}$$

The conserved current is

$$T^{\mu}{}_{\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial_\nu \phi - \delta^{\mu}{}_{\nu}\mathcal{L}, \tag{333}$$

which is known as the **stress-energy momentum tensor**. Despite its two indices, it is conserved like a current

$$\partial_\mu T^{\mu}{}_{\nu} = 0. \tag{334}$$

In this case we have four Noether charges, which defined by

$$-P_\nu = \int \mathrm{d}^3\boldsymbol{x} \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)}\partial_\nu \phi - \delta^{0}{}_{\nu}\mathcal{L} \right). \tag{335}$$

This is the total energy and momentum of the system. In particular

$$-P_0 = H = \int \mathrm{d}^3\boldsymbol{x} \left(\frac{\partial \mathcal{L}}{\partial(\partial_t \phi)}\partial_t \phi - \mathcal{L} \right), \tag{336}$$

is nothing other than the Hamiltonian.

5 Special Relativity

5.1 Galilean relativity

From a modern perspective, the transformation between different inertial frames is given by a "Galilean boost", which takes the form

$$t' = t + a^t, \tag{337}$$

$$\boldsymbol{x}' = R\boldsymbol{x} - \boldsymbol{v}t + \boldsymbol{a}, \tag{338}$$

where R is a rotation matrix and is an element of the group $O(3)$, i.e. $R^T R = 1$, a^t is a constant time translation and a is a constant space translation.

5.2 Minkowski spacetime

We can now define the geometry of flat or Minkowski space-time by defining a notion of spacetime interval which measures the space-time distance between two points in space-time at x^μ and $x^\mu + \delta x^\mu$

$$\Delta s^2 = -c^2\Delta t^2 + \Delta \boldsymbol{r}^2 = -c^2(\Delta x^0)^2 + (\Delta x^1)^2 + (\Delta x^2)^2 + (\Delta x^3)^2. \tag{339}$$

And we categorize the spacetime distance into three case:

$$\begin{aligned} \Delta s^2 &> 0, \quad \text{spacelike seperated}, \\ \Delta s^2 &= 0, \quad \text{null seperated}, \\ \Delta s^2 &< 0, \quad \text{timelike seperated}. \end{aligned}$$

To express this in the four-vector notation, we can introduce the four-dimensional **Minkowski metric** η defined as

$$\eta_{\mu\nu} = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{340}$$

which represents the metric in flat space-time. This metric is symmetric $\eta_{\nu\mu} = \eta_{\mu\nu}$. Using the Minkowski metric, the flat space-time interval can be written as

$$\Delta s^2 = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu. \tag{341}$$

We emphasize that when written in this form the multiplications are entirely commutative, i.e., $\Delta s^2 = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu = \Delta x^\mu \eta_{\mu\nu} \Delta x^\nu = \Delta x^\nu \eta_{\mu\nu} \Delta x^\mu = \Delta x^\mu \Delta x^\nu \eta_{\mu\nu}$.

5.3 Proper time

If a particle moving with a velocity v, the proper time $\Delta\tau$ is given by

$$\Delta\tau^2 = -\frac{1}{c^2}\eta_{ab}\Delta x^a \Delta x^b = \Delta t^2 - \frac{1}{c^2}\Delta x^2 = \left(1 - \frac{v^2}{c^2}\right)\Delta t^2. \tag{342}$$

The proper time of the particle is hence related to the physical time by

$$\Delta t = \gamma \Delta \tau, \tag{343}$$

where we have defined the usual **Lorentz factor**

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{344}$$

5.4 Lorentz transformation

In most of what follows, we work in 'natural' units where the speed of light in the vacuum, $c = 1$. Consider now a Lorentz boost in the direction $x^1 = x$ by velocity v, the time and space coordinates transform as

$$t' = \gamma\left(t - vx\right), \tag{345}$$
$$x' = \gamma(x - vt), \tag{346}$$
$$y' = y, \tag{347}$$
$$z' = z. \tag{348}$$

Using the four-vector notation, this can be written as a matrix multiplication

$$\begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}' = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}. \tag{349}$$

Denoting the transformation matrix by $\mathbf{\Lambda}$, and we will actually slightly abuse the notation and denote it by $\Lambda^\mu{}_\nu$,

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{350}$$

We can write the transformation in terms of the four-vector components as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu, \tag{351}$$

where $\Lambda^\mu{}_\nu$ denotes the elements of the matrix $\mathbf{\Lambda}$. The first index in $\Lambda^\mu{}_\nu$ refers to the row and the second index to the column.
Under a Lorentz transformation of the form, the flat space-time interval transforms as

$$\Delta s^2 = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu, \tag{352}$$
$$\Delta s'^2 = \eta_{\mu\nu}\Delta x'^\mu \Delta x'^\nu = \eta_{\mu\nu}\Lambda^\mu{}_\alpha \Delta x^\alpha \Lambda^\nu{}_\beta \Delta x^\beta = \Delta x^\alpha \Lambda^\mu{}_\alpha \eta_{\mu\nu} \Lambda^\nu{}_\beta \Delta x^\beta. \tag{353}$$

Therefore, a Lorentz transformation is one that leaves the Minkowski metric invariant, i.e.

$$\boxed{\Lambda^\mu{}_\alpha \eta_{\mu\nu} \Lambda^\nu{}_\beta = \eta_{\alpha\beta}.} \tag{354}$$

Actually we can use this as the definition of a **Lorentz transformation**.

5.5 Poincaré transformations

A Poincaré transformations include both Lorentz transformations (Λ) and translations a^μ

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \tag{355}$$

where $\Lambda^\mu{}_\nu$ and a^μ are constants. Two consecutive Poincaré transformations that map from S to S' and then S' to S'' can be described by a single Poincaré transformation that maps from S to S''. To see this, consider the two consecutive transformations,

$$x^\mu \xrightarrow{\Lambda,a} x'^\mu \xrightarrow{\Lambda',a'} x''^\mu, \tag{356}$$

and express their product as a third transformation

$$x^\mu \xrightarrow{\Lambda'',a''} x''^\mu. \tag{357}$$

We then have

$$\begin{aligned} x''^\mu =& \Lambda''^\mu{}_\alpha x^\alpha + a''^\mu \\ =& \Lambda'^\mu{}_\nu x'^\nu + a'^\mu = \Lambda'^\mu{}_\nu(\Lambda^\nu{}_\alpha x^\alpha + a^\nu) + a'^\mu = \Lambda'^\mu{}_\nu \Lambda^\nu{}_\alpha x^\alpha + (\Lambda'^\mu{}_\nu a^\nu + a'^\mu). \end{aligned} \tag{358}$$

So $\Lambda''^\mu{}_\alpha = \Lambda'^\mu{}_\nu \Lambda^\nu{}_\alpha$ and $a''^\mu = \Lambda'^\mu{}_\nu a^\nu + a'^\mu$.

5.5.1 Lorentz group

The Lorentz group is the 6 parameter group of Lorentz boosts (3) and rotations (3). The Lorentz matrix associated with a boost of velocity v in the x-direction is

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{359}$$

The Lorentz matrix associated with a boost of velocity v in the y-direction is

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & 0 & -v\gamma & 0 \\ 0 & 1 & 0 & 0 \\ -v\gamma & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{360}$$

The Lorentz matrix associated with a boost of velocity v in the z-direction is

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & 0 & 0 & -v\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -v\gamma & 0 & 0 & \gamma \end{pmatrix}. \tag{361}$$

The Lorentz matrix associated with a rotation $\boldsymbol{M}$ is

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \boldsymbol{M} & \\ 0 & & & \end{pmatrix}. \tag{362}$$

5.6 Lorentz scalars and vectors

If a quantity remains invariant under a Lorentz transformation we shall call it a **Lorentz scalar**. The space-time interval $x \cdot y$ is a Lorentz scalar:

$$x \cdot y = \eta_{\mu\nu} x^\mu y^\nu \to (x \cdot y)' = \eta_{\mu\nu} x'^\mu y'^\nu = \eta_{\mu\nu} \Lambda^\mu{}_\alpha x^\alpha \Lambda^\nu{}_\beta x^\beta = x^\alpha \eta_{\alpha\beta} x^\beta = x \cdot y. \quad (363)$$

To simplify the notation further, we define a **covariant vector** x_μ by

$$x_\mu = \eta_{\mu\nu} x^\nu = (-x^0, x^1, x^2, x^3), \quad (364)$$

and indicate it by using a subscript index. We say that we use the metric to lower the index. The original position four-vector x^μ with a superscript index is called a **contravariant vector**. The scalar product is then

$$x \cdot y = x^\mu y_\mu = x_\mu y^\mu. \quad (365)$$

To raise the index, i.e., turn a covariant vector back to a contravariant one, we need the inverse of the metric tensor, so that

$$x^\mu = (\eta^{-1})^{\mu\alpha} x_\alpha = (\eta^{-1})^{\mu\alpha} \eta_{\alpha\nu} x^\nu = \delta^\mu{}_\nu x^\nu. \quad (366)$$

Therefore

$$(\eta^{-1})^{\mu\alpha} \eta_{\alpha\nu} = \delta^\mu{}_\nu = \mathbb{1}, \quad (367)$$

and we say that $(\eta^{-1})^{\mu\nu}$ is the inverse matrix of $\eta_{\mu\nu}$. We can simply write

$$(\eta^{-1})^{\mu\nu} = \eta^{\mu\nu}, \quad (368)$$

and the expression for raising the index simplifies to

$$x^\mu = \eta^{\mu\nu} x_\nu. \quad (369)$$

Similarly, we can also find the **inverse** Lorentz transformation matrix

$$\boxed{(\Lambda^{-1})^\lambda{}_\nu = \Lambda_\nu{}^\lambda.} \quad (370)$$

5.7 Transformation law for tensors

Tensors can be thought of as linear relations between a number of four-vectors. For example, if the four-vector x^μ is related to the following four-vectors y^μ, z^μ and w^μ through a linear relation (i.e., a rank 4 tensor), this can be expressed as

$$x^\mu = M^\mu{}_{\nu\alpha\beta} y^\nu z^\alpha w^\beta, \quad (371)$$

and we find

$$\begin{aligned} x'^\mu =& \Lambda^\mu{}_\lambda x^\lambda = \Lambda^\mu{}_\lambda M^\lambda{}_{\nu\alpha\beta} y^\nu z^\alpha w^\beta = \Lambda^\mu{}_\lambda M^\lambda{}_{\nu\alpha\beta} \Lambda_\alpha{}^\nu y'^\alpha \Lambda_\beta{}^\rho z'^\beta \Lambda_\gamma{}^\sigma w'^\gamma \\ =& \Lambda^\mu{}_\lambda \Lambda_\alpha{}^\nu \Lambda_\beta{}^\rho \Lambda_\gamma{}^\sigma M^\lambda{}_{\nu\alpha\beta} y'^\alpha z'^\beta w'^\gamma = M'^\mu{}_{\alpha\beta\gamma} y'^\alpha z'^\beta w'^\gamma. \end{aligned} \quad (372)$$

So we have

$$M'^\mu{}_{\alpha\beta\gamma} = \Lambda^\mu{}_\lambda \Lambda_\alpha{}^\nu \Lambda_\beta{}^\rho \Lambda_\gamma{}^\sigma M^\lambda{}_{\nu\alpha\beta}. \quad (373)$$

5.8 Action for a relativistic particle

For a relativistic particle the energy (Hamiltonian) is

$$H = \frac{mc^2}{\sqrt{1 - \frac{\dot{\boldsymbol{r}}^2}{c^2}}} = \sqrt{\boldsymbol{p}^2 c^2 + m^2 c^4}, \tag{374}$$

and the momentum is

$$\boldsymbol{p} = \frac{m\dot{\boldsymbol{r}}}{\sqrt{1 - \frac{\dot{\boldsymbol{r}}^2}{c^2}}}. \tag{375}$$

So the Lagrangian is

$$L = \boldsymbol{p} \cdot \dot{\boldsymbol{r}} - H = \frac{m\dot{\boldsymbol{r}}^2}{\sqrt{1 - \frac{\dot{\boldsymbol{r}}^2}{c^2}}} - \frac{mc^2}{\sqrt{1 - \frac{\dot{\boldsymbol{r}}^2}{c^2}}} = -mc^2\sqrt{1 - \frac{\dot{\boldsymbol{r}}^2}{c^2}}. \tag{376}$$

Then the action is

$$S = -mc^2 \int \mathrm{d}t \sqrt{1 - \frac{\dot{\boldsymbol{r}}^2}{c^2}} = mc \int \sqrt{-\eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu} = -mc^2 \int \mathrm{d}\tau. \tag{377}$$

6 Relativistic Electromagnetism

6.1 Relativistic Lorentz law

To derive the relativistic formulation, we start from the expression for the Lorentz force for a massive particle of charge q and velocity $\boldsymbol{v}$

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \boldsymbol{F} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}). \tag{378}$$

The Lorentz force as the derivative of the momentum with respect to the proper time τ as

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}\tau} = \gamma\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = q(u^0\boldsymbol{E} + \boldsymbol{u} \times \boldsymbol{B}), \tag{379}$$

where $u^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau$. The time derivative of the energy E of the particle is

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \boldsymbol{F} \cdot \boldsymbol{v} = q\boldsymbol{E} \cdot \boldsymbol{v}, \tag{380}$$

from which we obtain the derivative with respect to the proper time as

$$\frac{\mathrm{d}E}{\mathrm{d}\tau} = q\boldsymbol{E} \cdot \boldsymbol{u}. \tag{381}$$

Here, E is also the 0^{th} component of the four-momentum, $p^0 = E$. We can now write the proper time derivative of the four- momentum $p^\mu = (E, \boldsymbol{p})$,

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} = q\begin{pmatrix} 0 & E_x & E_y & E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}\begin{pmatrix} u^0 \\ u^1 \\ u^2 \\ u^3 \end{pmatrix}. \tag{382}$$

The matrix appearing in this expression is called the **Faraday tensor** or the **field-strength tensor** and denoted by $F^\mu{}_\nu$, so we can write more compactly the relativistic Lorentz force equation as

$$\boxed{\frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = qF^\mu{}_\nu u^\nu.} \tag{383}$$

The field-strength tensor is often written with two contravariant or two covariant indices as

$$F^{\mu\nu} = F^\mu{}_\alpha \eta^{\alpha\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix}, \tag{384}$$

or

$$F_{\mu\nu} = \eta_{\mu\alpha}F^\alpha{}_\nu = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}. \tag{385}$$

The electric and magnetic fields must transform under Lorentz transformations:

$$\boldsymbol{E}'_{\parallel} = \boldsymbol{E}_{\parallel}, \tag{386}$$
$$\boldsymbol{B}'_{\parallel} = \boldsymbol{B}_{\parallel}, \tag{387}$$
$$\boldsymbol{E}'_{\perp} = \gamma(\boldsymbol{E}_{\perp} + \boldsymbol{v} \times \boldsymbol{B}), \tag{388}$$
$$\boldsymbol{B}'_{\perp} = \gamma(\boldsymbol{B}_{\perp} + \boldsymbol{v} \times \boldsymbol{E}), \tag{389}$$

where $\parallel$ refers to the component parallel to the boost velocity $\boldsymbol{v}$, and $\perp$ to the perpendicular components.

6.2 Four-vector potential

In terms of the electric potential ϕ and the vector potential $\boldsymbol{A}$, the field strength tensor is

$$\begin{aligned} F_{\mu\nu} &= \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \dot{A}_x + \partial_x\phi & \dot{A}_y + \partial_y\phi & \dot{A}_z + \partial_z\phi \\ -(\dot{A}_x + \partial_x\phi) & 0 & \partial_x A_y - \partial_y A_x & \partial_x A_z - \partial_z A_x \\ -(\dot{A}_x + \partial_x\phi) & -(\partial_x A_y - \partial_y A_x) & 0 & \partial_y A_z - \partial_z A_y \\ -(\dot{A}_x + \partial_x\phi) & -(\partial_x A_z - \partial_z A_x) & -(\partial_y A_z - \partial_z A_y) & 0 \end{pmatrix}. \end{aligned} \tag{390}$$

We can write this in the form:

$$\boxed{F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,} \tag{391}$$

where $A_\mu = (-\phi, \boldsymbol{A})$ or equivalent $A_\mu = (\phi/c^2, \boldsymbol{A})$. We make it explicit here that we are dealing with Lorentz transformations as we will see another type of transformations shortly, related to gauge ones:

$$x^\mu \;\to\; \Lambda^\mu{}_\nu x^\nu, \tag{392}$$
$$A_\mu \;\to\; \Lambda_\mu{}^\nu A_\nu, \tag{393}$$
$$F_{\mu\nu} \;\to\; \Lambda_\mu{}^\alpha \Lambda_\nu{}^\beta F_{\alpha\beta}. \tag{394}$$

6.3 Action for a relativistic charged particle

The relativistic action for an charged particle is

$$S = \int \mathrm{d}\lambda \left(-mc\sqrt{-\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}} + qA_\mu(x)\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \right). \tag{395}$$

The Euler-Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial L}{\partial\left(\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right)} \right) = \frac{\partial L}{\partial x^\mu}, \tag{396}$$

and

$$\frac{\partial L}{\partial\left(\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right)} = \frac{mc\eta_{\mu\nu}}{\sqrt{-\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} + qA_\mu(x) = m\eta_{\mu\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} + qA_\mu(x) = p_\mu + qA_\mu(x), \quad (397)$$

where $\mathrm{d}\tau = \frac{1}{c}\sqrt{-\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu}$. So the Euler-Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(p_\mu + qA_\mu(x)\right) = q\frac{\partial A_\nu}{\partial x^\mu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}. \quad (398)$$

Rearranging the expression, and then

$$\frac{\mathrm{d}p_\mu}{\mathrm{d}\lambda} = q\frac{\partial A_\nu}{\partial x^\mu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} - q\frac{\partial A_\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} = qF_{\mu\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}. \quad (399)$$

Here, $F_{\mu\nu}$ is called the Faraday tensor or the field strength tensor

$$\boxed{F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.} \quad (400)$$

6.3.1 Gauge fixed version

We choose gauge $\lambda = t$, then

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = qF_{i\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}t} = qF_{i0} + qF_{ij}v^j, \quad (401)$$

where

$$F_{i0} = \partial_i A_0 - \partial_0 A_i = -\partial_i\phi - \frac{\partial A_i}{\partial t} = E_i, \quad (402)$$

$$F_{ij} = \partial_i A_j - \partial_j A_i = \varepsilon_{ijk}B_k. \quad (403)$$

Therefore

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = qE_i + q\varepsilon_{ijk}v^j B_k. \quad (404)$$

6.4 Relativistic Maxwell's equations

The dynamics of the electromagnetic field is described by Maxwell's equations,

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = \frac{\rho}{\varepsilon_0}, \quad \text{(Gauss's law)} \quad (405)$$

$$\boldsymbol{\nabla}\times\boldsymbol{E} = -\frac{\partial\boldsymbol{B}}{\partial t}, \quad \text{(Faraday's law)} \quad (406)$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0, \quad \text{(magnetic Gauss's law)} \quad (407)$$

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \mu_0\boldsymbol{J} + \mu_0\varepsilon_0\frac{\partial\boldsymbol{E}}{\partial t}. \quad \text{(Ampère's law)} \quad (408)$$

Now we can derive these in terms of four-vector potential. Begin with the Faraday tensor, we can find the property

$$\boxed{\partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu} + \partial_\alpha F_{\mu\nu} = 0.} \quad (409)$$

So we have

$$\partial_z F_{xy} + \partial_x F_{yz} + \partial_y F_{zx} = \partial_x B_x + \partial_y B_y + \partial_z B_z = \boldsymbol{\nabla} \cdot \boldsymbol{B} = 0, \tag{410}$$

which is the **magnetic Gauss's law**. Similarly,

$$\partial_0 F_{xy} + \partial_x F_{y0} + \partial_y F_{0x} = \frac{\partial B_z}{\partial t} + \partial_x E_y - \partial_y E_x = 0, \tag{411}$$

is the **Faraday's law**.

$$\partial_\mu F^\mu{}_\nu = \frac{\partial}{\partial x^\mu} F^\mu{}_\nu = \begin{pmatrix} \partial_x E_x + \partial_y E_y + \partial_z E_z \\ \frac{1}{c^2}\partial_t E_x - \partial_y B_z + \partial_z B_y \\ \frac{1}{c^2}\partial_t E_y + \partial_x B_z - \partial_z B_x \\ \frac{1}{c^2}\partial_t E_z - \partial_x B_y + \partial_y B_x \end{pmatrix} = \begin{pmatrix} \boldsymbol{\nabla} \cdot \boldsymbol{E} \\ \frac{1}{c^2}\frac{\partial \boldsymbol{E}}{\partial t} - \boldsymbol{\nabla} \times \boldsymbol{B} \end{pmatrix}. \tag{412}$$

Combining the charge density ρ and the current $\boldsymbol{J}$ into a single four-current $J^\mu = (\rho, \boldsymbol{J})$, for which $J_\mu = (-\rho c^2, \boldsymbol{J})$ we therefore see that **Gauss's and Ampère's law** take the form

$$\boxed{\partial_\mu F^\mu{}_\nu = -\mu_0 J_\nu.} \tag{413}$$

The conservation law

$$\boxed{\partial_\mu J^\mu = 0,} \tag{414}$$

is equivalent to the continuity equation $\dot{\rho} + \boldsymbol{\nabla} \cdot \boldsymbol{J} = 0$.

6.5 Gauge transformations

Consider the following the transformations

$$\tilde{A}_\mu = A_\mu + \partial_\mu \chi. \tag{415}$$

This would change the field strength as

$$\tilde{F}_{\mu\nu} = \partial_\mu \tilde{A}_\nu - \partial_\nu \tilde{A}_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}. \tag{416}$$

Therefore the transformation is a symmetry.

6.5.1 Lorenz gauge

Lorenz gauge is defined by

$$\partial_\mu A^\mu = 0. \tag{417}$$

We can write

$$\partial_\mu F^\mu{}_\nu = \partial^\mu F_{\mu\nu} = \partial^\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) = \Box A_\nu - \partial_\nu(\partial^\mu A_\mu) = \Box A_\nu, \tag{418}$$

where $\Box = \eta^{\mu\mu}\partial_\mu\partial_\mu = -\frac{1}{c^2}\partial_t^2 + \boldsymbol{\nabla}^2$. So the Gauss's and Ampère's law can be written as

$$\Box A_\nu = -\mu_0 J_\nu. \tag{419}$$

6.6 Lagrangian for electrodynamics

The Lagrangian density $\mathcal{L}$ has to be a **Lorentz scalar**, and to be **gauge invariant**, which means the Lagrangian density is *invariant* under both Lorentz ($A_\mu \to A_\mu{}^\alpha A_\alpha$) and gauge transformations ($A_\mu \to A_\mu + \partial_\mu \chi$). The Lagrangian for electrodynamics is

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \mu_0 A_\mu J^\mu. \tag{420}$$

Here,

$$F^{\mu\nu} F_{\mu\nu} = (\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) = 2\partial^\mu A^\nu (\partial_\mu A_\nu - \partial_\nu A_\mu). \tag{421}$$