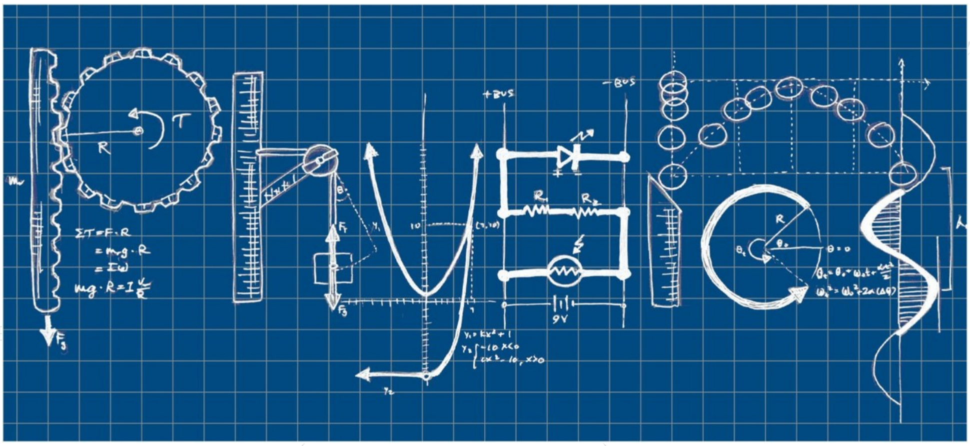
Physics

物理



Do not ask what it is. Ask what you can say about it.

Editor: Yuyang Songsheng

Date: October 13, 2016

Email: songshengyuyang@gmail.com

Version: 1.00

Contents

Part I

Classical Mechanics

Chapter 1
The formulation of Classical Mechanics

1.1 Lagrangian Formulation

$$S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)dt, \quad \delta q_i(t_1) = \delta q_i(t_2) = 0$$

$$\delta S = 0 \rightarrow \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}) - \frac{\partial L}{\partial q_i} = 0$$

1. If we transform the coordinates q to the Q as $q = q(Q, t)$, the new Lagrangian will be

   $$\bar{L}(Q, \dot{Q}, t) \equiv L(q(Q, t), \dot{q}(Q, \dot{Q}, t), t)$$

   We can verify that

   $$\frac{d}{dt}\frac{\partial \bar{L}}{\partial \dot{Q}} - \frac{\partial \bar{L}}{\partial Q} = 0$$

2. If $L_1 = L + \frac{d}{dt}f(q, t)$, then L and L_1 is equivalent and will generate the same dynamical equation.

Example:

1. The form of Lagrangian for an isolated system of particles in inertial frame:

   $$L = \sum_a \frac{1}{2} m_a v_a^2 - U(\vec{r}_1, \vec{r}_2, \cdots,)$$

   The equation of motion is

   $$m_i \ddot{\vec{r}}_i = -\nabla_{\vec{r}_i} U$$

   To get the form of Lagrangian for a system of interacting particles, we must assume:

   - Space and time are homogeneous and isotropic in inertial frame;
   - Galileo's relativity principle and Galilean transformation;
   - Spontaneous interaction between particles;

2. Consider a reference frame K. Suppose the K is moving with the velocity $\mathbf{V}(t)$ and rotating with angular velocity $\mathbf{\Omega}$ relative to the inertial reference frame. We use the

coordinates of the mass point in K as general coordinates, i.e. $\mathbf{r} = (x_k, y_k, z_k)$. Then the Lagrangian of the mass point will be

$$L = \frac{1}{2}m\mathbf{v}^2 + m\mathbf{v}\cdot(\mathbf{\Omega}\times\mathbf{r}) + \frac{m}{2}(\mathbf{\Omega}\times\mathbf{r})^2 - m\dot{\mathbf{V}}\cdot\mathbf{r} - U$$

The equation of motion will be

$$m\frac{d\mathbf{v}}{dt} = -\frac{\partial U}{\partial \mathbf{r}} - m\dot{\mathbf{V}} + m(\mathbf{r}\times\dot{\mathbf{\Omega}}) + 2m(\mathbf{v}\times\mathbf{\Omega}) + m[\mathbf{\Omega}\times(\mathbf{r}\times\mathbf{\Omega})]$$

1.2 Symmetry and Conservation Laws(1)

Theorem 1.1 Nother's theorem

For $q_i \to q_i + \delta q_i$ and $L \to L + \delta L$, if $\delta L = \frac{df(q,\dot{q},t)}{dt}$,then we get

$$\frac{d}{dt}(\sum_i p^i \delta q_i - f) = 0 \quad (p^i = \frac{\partial L}{\partial \dot{q}_i})$$

Example: For an isolated system of particles in inertial frame,
$\delta L = 0$ when $\delta\vec{r}_i \to \vec{r}_i + \delta\vec{a}$, so

$$\frac{d}{dt}(\sum_i \vec{p}_i) = 0$$

$\delta L = 0$ when $\delta\vec{r}_i \to \vec{r}_i + \vec{r}_i \times \delta\vec{\theta}$, so

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

Homogeneity of time If $\frac{\partial L}{\partial t} = 0$,then we get

$$\frac{dE}{dt} = 0 \quad (E = \sum_i \dot{q}_i p^i - L)$$

1.3 Hamilton formulation

$$p^i = \frac{\partial L}{\partial \dot{q}_i}$$

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

$$\dot{p^i} = -\frac{\partial H}{\partial q_i} \quad \dot{q}_i = \frac{\partial H}{\partial p^i}$$

Example: For an isolated system of particles in inertial frame,

$$\vec{p_i} = m_i \vec{v_i}$$

$$H(q,p,t) = \sum_i \frac{p_i^2}{2m} + U(\vec{r}_1, \vec{r}_2, \cdots)$$

$$\dot{\vec{p_i}} = -\nabla_{\vec{r}_i} U \quad \dot{r}_i = \frac{\vec{p_i}}{m_i}$$

1.3.1 Poisson Brackets

First, we assume the bracket operation has the following properties:

$$[f,g] = -[g,f]$$

$$[\alpha_1 f_1 + \alpha_2 f_2, \beta_1 g_1 + \beta_2 g_2] = \alpha_1\beta_1 [f_1, g_1] + \alpha_1\beta_2 [f_1, g_2] + \alpha_2\beta_1 [f_2, g_1] + \alpha_2\beta_2 [f_2, g_2]$$
$$[f_1 f_2, g_1 g_2] = f_1 [f_2, g_1] g_2 + f_1 g_1 [f_2, g_2] + g_1 [f_1, g_2] f_2 + [f_1, f_2] g_2 f_2$$
$$[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = 0$$

Here, f, g, h are functions of p^i, q_i, t. Then, we assume that

$$\left[q_i, p^k\right] = \delta_i^k$$

we can derive that

$$[f,g] = \sum_k \left(\frac{\partial f}{\partial q_k}\frac{\partial g}{\partial p^k} - \frac{\partial f}{\partial p^k}\frac{\partial g}{\partial q_k}\right)$$

So the Hamilton equation can be written as

$$\dot{p^i} = \left[p^i, H\right] \quad \dot{q}_i = [q_i, H]$$

And we can also get

$$\frac{df}{dt} = [f,H] + \frac{\partial f}{\partial t} \quad \frac{d}{dt}[f,g] = [\frac{df}{dt}, g] + [f, \frac{dg}{dt}]$$

Example: For an isolated system of particles in inertial frame,

$$[r_{ia}, p_{jb}] = \delta_{ab}\delta_{ij}$$

we define $l_a = \epsilon_{abc} r_a p_b$, then

$$[l_a, r_b] = \epsilon_{abc} r_c \quad [l_a, p_b] = \epsilon_{abc} p_c \quad [l_a, l_b] = \epsilon_{abc} l_c$$

1.3.2 Canonical transformations

In Hamiltonian mechanics, a canonical transformation is a change of canonical coordinates that preserves the form of Hamilton's equations (that is, the new Hamilton's equations resulting from the transformed Hamiltonian may be simply obtained by substituting the new coordinates for the old coordinates), although it might not preserve the Hamiltonian itself.

$$Q_i = Q_i(p,q,t) \quad P_i = P_i(p,q,t)$$

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

Proposition 1.1 Canonical condition

If $(q_i, p^i, H) \to (Q_i, P^i, H)$ is a canonical transformation, then there exists a generating function $F(q_i, Q_i, t)$ satisfying that

$$\sum_i p^i \dot{q}_i - H(p^i, q_i) = \sum_i P^i \dot{Q}_i - H'(Q_i, P^i) + \frac{dF}{dt}$$

Applying Legendre transformation, we can get four kinds of generating function.

1.
$$\frac{dF}{dt} = \sum_i p^i \dot{q}_i - \sum_i P^i \dot{Q}^i + (H' - H)$$
Assume $\Phi(q_i, Q_i, t) = F$, so
$$p^i = \frac{\partial \Phi}{\partial q_i} \quad P^i = -\frac{\partial \Phi}{\partial Q_i} \quad H' = H + \frac{\partial \Phi}{\partial t}$$

2.
$$\frac{d}{dt}(F + \sum_i P^i Q_i) = \sum_i p^i \dot{q}_i + \sum_i Q_i \dot{P}^i + (H' - H)$$
Assume $\Phi(q_i, P^i, t) = F + \sum_i P^i Q_i$, so
$$p^i = \frac{\partial \Phi}{\partial q_i} \quad Q_i = \frac{\partial \Phi}{\partial P^i} \quad H' = H + \frac{\partial \Phi}{\partial t}$$

3.
$$\frac{d}{dt}(F - \sum_i p^i q_i) = -\sum_i q_i \dot{p}^i - \sum_i P^i \dot{Q}_i + (H' - H)$$
Assume $\Phi(p^i, Q_i, t) = F - \sum_i p^i q_i$, so
$$q_i = -\frac{\partial \Phi}{\partial p^i} \quad P^i = -\frac{\partial \Phi}{\partial Q_i} \quad H' = H + \frac{\partial \Phi}{\partial t}$$

4.
$$\frac{d}{dt}(F+\sum_i P^iQ_i-\sum_i p^iq_i)=-\sum_i q_i\dot{p}^i+\sum_i Q_i\dot{P}^i+(H'-H)$$
Assume $\Phi(p^i,P^i,t)=F+\sum_i P^iQ_i-\sum_i p^iq_i$, so
$$q_i=-\frac{\partial\Phi}{\partial p^i}\quad Q_i=\frac{\partial\Phi}{\partial P^i}\quad H'=H+\frac{\partial\Phi}{\partial t}$$

Theorem 1.2 The invariance of Poisson Bracket

Suppose that $(q,p,H)\to(Q,P,H')$ is a canonical transformation and $f(q,p,t)=F(Q,P,t)$, $g(q,p,t)=G(Q,P,t)$, then
$$[f,g]_{q,p}=[F,G]_{Q,P}$$

♣

As a result, the condition for canonical transformation can also be stated as
$$[Q_i,Q_j]_{q,p}=0\quad [P^i,P^j]_{p,q}=0\quad [Q_i,P^j]_{q,p}=\delta_i^j$$

1.3.3 Evolution as canonical transformations

Let q_t,p_t be the values of the canonical variables at time t, and $q_{t+\tau}$,$p_{t+\tau}$ their values at another time $t+\tau$. The latter are some functions of the former:
$$q_{t+\tau}=q(q_t,p_t,t,\tau)\quad p_{t+\tau}=p(q_t,p_t,t,\tau)$$

If these formulae are regarded as a transformation from the variables q_t,p_t to $q_{t+\tau}$, $p_{t+\tau}$, then this transformation is canonical. This is evident from the expression
$$dS=p_tdq_t+p_{t+\tau}dq_{t+\tau}-(H_{t+\tau}-H_t)dt$$
for the differential of the action $S(q_t,q_{t+\tau},t,\tau)$, taken along the true path, passing through the points q, and $q_{t+\tau}$ at times t and $t+\tau$ for a given τ. $-S$ is the generating function of the transformation. So we have the following communication relation
$$[q_{i\,t+\tau},q_{j\,t+\tau}]_{q_t,p_t}=0\quad [p^i_{t+\tau},p^j_{t+\tau}]_{q_t,p_t}=0\quad [q_{i\,t+\tau},p^j_{t+\tau}]_{q_t,p_t}=\delta_i^j$$

1.3.4 Liouville's theorem

Lemma 1

Let D be the Jacobian of the canonical transformation
$$\frac{\partial(Q_1,\cdots,Q_s,P^1,\cdots,P^s)}{\partial(q_1,\cdots,q_s,p^1,\cdots,p^s)}$$
Then we have
$$D=1$$

♣

Theorem 1.3 Liouville's theorem

The phase-space distribution function is constant along the trajectories of the system

Proof: The phase volume is invariant under canonical transformation.The change in p and q during the motion can be regarded as a canonical transformation. Suppose that each point in the region of phase space moves in the course of time in accordance with the equations of motion of the mechanical system. The region as a whole therefore moves also, but its volume remains unchanged. □

1.4 Symmetry and Conservation Laws(2)

Suppose g is a function of p and q. If the transformation of q and p can be described as

$$q \to q + \epsilon[q, g]$$

$$p \to p + \epsilon[p, g]$$

We can prove that

$$H \to H + \epsilon[H, g]$$

So if H is invariant under the transformation, then $[H, g] = 0$, that means $\frac{dg}{dt} = 0$, i.e. g is a conserved quantity of the motion.

1.5 Hamilton-Jacobi equation

We define

$$S(q, t) = \left(\int_{q_0,t_0}^{q,t} Ldt\right)|_{extremum}$$

We can prove that

$$p = \frac{\partial S}{\partial q}, \quad H = -\frac{\partial S}{\partial t}$$

So, we have

$$-\frac{\partial S}{\partial t} = H(q, \frac{\partial S}{\partial q})$$

This is called Hamiltonian-Jacobi equation.

Suppose the complete integral of the Hamilton-Jacobi equation is

$$S = f(t, q_1, \cdots, q_s; \alpha^1, \cdots, \alpha^s) + A$$

where $\alpha^1, \cdots, \alpha^s$ and A are arbitrary constants. We effect a canonical transformation from the variables q, p to new variables, taking the function $f(t, q, \alpha)$ as the generating function, and the quantities $\alpha^1, \cdots, \alpha^s$ as the new momenta. Let the new co-ordinates be $\beta_1, \cdots, \beta_2$.

$$p^i = \frac{\partial f}{\partial q_i} \quad \beta_s = \frac{\partial f}{\partial \alpha_s} \quad H' = H + \frac{\partial f}{\partial t} = 0$$

So,

$$\alpha^s = \text{ constant }, \beta_s = \text{ constant}$$

By means of the s equations $\beta_s = \frac{\partial f}{\partial \alpha^s}$, the s coordinates q can be expressed in terms of the time and the $2s$ constants. This gives the general integral of the equations of motion.

1.6 Symmetry and Conservation Laws(3)

If S is invariant under transformation $q_i \to q_i + \delta q_i$, then

$$\delta S = (\sum_i p^i \delta q_i)|_{q_0,t_0}^{q,t} = 0$$

So, we have

$$\frac{d}{dt}(p^i \delta q_i) = 0$$

Further more, if

$$\delta S = (\sum_i p^i \delta q_i)|_{q_0,t_0}^{q,t} = f(q_i, \dot{q}_i, t)|_{q_0,t_0}^{q,t}$$

we will have conserved quantity

$$\frac{d}{dt}(p^i \delta q_i - f) = 0$$

Part II

Classical Field Theory

Chapter 2
Mechanics within special relativity

2.1 Basic Assumption

First, we assume there is an upper limit of velocity of propagation of interaction c. Second, we assume that inertial reference frame are all the same in describing the law of physics. Then, we can find the invariant intervals when transforming from one inertial reference frame to another, $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$. (In the following, we assume that $c = 1$.)This transformation is called Lorentz transformation, which can be written as

$$\bar{x}^\mu = \Lambda^\mu{}_\nu x^\nu$$

and it is easy to verify that

$$\eta_{\mu\nu}\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma = \eta_{\rho\sigma},$$

where,

$$\eta_{\mu\nu} = \begin{bmatrix} -1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1 \end{bmatrix}$$

and

$$(\Lambda^{-1})^\rho{}_\nu = \Lambda_\nu{}^\rho$$

In a special case when the new reference frame move along $\hat{1}$ direction with velocity β, we have

$$\bar{x}^0 = \gamma x^0 - \gamma\beta x^1$$

$$\bar{x}^1 = -\gamma\beta x^0 + \gamma x^1$$

Some physical quantity will behave like a tensor (vector,scalar) when transforming form one inertial frame to another. For example,

scalar proper time: $d\tau$, mass: m, electrical charge e

vector four velocity: $v^\mu = \frac{dx^\mu}{d\tau}$, four momentum: $p^\mu = mv^\mu$, four acceleration: $a^\mu = \frac{du^\mu}{d\tau}$, four force: $f^\mu = ma^\mu$

2.2 "Three vector"

three velocity: $\hat{u}^i = \frac{dx^i}{dt}$

$$u^0 = \gamma_v, u^i = \gamma \hat{u}^i$$

transformation of three velocity when we boost along $\hat{1}$ direction:

$$\bar{\hat{v}}^1 = \frac{\hat{v}^1 - \beta}{1 - \hat{v}^1 \beta}$$

$$\bar{\hat{v}}^2 = \frac{\hat{v}^2}{\gamma(1 - \hat{v}^2 \beta})$$

$$\bar{\hat{v}}^3 = \frac{\hat{v}^3}{\gamma(1 - \hat{v}^3 \beta})$$

three momentum: $\hat{p}^i = p^i$

$$\hat{p}^i \gamma_v \hat{v}^i$$

three acceleration: $\hat{a}^i = \frac{dv^i}{dt}$
three force: $\hat{f}^i = \frac{d\hat{p}^i}{dt}$

$$f^i = \gamma_v \hat{f}^i$$

Energy: $E = p^0 = mu^0 = \gamma_v m$

2.3 Mechanics

Revised newton's second law:

$$f^\mu = \frac{dp^\mu}{d\tau}$$

It can be written in three vector language as

$$\hat{f}^i = \gamma_v m \hat{a}^i + \gamma_v^3 (\hat{a}^j \hat{v}_j) m \hat{v}^i$$

2.4 Lagrangian formulation

$$S = -m \int_a^b d\tau, \quad \delta x^\mu(a) = \delta x^\mu(b) = 0$$

$$\delta S = 0 \Rightarrow m \frac{du^\mu}{d\tau} = 0$$

2.5 Hamiltonian formulation

$$S = -m \int_{t_1}^{t_2} \sqrt{1 - \dot{x}_i \dot{x}^i} dt$$
$$L = -m\sqrt{1 - \dot{x}_i \dot{x}^i}$$
$$\pi^i = \frac{\partial L}{\partial \dot{x}_i} = \gamma m \eta^{ij} \dot{x}_j$$
$$H = \pi^i \dot{x}_i - L = \gamma m = \sqrt{m^2 + \pi^i \pi_i}$$

Hamilton equation

$$\dot{\pi}^i = 0, \quad \dot{x}_i = \eta_{ij} \frac{\pi^j}{\sqrt{m^2 + \pi^k \pi_k}}$$

Hamiltonian-Jacobi equation

$$H = -\frac{\partial S}{\partial t}, \quad \pi^i = \frac{\partial S}{\partial x_i}$$

If we define $p^0 = H$, $p^i = \pi^i$, then we can verify that $p^\mu = \frac{\partial S}{\partial x_\mu}$. So, p^μ is a vector under Lorentz transformation. The Hamiltonian-Jacobi equation can be written as

$$(\frac{\partial S}{\partial t})^2 = m^2 + (\frac{\partial S}{\partial x})^2 + (\frac{\partial S}{\partial y})^2 + (\frac{\partial S}{\partial z})^2$$

2.6 Symmetry and conservation law

Translational symmetry and conservation of momentum

$$\bar{x}^\mu = x^\mu + \delta x^\mu$$
$$\delta S = \sum m u_\mu \delta x^\mu |_a^b = 0$$

$\sum p^\mu$ is conserved.

Rotational symmetry and conservation of angular momentum

$$\bar{x}^\mu = x^\mu + x_\nu \delta \Omega^{\mu\nu}$$
$$\delta S = \sum m u^\mu x^\nu \delta \Omega_{\mu\nu} |_a^b = 0$$

$\sum M^{\mu\nu}$ is conserved, where $M^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu$.

Chapter 3
Classical field theory

3.1 Lagrangian formulation

$$S = \int \mathcal{L}(\phi_a, \dot{\phi}_a, \nabla\phi_a) d^4x, \quad \delta\phi_a|_\Sigma = 0$$

$$\delta S = 0 \Rightarrow \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_a} = 0$$

Locality of the theory

There are no terms in the Lagrangian coupling $\phi(\vec{x}, t)$ directly to $\phi(\vec{y}, t)$ with $\vec{x} \neq \vec{y}$. The closet we get for the $\vec{x}$ label is coupling between $\phi(\vec{x}, t)$ and $\phi(\vec{x} + \delta\vec{x}, t)$ through the gradient term $\nabla\phi$.

Lorentz invariance

Scalar fields:

$$\bar{\phi}(x) = \phi(\Lambda^{-1}x)$$

Vector fields:

$$\bar{A}^\mu(x) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x)$$

$$\bar{A}_\mu(x) = (\Lambda^{-1})^\nu{}_\mu A_\nu(\Lambda^{-1}x) = \Lambda_\mu{}^\nu A_\nu(\Lambda^{-1}x)$$

$$\overline{\partial_\mu \phi}(x) = (\Lambda^{-1})^\nu{}_\mu \partial_\nu \phi(\Lambda^{-1}x) = \Lambda_\mu{}^\nu \partial_\nu \phi(\Lambda^{-1}x)$$

Lagrangian is a scalar, or more loosely, action is invariant under Lorentz transformation.

3.2 Symmetry and conservation law

Theorem 3.1 Noether's theorem

Every continuous symmetry of the Lagrangian gives rise to a conserved current $j^\mu(x)$ such that the equation of motion imply $\partial_\mu j^\mu = 0$. Suppose that the infinitesimal transformation is

$$\phi_a \to \phi_a + \delta\phi_a$$

$$\mathcal{L} \to +\mathcal{L} + \delta\mathcal{L}$$

and if $\delta\mathcal{L} = \partial_\mu K^\mu = 0$, we can get

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a - K^\mu$$

space-time translation

$\bar{x} = x - a$

$$j^\mu = -a_\nu T^{\mu\nu}$$

$$T^{\mu\nu} \equiv -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial^\nu\phi_a + \eta^{\mu\nu}\mathcal{L}$$

If we define $P^\mu = \int T^{0\mu}d^3x$,then we have

$$\frac{dP^\mu}{dt} = 0$$

Lorentz Transformation

$\bar{x}^\mu = x^\mu + \delta\omega^\mu{}_\nu x^\nu$
The infinitesimal Lorentz transformation can be written as $I + \delta\omega^\mu{}_\nu$

$$\delta\omega^\mu{}_\nu = \begin{bmatrix} 0 & \beta_1 & \beta_2 & \beta_3 \\ \beta_1 & 0 & -\theta_3 & \theta_2 \\ \beta_2 & \theta_3 & 0 & -\theta_1 \\ \beta_3 & -\theta_2 & \theta_1 & 0 \end{bmatrix}$$

This time, we assume that

$$\bar{\phi}_a(x) = S_a{}^b\phi_b(\Lambda^{-1}x)$$

In the limit of infinitesimal Lorentz transformation, we have

$$S_a{}^b = \delta_a{}^b + \frac{1}{2}\delta\omega_{\alpha\beta}(\Sigma^{\alpha\beta})_a{}^b$$

$$j^\mu = -\frac{1}{2}M^{\mu\nu\rho}\delta\omega_{\nu\rho}$$

$$M^{\mu\nu\rho} \equiv x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} (\Sigma^{\nu\rho})_a{}^b \phi_b$$

If we define $M^{\nu\rho} = \int M^{0\nu\rho} d^3x$, then we have

$$\frac{dM^{\nu\rho}}{dt} = 0$$

3.3 Functional derivatives

Definition 3.1 Functional derivatives

Given a manifold M representing (continuous/smooth) functions ρ (with certain boundary conditions etc.), and a functional F defined as

$$F\colon M \to \mathbb{R} \quad \text{or} \quad F\colon M \to \mathbb{C}\,,$$

the functional derivative of $F[\rho]$, denoted $\frac{\delta F}{\delta \rho}$,is defined by

$$\begin{aligned}\int \frac{\delta F}{\delta \rho}(x)\phi(x)\ dx &= \lim_{\varepsilon \to 0} \frac{F[\rho + \varepsilon\phi] - F[\rho]}{\varepsilon} \\ &= \left[\frac{d}{d\epsilon} F[\rho + \epsilon\phi]\right]_{\epsilon=0},\end{aligned}$$

where ϕ is an arbitrary function. The quantity $\epsilon\phi$ is called the variation of ρ.

Like the derivative of a function, the functional derivative satisfies the following properties, where $F[\rho]$ and $G[\rho]$ are functionals:
Linearity:

$$\frac{\delta(\lambda F + \mu G)[\rho]}{\delta \rho(x)} = \lambda \frac{\delta F[\rho]}{\delta \rho(x)} + \mu \frac{\delta G[\rho]}{\delta \rho(x)},$$

where λ, μ are constants.
Product rule:

$$\frac{\delta(FG)[\rho]}{\delta \rho(x)} = \frac{\delta F[\rho]}{\delta \rho(x)} G[\rho] + F[\rho] \frac{\delta G[\rho]}{\delta \rho(x)}\,,$$

Chain rules:
If F is a functional and G an operator, then

$$\frac{\delta F[G[\rho]]}{\delta \rho(y)} = \int dx \left.\frac{\delta F[G]}{\delta G(x)}\right|_{G=G[\rho]} \cdot \frac{\delta G[\rho](x)}{\delta \rho(y)}\,.$$

If G is an ordinary differentiable function g, then this reduces to

$$\frac{\delta F[g(\rho)]}{\delta \rho(y)} = \frac{\delta F[g(\rho)]}{\delta g[\rho(y)]} \frac{dg(\rho)}{d\rho(y)}\,.$$

Proposition 3.1 Properties of functional derivatives

$$\frac{\delta F}{\delta \rho}(y) = \lim_{\epsilon \to \infty} \frac{1}{\epsilon}\{F[\rho(x) + \epsilon\delta(x-y)] - F[\rho(x)]\}$$

$$\frac{\delta f(x)}{\delta f(y)} = \delta(x-y)$$

$$\frac{\delta}{\delta f(y)} \int g\left(f(x)\right) dx = g'\left(f(y)\right)$$

$$\frac{\delta f'(x)}{\delta f(y)} = \frac{d}{dx}\delta(x-y)$$

$$\frac{\delta}{\delta f(y)} \int g\left(f'(x)\right) dx = -\frac{d}{dy} g'\left(f'(y)\right)$$

3.4 Hamiltonian formulation

$$\pi^a(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a}$$

$$\mathcal{H}(\phi_a, \nabla\phi_a, \pi^a) = \pi^a \dot{\phi}_a - \mathcal{L}$$

$$H = \int \mathcal{H} d^3x$$

Now, we can get the Hamilton equation form $\delta S = 0$,

$$\dot{\phi}_a(\vec{x}, t) = \frac{\delta}{\delta \pi^a(\vec{x}, t)} H = \frac{\partial \mathcal{H}}{\partial \pi^a}$$

$$\dot{\pi^a}(\vec{x}, t) = -\frac{\delta}{\delta \phi_a(\vec{x}, t)} H = -\frac{\partial \mathcal{H}}{\partial \phi_a} + \left(\frac{\partial \mathcal{H}}{\partial \phi_{a,b}}\right)_{,b}$$

3.4.1 Poission bracket

First, we demand that

$$[\phi_a(\vec{x}), \phi_b(\vec{y})] = [\pi^a(\vec{x}), \phi_b(\vec{y})] = 0$$

$$[\phi_a(\vec{x}), \pi^b(\vec{y})] = \delta_a^b \delta(\vec{x} - \vec{y})$$

then, we assume the bracket operation has the same properties as the Poission bracket in classical mechanics. And we also assume that

$$[\partial_x A(\vec{x}), B(\vec{y})] = \partial_x [A(\vec{x}), B(\vec{y})]$$

and

$$\left[\int d^3x A(\vec{x}), B(\vec{y})\right] = \int d^3x [A(\vec{x}), B(\vec{y})]$$

We can verify that

$$[W[\phi(\vec{x}),\pi(\vec{x})],Z[\phi(\vec{x}),\pi(\vec{x})]]=\int d^3x\left\{\frac{\delta W}{\delta\phi(\vec{x})}\frac{\delta Z}{\delta\pi(\vec{x})}-\frac{\delta W}{\delta\pi(\vec{x})}\frac{\delta Z}{\delta\phi(\vec{x})}\right\}$$

Specially,

$$[\phi_a(\vec{x}),H]=\frac{\delta}{\delta\pi^a(\vec{x})}H,\quad [\pi^a(\vec{x}),H]=-\frac{\delta}{\delta\phi_a(\vec{x})}H$$

So, the Hamilton equation can be written as

$$\dot{\phi}_a=[\phi_a,H],\quad \dot{\pi^a}=[\pi^a,H]$$

Further more, we can prove

$$\frac{dO(\phi,\pi,t)}{dt}=[O,H]+\frac{\partial O}{\partial t}$$

and

$$\frac{d[A,B]}{dt}=[A,\frac{dB}{dt}]+[\frac{dA}{dt},B]$$

3.4.2 Momentum

It is easy to verify that

$$P^0=H,\quad P^i=\int -\pi^a\partial^i\phi_a d^3x$$

And we can get the commutation relationship that

$$\begin{aligned}[][\phi_a,P^\mu] &= -\partial^\mu\phi_a\\ [\pi^a,P^\mu] &= -\partial^\mu\pi^a\\ [P^\mu,P^\nu] &= 0\end{aligned}$$

3.4.3 Angular momentum

It is easy to verify that

$$M^{\mu\nu}=\int(x^\mu T^{0\nu}-x^\nu T^{0\mu}-\pi^a(\Sigma^{\mu\nu})_a{}^b\phi_b)d^3x$$

We denote that

$$M_L^{\mu\nu}=\int(x^\mu T^{0\nu}-x^\nu T^{0\mu})d^3x\quad M_S^{\mu\nu}=\int(-\pi^a(\Sigma^{\mu\nu})_a{}^b\phi_b)d^3x$$

$$(L^{\mu\nu})_a{}^b=-(x^\mu\partial^\nu-x^\nu\partial^\mu)\delta_a{}^b\quad (S^{\mu\nu})_a{}^b=-(\Sigma^{\mu\nu})_a{}^b$$

So, we have the commutation relationship that

$$M^{\mu\nu}=M_L^{\mu\nu}+M_S^{\mu\nu}$$

$$[\phi_a,M_L^{\mu\nu}]=(L^{\mu\nu})_a{}^b\phi_b\quad [\phi_a,M_S^{\mu\nu}]=(S^{\mu\nu})_a{}^b\phi_b$$

$$[\pi^a, M_L^{\mu\nu}] = (L^{\mu\nu})_b{}^a \pi^b \quad [\pi^a, M_S^{\mu\nu}] = -(S^{\mu\nu})_b{}^a \pi^b$$

Because $\frac{dM^{\mu\nu}}{dt} = 0$, we can prove that

$$[[\phi(x), M^{\mu\nu}], M^{\rho\sigma}] = (L^{\mu\nu} + S^{\mu\nu})(L^{\rho\sigma} + S^{\rho\sigma})\phi(x)$$

and then we can get the communication relationship from the Jacobi identity,

$$[\phi(x), [M^{\mu\nu}, M^{\rho\sigma}]] = (L^{\mu\nu}L^{\rho\sigma} - L^{\rho\sigma}L^{\mu\nu} + S^{\mu\nu}S^{\rho\sigma} - S^{\rho\sigma}S^{\mu\nu})\phi(x)$$

We can prove that

$$L^{\mu\nu}L^{\rho\sigma} - L^{\rho\sigma}L^{\mu\nu} = -g^{\nu\rho}L^{\mu\sigma} + g^{\sigma\mu}L^{\rho\nu} + g^{\mu\rho}L^{\nu\sigma} - g^{\sigma\nu}L^{\rho\mu}$$

If we demand that

$$S^{\mu\nu}S^{\rho\sigma} - S^{\rho\sigma}S^{\mu\nu} = -g^{\nu\rho}S^{\mu\sigma} + g^{\sigma\mu}S^{\rho\nu} + g^{\mu\rho}S^{\nu\sigma} - g^{\sigma\nu}S^{\rho\mu}$$

We can get get the communication relationship of the $M^{\mu\nu}$,

$$[M^{\mu\nu}, M^{\rho\sigma}] = -g^{\nu\rho}M^{\mu\sigma} + g^{\sigma\mu}M^{\rho\nu} + g^{\mu\rho}M^{\nu\sigma} - g^{\sigma\nu}M^{\rho\mu}$$

up to the possibility of a term on the right-hand side that commutes with $\phi(x)$ and its derivatives.

We now define $J_i \equiv \frac{1}{2}\epsilon_{ijk}M^{jk}$ and $K_i \equiv M^{i0}$,so

$$M^{\mu\nu} = \begin{bmatrix} 0 & -K_1 & -K_2 & -K_3 \\ K_1 & 0 & J_3 & -J_2 \\ K_2 & -J_3 & 0 & J_1 \\ K_3 & J_2 & -J_1 & 0 \end{bmatrix}$$

the communication relationship can be written as

$$\begin{aligned} [J_i, J_j] &= \epsilon_{ijk}J_k \\ [J_i, K_j] &= \epsilon_{ijk}K_k \\ [K_i, K_j] &= -\epsilon_{ijk}J_k \end{aligned}$$

We can use the similar method to derive that

$$[P^\mu, M^{\rho\sigma}] = g^{\mu\sigma}P^\mu - g^{\mu\rho}P^\sigma$$

It can also be written as

$$\begin{aligned} [J_i, H] &= 0 \\ [J_i, P_j] &= \epsilon_{ijk}P_k \\ [K_i, H] &= P_i \\ [K_i, P_j] &= \delta_{ij}H \end{aligned}$$

At last, we define $L_i \equiv \frac{1}{2}\epsilon_{ijk}M_L^{jk}$ and $S_i \equiv \frac{1}{2}\epsilon_{ijk}M_S^{jk}$ we can demonstrate that

$$\begin{aligned} [L_i, S_j] &= 0 \\ [S_i, P_j] &= 0 \\ [L_i, P_j] &= \epsilon_{ijk}P_k \end{aligned}$$

Part III

General relativity

Chapter 4
Elementary Differential Geometry

4.1 Fundamental conception on differential manifolds

Definition 4.1 Manifold

Manifold *Formally, a topological manifold is a second countable Hausdorff space that is locally homeomorphic to Euclidean space.*
Differentiable manifold *In formal terms, a differentiable manifold is a topological manifold with a globally defined differential structure.*
Tangent space *In mathematics, the tangent space of a manifold facilitates the generalization of vectors from affine spaces to general manifolds, since in the latter case one cannot simply subtract two points to obtain a vector pointing from one to the other.*
Cotangent space *Typically, the cotangent space is defined as the dual space of the tangent space at x.*

♡

Definition 4.2 Submanifold

Submanifold
Immersed submanifolds *An immersed submanifold of a manifold M is the image S of an immersion map $f : N \to M$; in general this image will not be a submanifold as a subset, and an immersion map need not even be injective (one-to-one) – it can have self-intersections.*
Injective immersion submanifolds *More narrowly, one can require that the map $f : N \to M$ be an inclusion (one-to-one), in which we call it an injective immersion, and define an immersed submanifold to be the image subset S together with a topology and differential structure such that S is a manifold and the inclusion f is a diffeomorphism: this is just the topology on N, which in general will not agree with the subset topology: in general the subset S is not a submanifold of M, in the subset topology.*
Open submanifolds
Closed submanifolds

♡

Definition 4.3 Embedded Submanifold

An embedded submanifold (also called a regular submanifold), is an immersed submanifold for which the inclusion map is a topological embedding. That is, the submanifold topology on S is the same as the subspace topology. Given any embedding $f : N \to M$ of a manifold N in M the image $f(N)$ naturally has the structure of an embedded submanifold. That is, embedded submanifolds are precisely the images of embeddings.

Proposition 4.1

If an n dimensional injective immersed submanifold N of a m dimensional manifold M is a closed submanifold of an open submanifold of M, then for every point $p \in f(N)$ there exists a chart ($U \subset M$,$\phi : U \to R_n$) containing p such that $\phi(f(N) \cap U)$ is the intersection of a n-dimensional plane with $\phi(U)$.
Closed submanifolds of an open submanifold are equal to embedded submanifolds.

4.2 Multi linear algebra

Definition 4.4 Tensor

Vector space
Dual space
In mathematics, any vector space V has a corresponding dual vector space (or just dual space for short) consisting of all linear functionals on V together with a naturally induced linear structure.
Tensor product

$$V \otimes W = Span\{v \otimes w\} = \mathcal{L}(V^*, W^*; F)$$

$$V^* \otimes W^* = Span\{v^* \otimes w^*\} = \mathcal{L}(V, W; F)$$

$$\mathcal{L}(V, W; Z) = \mathcal{L}(V \otimes W; Z)$$

$$(\phi \otimes \psi) \otimes \xi = \phi \otimes (\psi \otimes \xi)$$

Tensor

$$V_s^r = V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$$

$$x = {x^{i_1 \cdots i_r}}_{k_1 \cdots k_s} e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{*k_1} \otimes \cdots \otimes e^{*k_s}$$

$${(x \otimes y)^{i_1 \cdots i_{r_1+r_2}}}_{k_1 \cdots k_{s_1+s_2}} = {x^{i_1 \cdots i_{r_1}}}_{k_1 \cdots k_{s_1}} \cdot {y^{i_{r_1+1} \cdots i_{r_1+r_2}}}_{k_{s_1+1} \cdots k_{s_1+s_2}}$$

Definition 4.5 Symmtric and antisymmetric tensor

Permutation$(\sigma \in \mathcal{P}(r))$

$$\sigma x(v^{*1}, \cdots, v^{*r}) = x(v^{*\sigma(1)}, \cdots, v^{*\sigma(r)})$$

Symmetric contra-variant tensor

$$\sigma x = x$$

Antisymmetric contra-variant tensor

$$\sigma x = \mathit{sgn}\, \sigma \cdot x$$

Symmetrization operator

$$S_r(x) = \frac{1}{r!} \sum_{\sigma \in \mathcal{P}(x)} \sigma x$$

Antisymmetrization operator

$$A_r(x) = \frac{1}{r!} \sum_{\sigma \in \mathcal{P}(x)} \mathit{sgn} \cdot \sigma x$$

♡

Definition 4.6 Exterior vector space

Exterior vector space

$$\Lambda^r(V) = A_r(T^r(V))$$

$$\Lambda^0(V) = F \quad \Lambda^1(V) = V$$

Wedge product

$$\xi \wedge \eta \equiv \frac{(k+l)!}{k!l!} A_{k+l}(\xi \otimes \eta)$$

Pull-back mapping *$f : V \to W$ is a linear mapping, we define $f^* : \Lambda^r(W^*) \to \Lambda^r(V^*)$ as*

$$f^*\phi(v_1, \cdots, v_r) = \phi(f(v_1), \cdots, f(v_r)).$$

♡

Proposition 4.2 Properties of Wedge product

$$(\xi_1 + \xi_2) \wedge \eta = \xi_1 \wedge \eta + \xi_2 \wedge \eta$$

$$\xi \wedge (\eta_1 + \eta_2) = \xi \wedge \eta_1 + \xi \wedge \eta_2$$

$$\xi \wedge \eta = (-1)^{kl} \eta \wedge \xi$$

$$(\xi \wedge \eta) \wedge \zeta = \xi \wedge (\eta \wedge \zeta) = \frac{(k+l+h)!}{k!l!h!} A_{k+l+h}(\xi \otimes \eta \otimes \zeta)$$

$$f^*(\phi \wedge \psi) = f^*\phi \wedge f^*\psi$$

♠

Proposition 4.3 Properties of exterior space

$$e_{i_1}\wedge\cdots\wedge e_{i_r}(v^{*1},\cdots,v^{*r})=\det\langle e_{i_\alpha},v^{*\beta}\rangle$$
$$e_{i_1}\wedge\cdots\wedge e_{i_r}(e^{*j_1},\cdots,e^{*j_r})=\det\langle e_{i_\alpha},e^{*j_\beta}\rangle=\delta^{j_1\cdots j_r}_{i_1\cdots i_r}$$
$$\Lambda^r(V)=\text{ Span }\{e_{i_1}\wedge\cdots\wedge e_{i_r},1\le i_1<\cdots<i_r\le n\}$$
$$(\Lambda^r(V))^*=\Lambda^r(V^*)$$

4.3 Vector Bundle

Definition 4.7 Fiber bundle

Fiber bundle *In mathematics, and particularly topology, a fiber bundle is a space that is locally a product space, but globally may have a different topological structure. Specifically, the similarity between a space E and a product space B × F is defined using a continuous surjective map $\pi : E\to B$ that in small regions of E behaves just like a projection from corresponding regions of $B\times F$ to B. The map π, called the projection or submersion of the bundle, is regarded as part of the structure of the bundle. The space E is known as the total space of the fiber bundle, B as the base space, and F the fiber.*

Vector Bundle *In mathematics, a vector bundle is a topological construction that makes precise the idea of a family of vector spaces parameterized by another space X (for example X could be a topological space, a manifold, or an algebraic variety): to every point x of the space X we associate (or "attach") a vector space $V(x)$ in such a way that these vector spaces fit together to form another space of the same kind as X (e.g. a topological space, manifold, or algebraic variety), which is then called a vector bundle over X.*

Tangent bundle *In differential geometry, the tangent bundle of a differentiable manifold M is a manifold TM, which assembles all the tangent vectors in M. As a set, it is given by the disjoint union of the tangent spaces of M. That is,*

$$TM=\bigsqcup_{x\in M}T_xM=\bigcup_{x\in M}\{x\}\times T_xM=\bigcup_{x\in M}\{(x,y)|y\in T_xM\}$$

where T_xM denotes the tangent space to M at the point x. So, an element of TM can be thought of as a pair (x,v) , where x is a point in M and v is a tangent vector to M at x. There is a natural projection $\pi : TM\to M$ defined by $\pi(x,v)=x$. This projection maps each tangent space T_xM to the single point x. A section of TM is a vector field on M, and the dual bundle to TM is the cotangent bundle, which is the disjoint union of the cotangent spaces of M.

Cotangent bundle $T^*M=\bigcup_{x\in M}T^*_xM$

Tensor bundle $T^r_sM=\bigcup_{x\in M}T^r_{sx}M$

4.4 Tangent vector field

Theorem 4.1

Let M be a smooth manifold, and let $Y : M \to TM$ be a vector field. If $(U, (X_i))$ is any smooth coordinate chart on M, then Y is smooth on U if and only if its component functions with respect to this chart are smooth.

Theorem 4.2

Let M be a m dimensional smooth manifold and v a smooth tangent vector field on M. $v : C^\infty(M) \to C^\infty$ satisfy that
(1) $\forall f, g \in C^\infty(M), v(f + g) = v(f) + v(g)$;
(2) $\forall f \in C^\infty(M), \alpha \in \mathbf{R}, v(\alpha f) = \alpha \cdot v(f)$;
(3) $\forall f, g \in C^\infty(M), v(f, g) = f \cdot v(g) + g \cdot v(f)$.
If $\alpha : C^\infty(M) \to C^\infty(M)$ satisfy the three conditions above, there exists a unique smooth vector field v on M that $\forall f \in C^\infty(M), v(f) = \alpha(f)$.

Theorem 4.3

$\forall X, Y \in \mathcal{H}(M), [X, Y] = X \circ Y - Y \circ X \in \mathcal{H}(M)$.

Proposition 4.4

(1) $[aX + bY, Z] = a[X, Z] + b[Y, Z]$;$[Z, aX + bY] = a[Z, X] + b[Z, Y]$;
(2) $[X, Y] = -[Y, X]$;
(3) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$;
(4) $[X, Y]|_U = [X|_U, Y|_U] = (X_i \frac{\partial Y^j}{\partial u^i} - Y^i \frac{\partial X^j}{\partial u^i}) \frac{\partial}{\partial u^j}$;
(5) $f_*[X, Y] = [f_*X, f_*Y]$;

Definition 4.8 One parameter differentiable transformation group

Let M be a smooth manifold and $\phi : \mathbf{R} \times M \to M$ a smooth mapping, and $\forall (t, p) \in \mathbf{R} \times M$,denote $\phi_t(p) = \phi(t, p)$. If ϕ satisfy that
(1) $\phi_0 = \mathrm{id} : M \to M$;
(2) $\forall s, t \in \mathbf{R}, \phi_s \circ \phi_t = \phi_{s+t}$;
then ϕ is called a one parameter differentiable transformation group acting on M.
Trajectory of ϕ through p on M:$\gamma_p(t) = \phi(t, p)$.
Vector field induced by ϕ: $X_p(f) = \langle \gamma_p, f \rangle$.

Proposition 4.5

(1) $\gamma_q(t) = \phi(t, \phi(s, p)) = \phi(t + s, p) = \gamma_p(t + s)$;
(2) $(\phi_s)_* X_p = X_{\phi_s(p)}$;
(3) $\psi_* X_p = \tilde{X}_{\psi(p)}$ if X is induced by ϕ and $\tilde{X}$ is induced by $\psi \circ \phi \circ \psi^{-1}$. ψ is a smooth homeomorphism.
(4) $[X, Y] = \lim_{t\to 0} \frac{Y_p - (\phi_t)_* Y_{\phi_{-t}(p)}}{t} = \lim_{t\to 0} \frac{(\phi_{-t})_* Y_{\phi_t(p)} - Y}{t}$ if X is induced by ϕ.

Definition 4.9 Lie derivative

$$\mathcal{L}_X Y \equiv \lim_{t\to 0} \frac{(\phi_{-t})_* Y_{\phi_t(p)} - Y}{t} = [X, Y]$$
$$\mathcal{L}_X f \equiv X(f)$$

Proposition 4.6

$$\mathcal{L}_X(Y + \lambda Z) = \mathcal{L}_X Y + \lambda \mathcal{L}_X Z$$
$$\mathcal{L}_X(f \cdot Y) = \mathcal{L}_X(f) \cdot Y + f \mathcal{L}_X Y$$
$$\mathcal{L}_X([Y, Z]) = [\mathcal{L}_X Y, Z] + [Y, \mathcal{L}_X Z]$$

Theorem 4.4

Let M be a n-dimensional smooth manifold and $X \in \mathcal{H}(M)$. If $p \in M$ and $X_p \neq 0$, $\exists (V, x^i)$ and $p \in V$ that $X|_V = \frac{\partial}{\partial y^1}$

Definition 4.10 Distribution

Distribution Let M be a C^∞ manifold of dimension m, and let $n \leq m$. Suppose that for each $x \in M$, we assign an n-dimensional subspace $\Delta_x \subset T_x(M)$ of the tangent space in such a way that for a neighbourhood $N_x \subset M$ of x there exist n linearly independent smooth vector fields $X_1, \ldots, X_n$ such that for any point $y \in N_x$, $X_1(y), \ldots, X_n(y)$ span Δ_y. We let Δ refer to the collection of all the Δ_x for all $x \in M$ and we then call Δ a distribution of dimension n on M, or sometimes a C^∞ n-plane distribution on M. The set of smooth vector fields $\{X_1, \ldots, X_n\}$ is called a local basis of Δ.

Definition 4.11 Involutive distributions

We say that a distribution Δ on M is involutive if for every point $x \in M$ there exists a local basis $\{X_1, \ldots, X_n\}$ of the distribution in a neighbourhood of x such that for all $1 \leq i, j \leq n$, $[X_i, X_j]$ is in the span of $\{X_1, \ldots, X_n\}$.That is, if $[X_i, X_j]$ is a linear combination of $\{X_1, \ldots, X_n\}$. Normally this is written as $[\Delta, \Delta] \subset \Delta$.
Involutive distributions are the tangent spaces to foliations. Involutive distributions are important in that they satisfy the conditions of the Frobenius theorem, and thus lead to integrable systems. A related idea occurs in Hamiltonian mechanics: two functions f and g on a symplectic manifold are said to be in mutual involution if their Poisson bracket vanishes.

♡

Theorem 4.5 Frobenius Theorem

If distribution Δ on M is involutive, then $\forall p \in M, \exists (V, x^i)$ and $p \in V$ that $\Delta|_V = \text{Span}\{\frac{\partial}{\partial y^1}, \cdots, \frac{\partial}{\partial y^h}\}$.

♣

Definition 4.12 Integrable manifold

Let L^h be a smooth distribution on M. If $\phi : N \to M$ is an injective immersion manifold, and $\forall p \in N$, $\phi_*(T_pN) \subset L^h(\phi(p))$, then (ϕ, N) is called an integrable manifold of L^h. If $\forall q \in M$, there is an integrable manifold of L^h through it, we say that L^h is completely integrable.

♡

Theorem 4.6

Let

$$\tau : \underbrace{A^1(M) \times \cdots \times A^1(M)}_{p} \times \underbrace{\mathcal{H}(M) \times \cdots \times \mathcal{H}(M)}_{q} \to c^\infty(M)$$

be a $p + q$ multi-linear mapping, if $\forall 1 \leq a \leq p, 1 \leq b \leq q$ and $\mu \in C^\infty(M)$,

$$\begin{aligned} & \tau(\alpha^1, \cdots, \mu\alpha^a, \cdots, \alpha^p, v_1, \cdots, v_q) \\ = \; & \tau(\alpha^1, \cdots, \alpha^p, v_1, \cdots, \mu v_b, \cdots, v_q) \\ = \; & \mu \cdot \tau(\alpha^1, \cdots, \alpha^p, v_1, \cdots, v_q) \end{aligned}$$

then the mapping τ define a (p, q) tensor for all $x \in M$ smoothly.

♣

Definition 4.13 Lie derivatives

Let X be a smooth tangent vector field on M and ϕ_t the one parameter differentiable transformation group inducing it. Denote the trajectory of ϕ_t through x by $\gamma_x(t)$. So we have linear isomorphism

$$(\phi_t^{-1})_* = (\phi_{-t})_* : T_{\gamma_x(t)}M \to T_x M$$

$$(\phi_t)^* : T^*_{\gamma_x(t)} \to T_x M$$

So we can induce the linear isomorphism

$$\Phi_t : T^p_q(\gamma_x(t)) \to T^p_q(x)$$

If S and T are smooth tensor fields on M,
(1) for all t which is small enough, $\Phi_t S$ is a smooth tensor field on M which has the same type as S ,and $\lim_{t\to 0} \Phi_t(S(\gamma_p(t))) = S(p), \forall p \in M$.
(2)$\Phi_t(S \otimes T) = \Phi_t S \otimes \Phi_t T$.
(3)$\Phi_t(C^a_b(S)) = C^a_b(\Phi_t(S))$, C^a_b is a tag for contraction.
So,we can define the Lie derivative for smooth tensor field τ on M as

$$\mathcal{L}_X(\tau) = \lim_{t\to 0} \frac{\Phi_t(\tau) - \tau}{t}$$

Proposition 4.7

$$\mathcal{L}_X(\tau_1 + \lambda\tau_2) = \mathcal{L}_X\tau_1 + \lambda\mathcal{L}_X\tau_2$$
$$\mathcal{L}_X(\tau_1 \otimes \tau_2) = \mathcal{L}_X\tau_1 \otimes \tau_2 + \tau_1 \otimes \mathcal{L}_X\tau_2$$
$$C^r_s(\mathcal{L}_X\tau) = \mathcal{L}_X(C^r_s(\tau))$$
$$(\mathcal{L}_X\omega)(Y) = X(\omega(Y)) - \omega([X,Y])$$
$$\mathcal{L}_{[X,Y]} = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X$$
$$\mathcal{L}_{X+Y} = \mathcal{L}_X + \mathcal{L}_Y$$

Proposition 4.8

$$((\mathcal{L}_X\tau)|_U)^{\mu_1,\cdots,\mu_p}_{v_1,\cdots,v_q} = X^\alpha \partial_\alpha \tau^{\mu_1,\cdots,\mu_p}_{v_1,\cdots,v_q} - \sum_{i=1}^{p} \tau^{\mu_1,\cdots,\alpha,\cdots,\mu_p}_{v_1,\cdots,v_q} \partial_\alpha X^{\mu_i} + \sum_{j=1}^{q} \tau^{\mu_1,\cdots,\mu_p}_{v_1,\cdots,\alpha,\cdots,v_q} \partial_{v_j} X^\alpha$$

4.5 Exterior differential

Definition 4.14 Exterior form space

$A(M) = \sum_{r=0}^{m} A^r(M)$

For $\tau \in A^r(M)$,

$$\tau|_U = \frac{1}{r!}\tau_{i_1\cdots i_r} dx^{i_1} \wedge \cdots \wedge dx^{i_r} = \tau_{|i_1\cdots i_r|} dx^{i_1} \wedge \cdots \wedge dx^{i_r}$$

$$\tau_{i_1\cdots i_r} = \tau(\frac{\partial}{\partial x^{i_1}}, \cdots, \frac{\partial}{\partial x^{i_r}})$$

$$\begin{aligned} \tau(v_1, \cdots, v_r)|_U &= \tau_{|i_1\cdots i_r|} dx^{i_1} \wedge \cdots \wedge dx^{i_r}(v_1, \cdots, v_r) \\ &= \tau_{|i_1\cdots i_r|} \begin{vmatrix} v_1^{i_1} & \cdots & v_r^{i_1} \\ \vdots & & \vdots \\ v_1^{i_r} & \cdots & v_r^{i_r} \end{vmatrix} \end{aligned}$$

It is a r multi-linear mapping, and for every variable, it is $C^\infty(M)$ linear.

Proposition 4.9 Pullback mapping

$f : M \to N \Rightarrow f_* : T_pM \to T_{f(p)}N \Rightarrow f^* : \wedge^r(T^*_{f(p)}N) \to \wedge^r(T^*_pM)$

$$f^*\phi(v_1, \cdots, v_r) = \phi(f_*v_1, \cdots, f_*v_r)$$

$$f^*\phi|_U = \frac{1}{r!}(\phi_{\alpha_1\cdots\alpha_r} \circ f) \cdot \frac{\partial f^{\alpha_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{\alpha_r}}{\partial x^{i_r}} dx^{i_1} \wedge \cdots \wedge dx^{i_r}$$

$$f^*(\phi \wedge \psi) = f^*\phi \wedge f^*\psi$$

Definition 4.15 Exterior differential

Let M be a m-dimensional smooth manifold. Then $\exists$ a unique mapping $d : A(M) \to A(M)$ satisfy that

(1) $d(A^r(M)) \subset A^{r+1}(M)$

(2) $\forall \omega_1, \omega_2 \in A(M), d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$

(3) if $\omega_1 \in A^r(M)$, then $d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^r \omega_1 \wedge d\omega_2$

(4) $f \in A^0(M)$, df is just the differential of f

(5) $\forall f \in A^0(M)$, $d(df) = 0$

d is called exterior differential.

Theorem 4.7

$\forall \omega \in A^1(M), X, Y \in \mathcal{H}(M)$,

$$d\omega(X,Y) = X\langle Y,\omega\rangle - Y\langle X,\omega\rangle - \langle [X,Y],\omega\rangle$$

$\forall \omega \in A^r(M), X_1, \cdots, X_{r+1} \in \mathcal{H}(M)$,

$$\begin{aligned} d\omega(X_1,\cdots,X_{r+1}) &= \sum_{i=1}^{r+1}(-1)^{i+1}X_i(\langle X_1\wedge\cdots\wedge\hat{X}_i\wedge\cdots\wedge X_{r+1},\omega\rangle) \\ &+ \sum_{1\le i<j\le r+1}(-1)^{i+j}\langle [X_i,X_j]\wedge\cdots\wedge\hat{X}_i\wedge\cdots\wedge\hat{X}_j\wedge\cdots X_{r+1},\omega\rangle \end{aligned}$$

♣

Theorem 4.8

$$f^*(d\omega) = d(f^*\omega)$$

♣

Lemma 1 Poincare Lemma

1. $d^2 = 0$
2. Let $U = B_0(r)$ be a spherical neighbourhood with center origin O and radius r in R^n. $\forall \omega \in A^r(U)$ and $d\omega = 0$, $\exists \tau \in A^{r-1}(U)$,satisfy that $\omega = d\tau$.

♣

Definition 4.16 Pfaff euqations

Let $\omega^\alpha(1 \le \alpha \le r) \in A^1(U)$ and U is an open set of m-dimensional smooth manifold M. Differential equation set $\omega^\alpha = 0$ is called Pfaff equations.

♡

Definition 4.17 Integral manifold of Pfaff equations

If there is an injective immersion submanifold $\phi : N \to U$ satisfying that $\phi^*\omega^\alpha = 0$,(ϕ, N) is called an integral manifold of Pfaff eqation set.

♡

Proposition 4.10 Partial differential equations and Pfaff equations

There is a set of first order partial differential equations

$$\frac{\partial y^\alpha}{\partial x^i} = f_i^\alpha(x^1, \cdots, x^m, y^1, \cdots, y^n) \quad (1 \le i \le m, 1 \le \alpha \le n)$$

$f_i^\alpha(x, y)$ is a smooth function on the open set $U \times V \subset R^m \times R^n$. The equations sets can be written as Pfaff equations on $U \times V$

$$\omega^\alpha \equiv dy^\alpha - f_i^\alpha(x, y)dx^i = 0$$

If the partial differential equations have solution

$$y^\alpha = g^\alpha(x^1, \ldots, x^m)$$

then the submanifold $\phi : U \to U \times V$,

$$\phi(x^1, \ldots, x^m) = (x^1, \ldots, x^m, g^1(x), \ldots, g^n(x))$$

is an integral manifold of the Pfaff equations , i.e. $\phi^*\omega^\alpha = 0$

Proposition 4.11 Distribution and Pfaff equations

Pfaff equations $\omega^\alpha = 0$ on open set $V \in M$ with rank r is equivalent to a $h = m - r$ dimensional smooth distribution locally.

$$\Delta^h(p) = \{v \in T_pM : \omega^\alpha(v) = 0, 1 \le \alpha \le r\}$$

If $\phi : N \to V$ is an integral manifold of ω^α,$\forall X \in T_pN$,$\omega^\alpha(\phi_* X) = \phi^*\omega_\alpha(X) = 0$. So $\phi_* X \in \Delta^h(p)$, and so $\phi : N \to V$ is an integral manifold of Δ^h.

Definition 4.18 Completely integrable

Suppose ω^α is a set of r linearly independent 1 forms defined on an open set $U \subset M$. If $\forall p \in U$, Paffa equations

$$\omega^\alpha = 0 \quad (1 \le \alpha \le r)$$

has an $h = \dim M - r$ dimensional integral manifold $\phi : N \to V$ such that $p \in V$, Paffa equations are called completely integrable.

Definition 4.19 Frobenius condition

Frobenius condition for Pfaff equations $\omega^\alpha = 0 (1 \leq \alpha \leq r)$ *is that*

$$d\omega^\alpha \equiv 0(\mathrm{mod}(\omega^1, \ldots, \omega^r))$$

♡

Theorem 4.9 Frobenius theorem

Pfaff equations satisfying Frobenius condition is completely integrable.

♣

Definition 4.20 Orientation of manifold

Let $\alpha : [0,1] \to M$ *be a path on* M. $\forall t \in [0,1]$, *assign an orientation for* $T_{\alpha(t)}M$, *denoted by* μ_t. *If for* $t_0 \in [0,1]$, *there is a local coordinate* $(U; x_i)$ *of* $\alpha(t_0)$ *and a neighbourhood* $[t_0 - \delta_1, t_0 + \delta_2]$ *of* t_0 *that*

$$\alpha([t_0 - \delta_1, t_0 + \delta_2]) \subset U$$

and

$$\left\{ \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m} \right\} |_{\alpha(t)} \in \mu_t, \forall t \in [t_0 - \delta_1, t_0 + \delta_2],$$

μ *is called a continuous topological orientation of* α

♡

Definition 4.21 The propagation of orientation

Let $p, q \in M$ *and* $\alpha : [0,1] \to M$ *a path connecting* p, q. *Assign an orientation* λ *of* T_pM. *If there is a continuous topological orientation of* α μ *satisfying that* $\mu_0 = \lambda$, *then orientation* μ_1 *of* T_qM *is called the propagation of orientation* λ *along* α. *The orientation of* μ_1 *is unique.*

♡

Definition 4.22 Orientable manifold

Let M *be a* m *dimensional smooth manifold. If there is an atlases* $(\mathcal{A}, = \{(U_\alpha, \phi_\alpha)\})$, *making that if* $U_\alpha \cap U_\beta \neq \emptyset$, *the Jacobian of*

$$\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$$

is positive. Then M *is called orientable manifold.*

♡

Theorem 4.10

Let M be a orientable connected manifold. $\forall p \in M$,assign an orientation λ for T_pM, then for all point $q \in M$, the propagation of λ along an arbitrary path define a unique orientation μ for T_qM.

Definition 4.23 Manifold with boundary

A topological manifold with boundary is a Hausdorff space in which every point has a neighbourhood homeomorphic to an open subset of Euclidean half-space (for a fixed n):

$$\mathbb{R}^n_+ = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n \geq 0\}$$

Definition 4.24 Boundary and interior

Let M be a manifold with boundary. The interior of M, denoted $Int\ M$, is the set of points in M which have neighbourhoods homeomorphic to an open subset of $\mathbb{R}^n$. The boundary of M, denoted ∂M, is the complement of $Int\ M$ in M. The boundary points can be characterized as those points which land on the boundary hyperplane ($x^n = 0$) of $\mathbb{R}^n_+$ under some coordinate chart.
If M is a manifold with boundary of dimension n, then $Int\ M$ is a manifold (without boundary) of dimension n and ∂M is a manifold (without boundary) of dimension $n-1$.

Theorem 4.11

Let M be a smooth manifold with boundary and $\partial M \neq \emptyset$. The differential structure of ∂M can be deduced from the M, making ∂M a $m-1$ dimensional smooth manifold and the inclusion map $i : \partial M \to M$ is embedding map. If M is orientable, then ∂M is also orientable.

Definition 4.25 Induced orientation

Let M be an orientable m dimensional smooth manifold with boundary and $\partial M \neq \emptyset$. $\mathcal{A}$ is the orientation of M. For local coordinates $(U; x^i) \in \mathcal{A}$, when

$$\tilde{U} = U \cap \partial M = \{(x^1, \ldots, x^m) \in U : x^m = 0\} \neq \emptyset$$

assign a local coordinate system $((-1)^m \cdot x^1, x^2, \ldots, x^{m-1})$ on $\tilde{U}$. The orientation defined by this local coordinate system is called induced orientation of ∂M.

Definition 4.26 Support set

Let M be a m dimensional orientable smooth manifold. $\omega \in A^r(M)$, the support set of ω can be defined as

$$Supp\,\omega = \overline{\{p \in M : \omega(p) \neq 0\}}$$

All the r-form with compact support set is denoted as $A_0^r(M)$.

Definition 4.27 Partition of unity

Let Σ be an open cover of M. Then there is a family of smooth function g_α on M that

1. $\forall \alpha, 0 \leq g_\alpha \leq 1, \operatorname{supp} g_\alpha$ *is compact and there is an open set* $W_i \in \Sigma$ *that* $\operatorname{supp} g_\alpha \subset W_i$.
2. $\forall p \in M$, *it has a neighbourhood* U *which intersect finite* $\operatorname{supp} g_\alpha$.
3. $\sum_\alpha g_\alpha = 1$

Definition 4.28 Integral of differential form with compact support

$$\phi = (\sum_\alpha g_\alpha) \cdot \phi = \sum_\alpha (g_\alpha \cdot \phi)$$

$$\int_M g_\alpha \cdot \phi = \int_{W_i} g_\alpha \cdot \phi = \int_{W_i} f(u^1, \cdots, u^m) du^1 \wedge \cdots \wedge du^m = \int_{W_i} f(u^1, \cdots, u^m) du^1 \cdots du^m$$

$$\int_M \phi = \sum_\alpha \int_M g_\alpha \cdot \phi$$

Theorem 4.12 Stokes Theorem

Let M be an orientable m dimensional smooth manifold with boundary and $\omega \in A_0^{m-1}(M)$,then

$$\int_M d\omega = \int_{\partial M} i^* \omega$$

Here, ∂M has an orientation induced by M and i is embedding mapping.

4.6 Connection

Definition 4.29 Connection

Let M be a smooth manifold and E a q dimensional real vector bundle on M. $\Gamma(E)$ is the set of all smooth sections of E on M. The connection on E is a mapping:

$$D : \Gamma(E) \to \Gamma(T^*(M) \otimes E)$$

it satisfies that

1. $\forall s_1, s_2 \in \Gamma(E)$, $D(s_1 + s_2) = Ds_1 + Ds_2$
2. $\forall s \in \Gamma(E)$ *and* $\alpha \in C^\infty(M)$, $D(\alpha s) = d\alpha \otimes s + \alpha Ds$

If X is a smooth tangent vector field on M, $s \in \Gamma(E)$, then $D_X s = \langle X, Ds \rangle$, called absolute derivative of s along X.

♡

Proposition 4.12

Local representation of connection is:

$$Ds_\alpha = \sum_{1 \le i \le m, 1 \le \beta \le q} \Gamma^\beta_{\alpha i} du^i \otimes s_\beta$$

$$\omega^\beta_\alpha = \sum_{1 \le i \le m} \Gamma^\beta_{\alpha i} du^i$$

$$Ds_\alpha = \sum_{\beta=1}^{q} \omega^\beta_\alpha \otimes s_\beta$$

$$DS = \omega \otimes S$$

Transformation law of connection is:

$$S' = A \cdot S$$

$$\begin{aligned} DS' &= dA \otimes S + A \cdot DS \\ &= (dA + A \cdot \omega) \otimes S \\ &= (dA \cdot A^{-1} + A \cdot \omega \cdot A^{-1}) \otimes S' \end{aligned}$$

$$\omega' = dA \cdot A^{-1} + A \cdot \omega \cdot A^{-1}$$

♠

Theorem 4.13

For an arbitrary vector bundle, connection always exists.

Theorem 4.14

Let D be a connection of vector bundle E. $\forall p \in M$, there exists a local frame field S on the neighbourhood of p that $\omega(p) = 0$.

Definition 4.30 Curvature matrix

$$\Omega \equiv d\omega - \omega \wedge \omega$$

Proposition 4.13 Transformation law of curvature matrix

$$\Omega' = A \cdot \Omega \cdot A^{-1}$$

Definition 4.31 Curvature operator

$$s = \sum_{\alpha=1}^{q} \lambda^\alpha s_{\alpha|p}$$

$$R(X,Y)s = \sum_{\alpha,\beta=1}^{q} \lambda^\alpha \Omega_\alpha^\beta(X,Y) s_{\beta|p}$$

$$(R(X,Y)s)(p) = R(X_p, Y_p)s_p$$

Proposition 4.14

$$R(X,Y) = D_X D_Y - D_Y D_X - D_{[X,Y]}$$

Theorem 4.15 Bianchi equation

$$d\Omega = \omega \wedge \Omega - \Omega \wedge \omega$$

Definition 4.32 Induced connection

$$d\langle s, s^*\rangle = \langle Ds, s^*\rangle + \langle s, Ds^*\rangle$$

$$\langle s_\alpha, s^{*\beta}\rangle = \delta_\alpha^\beta \Rightarrow Ds^{*\beta} = -\sum_{\alpha=1}^{q} \omega_\alpha^\beta \otimes s^{*\alpha}$$

$$D(s_1 \oplus s_2) \equiv Ds_1 \oplus Ds_2$$

$$D(s_1 \otimes s_2) \equiv Ds_1 \otimes s_2 + s_1 \otimes Ds_2$$

♡

Definition 4.33 Affine connection

$$D\frac{\partial}{\partial u^i} \equiv \omega_i^j \otimes \frac{\partial}{\partial u^j} \equiv \Gamma_{ik}^j du^k \otimes \frac{\partial}{\partial u^j}$$

♡

Proposition 4.15

$$\Gamma'^j_{ik} = \Gamma^q_{pr}\frac{\partial w^j}{\partial u^q}\frac{\partial u^p}{\partial w^i}\frac{\partial u^r}{\partial w^k} + \frac{\partial^2 u^p}{\partial w^i \partial w^k}\frac{\partial w^j}{\partial u^p}$$

$$DX = (dx^i + x^j\omega_j^i) \otimes \frac{\partial}{\partial u^i} = (x^i_{,j} + x^k\Gamma^i_{kj})du^j \otimes \frac{\partial}{\partial u^i} = x^i_{;j}du^j \otimes \frac{\partial}{\partial u^i}$$

$$D\alpha = (d\alpha_i - \alpha_j\omega_i^j) \otimes du^i = (\alpha_{i,j} - \alpha_k\Gamma^k_{ij})du^j \otimes du^i = \alpha_{i;j}du^j \otimes du^i$$

♠

Definition 4.34 Geodesic equation

$$\frac{D(\frac{du^i(t)}{dt}\frac{\partial}{\partial u^i})}{dt} = 0$$

$$\frac{d^2u^i}{dt^2} + \Gamma^i_{jk}\frac{du^j}{dt}\frac{du^k}{dt} = 0$$

♡

Definition 4.35 Curvature tensor

$$\Omega_i^j = \frac{1}{2}R^j_{ikl}du^k \wedge du^l$$

$$R \equiv R^j_{ikl}du^i \otimes \frac{\partial}{\partial u^j} \otimes du^k \otimes du^l$$

♡

Proposition 4.16

$$R^j_{ikl} = \frac{\partial \Gamma^j_{il}}{\partial u^k} - \frac{\partial \Gamma^j_{ik}}{\partial u^l} + \Gamma^h_{il}\Gamma^j_{hk} - \Gamma^h_{ik}\Gamma^j_{hl}$$

$$R'^j_{ikl} = R^q_{prs}\frac{\partial w^j}{\partial u^q}\frac{\partial u^p}{\partial w^i}\frac{\partial u^r}{\partial w^k}\frac{\partial u^s}{\partial w^l}$$

$$R(X, \alpha_Y, Z, W) = \langle \alpha_Y, R(Z, W)X\rangle$$

$$R^j_{ikl} = \langle R(\frac{\partial}{\partial u^k}, \frac{\partial}{\partial u^l})\frac{\partial}{\partial u^i}, du^j\rangle$$

Definition 4.36 Torsion tensor

$$T^j_{ik} = \Gamma^j_{ki} - \Gamma^j_{ik}$$

$$T = T^j_{ik}\frac{\partial}{\partial u^j} \otimes du^i \otimes du^k$$

Proposition 4.17

$$T(X, Y) = T^k_{ij}X^iY^j\frac{\partial}{\partial u^k}$$

$$T(X, Y) = D_XY - D_YX - [X, Y]$$

Theorem 4.16

Let D be an affine connection without torsion on M. $\forall p \in M$, there is a local coordinate system that $\Gamma^j_{ik}(p)$ vanishes.

Theorem 4.17

Let D be an affine connection without torsion on M. Then we have Bianchi equation

$$R^j_{ikl;h} + R^j_{ihk;l} + R^j_{ilh;k} = 0$$

4.7 Riemann manifold

Definition 4.37 Riemann manifold

Let M be a smooth manifold equipped with a smooth non-degenerate symmetric second order covariant tensor field G, then M is called general Riemann manifold and G is called the metric tensor of M.
If G is positive definite, then M is called Riemann manifold.

Theorem 4.18

There must be a Riemann metric on m dimensional manifold M.

Definition 4.38 Index lifting

$$f : T_p(M) \to T_p^*(M) \quad \alpha_X(Y) \equiv G(X, Y)$$

Definition 4.39 Adapted connection

Let (M, G) be a general Riemann manifold and D a connection on M. If $DG = 0$,then D is called adapted connection on M.

Proposition 4.18 Christoffel-Levi-Civita connection

Let M be a general Riemann manifold, then there is a unique adapted connection without torsion on M, called Christoffel-Levi-Civita connection.
As $\omega_i^j = \Gamma_{ik}^j du^k, dg_{ij} = g_{ik}\omega_j^k + g_{kj}\omega_i^k$. If we denote that $\omega_{ij} = \omega_i^j g_{jk}$ and $\Gamma_{ijk} = \Gamma_{ik}^l g_{lj}$, we have $\omega_{ij} = \Gamma_{ijk} du^k$ and $dg_{ij} = \omega_{ji} + \omega_{ij}$. At last, we have

$$\Gamma_{ij}^k = \frac{1}{2} g^{kl} \left(\frac{\partial g_{il}}{\partial u^j} + \frac{\partial g_{jl}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^l}\right)$$

Proposition 4.19 Curvature tensor

If we denote that $\Omega_{ij} = \Omega_i^k g_{kj}$ and $R_{ijkl} = R_{ikl}^h g_{hj}$, we will have that

$$\Omega_{ij} + \Omega_{ji} = 0, \ \ \Omega_{ij} = d\omega_{ij} + \omega_i^l \wedge \omega_{jl}, \ \ \Omega_{ij} = \frac{1}{2} R_{ijkl} du^k \wedge du^l$$

The properties of curvature tensor:

$$R_{ijkl} = -R_{jikl} = -R_{ijlk}$$

$$R_{ijkl} + R_{iklj} + R_{iljk} = 0$$

$$R_{ijkl} = R_{klij}$$

Definition 4.40 Normal coordinates

In differential geometry, normal coordinates at a point p in a differentiable manifold equipped with a symmetric affine connection are a local coordinate system in a neighbourhood of p obtained by applying the exponential map to the tangent space at p. In a normal coordinate system, the Christoffel symbols of the connection vanish at the point p, thus often simplifying local calculations. In normal coordinates associated to the Levi-Civita connection of a Riemann manifold, one can additionally arrange that the metric tensor is the Kronecker delta at the point p, and that the first partial derivatives of the metric at p vanish.

The properties of normal coordinates often simplify computations. In the following, assume that U is a normal neighbourhood centred at p in M and (x_i) are normal coordinates on U.

Let V be some vector from T_pM with components V^i in local coordinates, and γ_V be the geodesic with starting point p and velocity vector V, then γ_V is represented in normal coordinates by $\gamma_V(t) = (tV^1, ..., tV^n)$ as long as it is in U.

The coordinates of p are $(0, \cdots, 0)$

In Riemann normal coordinates at p the components of the Riemann metric g simplify to δ_{ij}.

The Christoffel symbols vanish at p. In the Riemann case, so do the first partial derivatives of g_{ij}.

Theorem 4.19

Let M be a differentiable manifold equipped with a symmetric affine connection. $\forall x_0 \in M$, there is a neighbourhood W that for every point in W, there is a neighbourhood equipped with a normal coordinate system which contains W.

Theorem 4.20

Let M be a Riemann manifold. $\forall O \in M$,there is a neighbourhood with normal coordinates W that:
(1) For every point in W, there is a neighbourhood equipped with a normal coordinates which contains W.
(2) The geodesic connecting O and $p \in W$ is the only shortest path connecting these two points in W.

Theorem 4.21

Let U be the neighbourhood with normal coordinates of O. $\exists \epsilon > 0, \forall \delta \in (0, \epsilon)$,the surface

$$\Sigma_\delta = p \in U | \sum_{i=1}^{m} (u^i(p))^2 = \delta^2$$

has following properties:
(1) $\forall p \in \Sigma_\delta$, there is a unique shortest geodesic connecting p and O in U.
(2) For all geodesics tangent to Σ_δ, there is a neighbourhood of the cut point in which the geodesics lies outside of Σ_δ

Theorem 4.22

Let M be a Riemann manifold and $\forall p \in M$, there is a η-spherical neighbourhood W that for arbitrary two points in W, there is a unique geodesic connecting these two points.

Definition 4.41 Cross section curvature

$$R(X, Y, Z, W) \equiv R_{ijkl} X^i Y^j Z^k W^l$$

$$R(X, Y, Z, W) = (R(Z, W)X) \cdot Y$$

$$G(X, Y, Z, W) \equiv G(X, Z)G(Y, W) - G(X, W)G(Y, Z)$$

Let E be a two dimensional subspace of $T_p(M)$ and X, Y two linearly independent tangent vector of E,then

$$K(E) = -\frac{R(X, Y, X, Y)}{G(X, Y, X, Y)}$$

is a function of E, which is independent of the choice of X, Y,called cross section curvature.

Theorem 4.23

Let M be a Riemann space. The curvature tensor of $p \in M$ is uniquely determined by the cross section curvature of all the two dimensional subspace of $T_p(M)$.

Definition 4.42 Constant curvature Riemann manifold

Let M be a Riemann manifold. If all of $K(E)$ on p is constant, then M is called isotropic on p.

If M is isotropic every where and $K(p)$ is constant over M, then M is called constant curvature Riemann manifold.

Theorem 4.24 F.Schur theorem

Let M be a m-dimensional connected Riemann manifold that is isotropic every where. If $m \geq 3$, then M is constant curvature Riemann manifold.

Chapter 5
A Geometrical Description of Newton Theory

5.1 Introduction

We choose Euclidean coordinates for our absolute space and an absolute time t, than the equation of motion can be written as

$$\frac{d^2t}{d\lambda^2} = 0$$

$$\frac{d^2x^i}{d\lambda^2} + \frac{\partial\Phi}{\partial x^i}(\frac{d\lambda}{dt})^2 = 0$$

It is convenient to define that $\Gamma^i_{00} = \frac{\partial\Phi}{\partial x^i}$,and all other $\Gamma^\alpha_{\beta\gamma}$ vanish.Then we can write the equation of motion as

$$\frac{d^2x^\alpha}{d\lambda^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\lambda}\frac{dx^\gamma}{d\lambda} = 0$$

Next, we can get the Riemann tensor given the connection above

$$R^i_{0j0} = -R^i_{00j} = \frac{\partial\Phi}{\partial x^i\partial x^j},$$

and all other terms vanish. It is straight forward to derive the expression of Ricci tensor,

$$R_{00} = \Phi_{ii} = \nabla^2\Phi,$$

and all other terms vanish. So, newton gravity law can be written as

$$R_{00} = 4\pi\rho$$

5.2 Geometry structure of Newtonian Space-time

Stratification of space-time

Regard absolute time t as a scalar field defined once and for all in Newtonian space-time $t = t(\mathcal{P})$.The layers of space-time are the slices of constant t-the "space slices"-each of which has an identical geometric structure: the old "absolute space."

Flat Euclidean space

A given space slice is endowed with basis vectors $\mathbf{e}_i = \frac{\partial}{\partial x^i}$; and this basis has vanishing connection coefficients,$\Gamma^i_{jk} = 0$. Consequently,the geometry of each space slice is completely flat. Absolute space is Euclidean in its geometry.Each space slice is endowed with a three-dimensional metric, and its Galilean coordinate basis is orthonormal,$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$.

Curvature of space-time

Parallel transport a vector around a closed curve lying entirely in a space slice; it will return to its starting point unchanged. But transport it forward in time by Δt, northerly in space by Δx_k, back in time by $-\Delta t$, and southerly by $-\Delta x_k$ to its starting point; it will return changed by

$$\delta \mathbf{A} = -\mathcal{R}(\Delta t \frac{\partial}{\partial t}, \Delta x_k \frac{\partial}{\partial x_k})\mathbf{A}$$

Geodesics of a space slice (Euclidean straight lines) that are initially parallel remain always parallel. But geodesics of space-time (trajectories of freely falling particles) initially parallel get pried apart or pushed together by space-time curvature,

$$\nabla_{\mathbf{u}} \nabla_{\mathbf{u}} \mathbf{n} + \mathcal{R}(\mathbf{n}, \mathbf{u})\mathbf{u} = 0$$

5.3 Geometry formulation of Newtonian gravity

1. There exists a function t called "universal time", and a symmetric covariant derivative ∇.

2. The 1-form $\mathbf{d}t$ is covariant constant, i.e.,

$$\nabla_{\mathbf{u}} \mathbf{d}t = 0 \text{ for all } \mathbf{u}.$$

Note: *if $\mathbf{w}$ is a spatial vector field, then $\nabla_{\mathbf{u}} \mathbf{w}$ is also spatial for every $\mathbf{u}$.*

3. Spatial vectors are unchanged by parallel transport around infinitesimal closed curves; i.e.,

$$\mathcal{R}(\mathbf{n}, \mathbf{u})\mathbf{w} = 0 \text{ if } \mathbf{w} \text{ is spatial, for every } \mathbf{u} \text{ and } \mathbf{n}.$$

4. All vectors are unchanged by parallel transport around infinitesimal, spatial, closed curves; i.e.,

$$\mathcal{R}(\mathbf{v}, \mathbf{w}) = 0 \text{ for every spatial } \mathbf{v} \text{ and } \mathbf{w}.$$

5. The Ricci curvature tensor has the form

$$\mathbf{Ricci} = 4\pi\rho \mathbf{d}t \otimes \mathbf{d}t$$

where ρ is the density of mass.

6. There exists a metric $\cdot$ defined on spatial vectors only, which is compatible with the covariant derivative in this sense: for any spatial $\mathbf{w}$ and $\mathbf{v}$, and for any $\mathbf{u}$ whatsoever,

$$\nabla_{\mathbf{u}}(\mathbf{w} \cdot \mathbf{v}) = (\nabla_{\mathbf{u}} \mathbf{w}) \cdot \mathbf{v} + \mathbf{w} \cdot (\nabla_{\mathbf{u}} \mathbf{v}).$$

Note: *Axioms (1), (2), and (3) guarantee that such a spatial metric can exist.*

7. The Jacobi curvature operator $\mathcal{J}(\mathbf{u},\mathbf{v})$, defined for any vectors $\mathbf{u},\mathbf{n},\mathbf{p}$ by
$$\mathcal{J}(\mathbf{u},\mathbf{n})\mathbf{p}=\frac{1}{2}[\mathcal{R}(\mathbf{p},\mathbf{n})\mathbf{u}+\mathcal{R}(\mathbf{p},\mathbf{u})\mathbf{n}]$$
is "self-ad-joint" when operating on spatial vectors,i.e.,
$$\mathbf{v}\cdot[\mathcal{R}(\mathbf{u},\mathbf{n})\mathbf{w}]=\mathbf{w}\cdot[\mathcal{R}(\mathbf{u},\mathbf{n})\mathbf{v}]\text{ for all spactial }\mathbf{v},\mathbf{w};\ \text{ and for any }\mathbf{u},\mathbf{n}.$$

8. "Ideal rods" measure the lengths that are calculated with the spatial metric; "ideal clocks" measure universal time t (or some multiple thereof); and "freely falling particles" move along geodesics of ∇.

5.4 Standard formulation of Newtonian gravity

1. There exist a universal time t, a set of Cartesian space coordinates x_i (called "Galilean coordinates"), and a Newtonian gravitational potential Φ.
2. The density of mass ρ generates the Newtonian potential by Poisson's equation,
$$\nabla^2\Phi=\frac{\partial\Phi}{\partial x^i\partial x^i}=4\pi\rho.$$
3. The equation of motion for a freely falling particle is
$$\frac{d^2x^i}{dt^2}+\frac{\partial\Phi}{\partial x^i}=0.$$
4. "Ideal rods" measure the Galilean coordinate lengths; "ideal clocks" measure universal time.

5.5 Galilean coordinate system

The features of Galilean coordinate systems are
$$x^0(\mathcal{P})=t(\mathcal{P})$$
$$\frac{\partial}{\partial x^i}\cdot\frac{\partial}{\partial x^j}=\delta_{ij}$$
$$\Gamma^j_{00}=\Phi_{,j}\text{ for some scalar field }\Phi,\ \text{ and all other }\Gamma^\alpha_{\beta\gamma}\text{vanish.}$$

Consider following coordinate transformation:
(1)$x^{0'}=x^0=t$,both time coordinates must be universal time.
(2)at fixed t,both sets of space coordinates must be Euclidean, so they must be related by a rotation and a translation:
$$\bar{x}^{i'}(t)=A_{i'j}(t)x^j+\bar{a}^{i'}(t)$$

We can get

$$\bar{\Gamma}^{i'}_{0j'} = \bar{\Gamma}^{i'}_{j'0} = A_{i'l}\dot{A}_{j'l}$$

$$\bar{\Gamma}^{i'}_{00} = \Phi_{,i'} + A_{i'j}(\ddot{A}_{l'j}\bar{x}^{l'} - \ddot{a}^{j}), \text{ here, } a^{j} = \bar{a}^{l'}A_{l'j}$$

and all other terms vanish.So, new coordinates have the standard Galilean form if and only if

$$\dot{A}_{i'j} = 0, \ \ \Phi' = \Phi - \ddot{\bar{a}}^{i'}x^{i'} + C$$

Were all the matter in the universe concentrated in a finite region of space and surrounded by emptiness (”island universe”), then one could impose the global boundary condition $\Phi \to 0$ as $r \equiv (x^i x^i)^{\frac{1}{2}} \to \infty$. This would single out a subclass of Galilean coordinates (”absolute” Galilean coordinates), with a unique, common Newtonian potential. The transformation from one absolute Galilean coordinate system to any other is called Galilean transformation.

5.6 Coordinate transformation in space

We now consider a coordinate transformation of Galilean coordinate system purely in space without any terms related with time. That means that $\bar{x}^{i'} = y^{i'}(x^i)$ and $t' = t$. We can calculate the connection term in the new coordinate system.

$$\bar{\Gamma}^{i'}_{00} = \Gamma^{i}_{00}\frac{\partial y^{i'}}{\partial x^{i}}$$

$$\bar{\Gamma}^{i'}_{j'k'} = \frac{\partial^2 x^m}{\partial y^{i'}\partial y^{k'}}\frac{\partial y^{i'}}{\partial x^m}$$

The equation of motion of free fall body is that

$$\frac{d^2 t'}{d\lambda^2} = 0$$

$$\frac{d^2\bar{x}^{i'}}{d\lambda^2} + \bar{\Gamma}^{i'}_{j'k'}\frac{d\bar{x}^{j'}}{d\lambda}\frac{d\bar{x}^{k'}}{d\lambda} + \bar{\Gamma}^{i'}_{00}\frac{dt'}{d\lambda}\frac{dt'}{d\lambda} = 0$$

We can write it compactly as

$$\frac{d^2\bar{x}^{i'}}{dt^2} + \bar{\Gamma}^{i'}_{j'k'}\frac{d\bar{x}^{j'}}{dt}\frac{d\bar{x}^{k'}}{dt} + \bar{\Gamma}^{i'}_{00} = 0$$

We can demonstrate that

$$\bar{\Gamma}^{i'}_{j'k'} = \frac{1}{2}\bar{g}^{i'p'}(\partial_{k'}\bar{g}_{j'p'} + \partial_{j'}\bar{g}_{k'p'} - \partial_{p'}\bar{g}_{j'k'})$$

and

$$\bar{\Gamma}^{i'}_{00} = \bar{g}^{i'j'}\partial_{j'}\Phi$$

Here,$\bar{g}$ is the metric of the space in new coordinate system.

Chapter 6
Geometry of Space-time

6.1 More on the manifold of space-time

6.1.1 Hodge dual

Definition 6.1 Hodge dual operator

The Hodge star operator on a vector space V with a non-degenerate symmetric bilinear form (herein referred to as the inner product) is a linear operator on the exterior algebra of V, mapping k-vectors to $(n-k)$-vectors where $n = \dim V$, for $0 \leq k \leq n$. It has the following property, which defines it completely: given two k-vectors α, β,

$$\alpha \wedge (\star\beta) = \langle \alpha, \beta \rangle \omega$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on k-vectors and ω is the preferred unit n-vector. The inner product $\langle \cdot, \cdot \rangle$ on k-vectors is extended from that on V by requiring that

$$\langle \alpha, \beta \rangle = \det \left[\langle \alpha_i, \beta_j \rangle \right]$$

for any decomposable k-vectors $\alpha = \alpha_1 \wedge \cdots \wedge \alpha_k$ and $\beta = \beta_1 \wedge \cdots \wedge \beta_k$. The unit n-vector ω is unique up to a sign. The preferred choice of ω defines an orientation on V.

Given an orthonormal basis $(e_1, \cdots, e_n)$ ordered such that $\omega = e_1 \wedge \cdots \wedge e_n$, we see that

$$\star(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}) = e_{i_{k+1}} \wedge e_{i_{k+2}} \wedge \cdots \wedge e_{i_n}$$

where $(i_1, i_2, \cdots, i_n)$ is an even permutation of $\{1, 2, \cdots, n\}$.Of these $\frac{n!}{2}$, only $\binom{n}{k}$ are independent. The first one in the usual lexicographical order reads

$$\star(e_1 \wedge e_2 \wedge \cdots \wedge e_k) = e_{k+1} \wedge e_{k+2} \wedge \cdots \wedge e_n$$

6.1.2 Levi-Civita tensor

Definition 6.2 Levi-Civita tensor

$$\epsilon_{i_1,\cdots,i_n} \equiv |g|^{\frac{1}{2}}\tilde{\epsilon}_{i_1,\cdots,i_n},$$

where ϵ is Levi-Civita symbol.

Proposition 6.1

$$\epsilon^{i_1,\cdots,i_n} = g^{i_1j_1}\cdots g^{i_nj_n}\epsilon_{j_1,\cdots,j_n} = \frac{|g|^{\frac{1}{2}}}{g}\tilde{\epsilon}^{i_1,\cdots,i_n} = sgn(g)\frac{1}{|g|^{\frac{1}{2}}}\tilde{\epsilon}^{i_1,\cdots,i_n}$$

Using tensor index notation, the Hodge dual is obtained by contracting the indices of a k-form with the n-dimensional completely antisymmetric Levi-Civita tensor.

Proposition 6.2

$$(\star\eta)_{i_1,i_2,\ldots,i_{n-k}} = \frac{1}{(n-k)!}\eta^{j_1,\ldots,j_k}\epsilon_{j_1,\ldots,j_k,i_1,\ldots,i_{n-k}},$$

where η is an arbitrary antisymmetric tensor in k indices.

6.1.3 Metric-induced properties of Riemann curvature tensor

1. In a n dimensional manifold with torsion-free affine connection, the number of independent components of Riemann tensor is

$$\frac{n^3(n-1)}{2} - \frac{n^2(n-1)(n-2)}{6} = \frac{(n^2-1)n^2}{3}$$

   In a n dimensional Riemann manifold, the number of independent components of Riemann tensor is

$$(\frac{n(n-1)}{2})^2 - \frac{n^2(n-1)(n-2)}{6} = \frac{(n^2-1)n^2}{12}$$

2. The double dual of Riemann tensor

$$\bar{G}^{\alpha\beta}{}_{\gamma\delta} = \frac{1}{2}\tilde{\epsilon}^{\alpha\beta\mu\nu}R_{\mu\nu}{}^{\rho\sigma}\frac{1}{2}\tilde{\epsilon}_{\rho\sigma\gamma\delta} = -\frac{1}{4}\delta^{\alpha\beta\mu\nu}_{\rho\sigma\gamma\delta}R_{\mu\nu}{}^{\rho\sigma}$$

   contains precisely the same amount of information as Riemann tensor, and satisfies precisely the same set of symmetries.

3. The Einstein curvature tensor, which is symmetric

$$G^{\beta}{}_{\delta} = \bar{G}^{\mu\beta}{}_{\mu\delta}; \quad G_{\beta\delta} = G_{\delta\beta}$$

4. The Bianchi identity takes a particularly simple form when rewritten in Bianchi identities terms of the double dual $\bar{G}$:

$$\bar{G}^{\alpha\beta}{}_{\gamma\delta}{}^{;\delta} = 0$$

and it has the obvious consequence

$$G_{\beta\delta}{}^{;\delta} = 0$$

5. The Ricci curvature tensor, which is symmetric, and the curvature scalar

$$R^{\beta}{}_{\delta} = R^{\mu\beta}{}_{\mu\delta}; \quad R_{\beta\delta} = R_{\delta\beta}; \quad R = R^{\beta}{}_{\beta}$$

which are related to the Einstein tensor by

$$G^{\beta}{}_{\delta} = R^{\beta}{}_{\delta} - \frac{1}{2}\delta^{\beta}_{\delta}R$$

6. The Weyl conformal tensor

$$C^{\alpha\beta}{}_{\gamma\delta} = R^{\alpha\beta}{}_{\gamma\delta} - 2\delta^{[\alpha}{}_{[\gamma}R^{\beta]}{}_{\delta]} + \frac{1}{3}\delta^{[\alpha}{}_{[\gamma}\delta^{\beta]}{}_{\delta]}R$$

possesses the same symmetries as the Riemann tensor. Weyl tensor is completely "trace-free"; i.e., that contraction of $C_{\alpha\beta\gamma\delta}$ on any pair of slots vanishes. Thus, $C_{\alpha\beta\gamma\delta}$ can be regarded as the trace-free part of Riemann, and $R_{\alpha\beta}$ can be regarded as the trace of Riemann. Riemann is determined entirely by its trace-free part $C_{\alpha\beta\gamma\delta}$ and its trace $R_{\alpha\beta}$

6.2 The coordinates of observer

6.2.1 Riemann normal coordinates

$$g_{\alpha\beta}(\mathcal{P}_0) = \eta_{\alpha\beta}$$

$$g_{\alpha\beta,\gamma}(\mathcal{P}_0) = 0$$

$$\Gamma^{\alpha}{}_{\beta\gamma}(\mathcal{P}_0) = 0$$

$$\Gamma^{\alpha}{}_{\beta\gamma,\mu}(\mathcal{P}_0) = -\frac{1}{3}(R^{\alpha}{}_{\beta\gamma\mu} + R^{\alpha}{}_{\gamma\beta\mu})$$

$$g_{\alpha\beta,\mu\nu}(\mathcal{P}_0) = -\frac{1}{3}(R_{\alpha\mu\beta\nu} + R_{\alpha\nu\beta\mu})$$

$$R_{\alpha\beta\gamma\delta}(\mathcal{P}_0) = g_{\alpha\delta,\beta\gamma} - g_{\alpha\gamma,\beta\mu}$$

6.2.2 The proper reference frame of an accelerated observer

1. Let τ be proper time as measured by the accelerated observer's clock .Let $\mathcal{P} = \mathcal{P}_0(\tau)$ be the observer's world line.

2. The observer carries with himself an orthonormal tetrad $\{\mathbf{e}_{\hat{\alpha}}\}$ with
$$\mathbf{e}_{\hat{0}} = \mathbf{u} = \frac{d\mathcal{P}_0}{d\tau}$$
and with
$$\mathbf{e}_{\hat{\alpha}} \cdot \mathbf{e}_{\hat{\beta}} = \eta_{\alpha\beta}$$

3. The tetrad changes from point to point along the observer's world line, relative to parallel transport:
$$\nabla_{\mathbf{u}} \mathbf{e}_{\hat{\alpha}} = -\mathbf{\Omega} \cdot \mathbf{e}_{\hat{\alpha}}$$
$$\Omega^{\mu\nu} = a^{\mu} u^{\nu} - u^{\mu} a^{\nu} + u_{\alpha} \omega_{\beta} \epsilon^{\alpha\beta\mu\nu}$$
This transport law has the same form in curved space-time as in flat because curvature can only be felt over finite distances, not over the infinitesimal distance involved in the "first time-rate of change of a vector" (equivalence principle).
$$\mathbf{a} = \nabla_{\mathbf{u}} \mathbf{u}$$
$$\mathbf{u} \cdot \mathbf{a} = \mathbf{u} \cdot \omega = 0$$
If ω were zero, the observer would be Fermi-Walker-transporting his tetrad (gyroscope-type transport). If both $\mathbf{a}$ and ω were zero, he would be freely falling (geodesic motion) and would be parallel-transporting his tetrad.

4. The observer constructs his proper reference frame (local coordinate system) in a manner analogous to the Riemann-normal construction. From each event $\mathcal{P}_0(\tau)$ on his world line, he sends out purely spatial geodesics (geodesics orthogonal to $\mathbf{u}$), with affine parameter equal to proper length. The tangent vector has unit length, because the chosen affine parameter is proper length.

5. Each event near the observer's world line is intersected by precisely one of the geodesics $\mathcal{G}[\tau, \mathbf{n}, s]$. [Far away, this is not true; the geodesics may cross, either because of the observer's acceleration].

6. Pick an event $\mathcal{P}$ near the observer's world line. The geodesic through it originated on the observer's world line at a specific time τ, had original direction $\mathbf{n} = n^{\hat{j}} \mathbf{e}_{\hat{j}}$; and needed to extend a distance s before reaching $\mathcal{P}$. Hence, the four numbers
$$(x^{\hat{0}}, x^{\hat{1}}, x^{\hat{2}}, x^{\hat{3}}) \equiv (\tau, sn^{\hat{1}}, sn^{\hat{2}}, sn^{\hat{3}})$$
are a natural way of identifying the event $\mathcal{P}$. These are the coordinates of $\mathcal{P}$ in the observer's proper reference frame.

Along the world line of observer:

$$\frac{\partial}{\partial x^{\hat{\alpha}}} = \mathbf{e}_{\hat{\alpha}}, \quad g_{\hat{\alpha}\hat{\beta}} = \mathbf{e}_{\hat{\alpha}} \cdot \mathbf{e}_{\hat{\beta}} = \eta_{\hat{\alpha}\hat{\beta}}$$

$$\Gamma^{\hat{\beta}}_{\hat{\alpha}\hat{0}} = -\Omega^{\hat{\beta}}_{\hat{\alpha}}, \quad \Gamma^{\hat{0}}_{\hat{0}\hat{0}} = 0$$

$$\Gamma^{\hat{0}}_{\hat{j}\hat{0}} = \Gamma^{\hat{j}}_{\hat{0}\hat{0}} = a^{\hat{j}}, \quad \Gamma^{\hat{j}}_{\hat{k}\hat{0}} = -\omega^{\hat{i}}\epsilon_{0\hat{i}\hat{j}\hat{k}}$$

$$\Gamma^{\hat{\alpha}}_{\hat{j}\hat{k}} = 0$$

$$g_{\hat{\alpha}\hat{\beta},\hat{0}} = 0, \quad g_{\hat{j}\hat{k},\hat{l}} = 0$$

$$g_{\hat{0}\hat{0},\hat{j}} = -2a^{\hat{j}}, \quad g_{\hat{0}\hat{j},\hat{k}} = -\epsilon_{0\hat{j}\hat{k}\hat{l}}\omega^{\hat{l}}$$

6.3 Hypersurfaces

6.3.1 Description of hypersurfaces

Note: *We only discuss timelike and spacelile hypersurfaces in this section.*

Normal vector

$$n^{\alpha}n_{\alpha} = \epsilon \equiv \begin{cases} -1 \text{ if } \Sigma \text{ is spacelike} \\ +1 \text{ if } \Sigma \text{ is timelike} \end{cases}$$

Induced metric

Suppose that the hypersurface is parametrized with equation $x^{\alpha} = x^{\alpha}(y^a)$, Then

$$e^{\alpha}_a = \frac{\partial x^{\alpha}}{\partial y^a}.$$

For displacements within Σ, we have

$$\begin{aligned} ds^2_{\Sigma} &= g_{\alpha\beta}dx^{\alpha}dx^{\beta} \\ &= g_{\alpha\beta}(\frac{\partial x^{\alpha}}{\partial y^a}dy^a)(\frac{\partial x^{\beta}}{\partial y^b}dy^b) \\ &= h_{ab}dy^a dy^b \end{aligned}$$

where $h_{ab} = g_{\alpha\beta}e^{\alpha}_a e^{\beta}_b$. The completeness relation can be written as

$$g^{\alpha\beta} = \epsilon n^{\alpha}n^{\beta} + h^{ab}e^{\alpha}_a e^{\beta}_b$$

6.3.2 Integration on hypersurfaces

The positive volume element of the whole space time is $dx^0 \wedge \cdots \wedge dx^{m-1}$, the positive volume element of the hypersurfaces is $dy^1 \wedge \cdots \wedge dy^{m-1}$. Suppose that the coordinate in hypersurfaces is compatible with the coordinate of the whole space-time, which means that $-dy^m \wedge dy^1 \wedge \cdots \wedge dy^{m-1}$ has the same orientation as $dx^0 \wedge \cdots \wedge dx^{m-1}$. Then we have that

$$\tilde{\epsilon}_{\alpha_m \alpha_1 \cdots \alpha_{m-1}} \frac{\partial x^{\alpha_m}}{\partial y^m} e_1^{\alpha_1} \cdots e_{m-1}^{\alpha_{m-1}} < 0$$

If we demand that the direction of n^α is the opposite of $\frac{\partial x^\alpha}{\partial y^m}$, then we have

$$\tilde{\epsilon}_{\alpha_m \alpha_1 \cdots \alpha_{m-1}} n^{\alpha_m} e_1^{\alpha_1} \cdots e_{m-1}^{\alpha_{m-1}} > 0.$$

Surface element

We define the surface element of a hypersurface as

$$d\Sigma_\mu = \epsilon_{\mu\alpha\beta\gamma} e_1^\alpha e_2^\beta e_3^\gamma dy^1 \wedge dy^2 \wedge dy^3$$

It is easy to verify that

$$f^*(\sqrt{-g} dx^1 \wedge dx^2 \wedge dx^3) = d\Sigma_0$$
$$f^*(-\sqrt{-g} dx^0 \wedge dx^2 \wedge dx^3) = d\Sigma_1$$

and so on. We can demonstrate that

$$d\Sigma_\mu = \epsilon n_\mu |h|^{\frac{1}{2}} dy^1 \wedge dy^2 \wedge dy^3$$

Element of two-surface

Within the hypersurface Σ, we can define a two-surface S, which is parametrized with $y^a = y^a(\theta_A)$, then

$$e_A^a = \frac{\partial y^a}{\partial \theta^A}, \quad e_A^\alpha = \frac{\partial x^\alpha}{\partial \theta^A} = e_a^\alpha e_A^a$$
$$\sigma_{AB} = h_{AB} e_A^a e_B^b = g_{\alpha\beta} e_A^\alpha e_B^\beta$$
$$h^{ab} = \epsilon_r r^a r^b + \sigma^{AB} e_A^a e_B^b$$
$$g^{\alpha\beta} = \epsilon_n n^\alpha n^\beta + \epsilon_r r^\alpha r^\beta + \sigma^{AB} e_A^\alpha e_B^\beta$$

If we demand that the direction r^a is the opposite of that of $\frac{\partial y^a}{\partial \theta^1}$, then the condition of compatibility can be written as

$$\epsilon_{\mu\nu\beta\gamma} n^\mu r^\nu e_2^\beta e_3^\gamma > 0$$

We define the surface element of a two-surface as

$$dS_{\mu\nu} = \epsilon_{\mu\nu\beta\gamma} e_2^\beta e_3^\gamma d\theta^2 \wedge d\theta^3$$

It is easy to verify that

$$f^*(\sqrt{-g} dx^2 \wedge dx^3) = dS_{01}$$

and so on. We can demonstrate that

$$dS_{\alpha\beta} = \epsilon_n \epsilon_r (n_\alpha r_\beta - n_\beta r_\alpha)\sqrt{\sigma} d\theta^2 \wedge d\theta^3$$

Theorem 6.1 Gauss-Stokes theorem

1.
$$\int_M d\omega = \int_{\partial M} i^* \omega$$
2.
$$\int_{\mathcal{V}} A^\alpha_{;\alpha} \sqrt{-g} dx^4 = \oint_{\partial\mathcal{V}} A^\alpha d\Sigma_\alpha$$
3.
$$\int_\Sigma B^{\alpha\beta}{}_{;\beta} d\Sigma_\alpha = \frac{1}{2}\oint_{\partial\Sigma} B^{\alpha\beta} d\Sigma_{\alpha\beta},$$
where $B_{\alpha\beta}$ is an antisymmetric tensor

6.3.3 Differentiation of tangent vector fields

Tangent tensor field

$$A^{\alpha\beta\cdots} = A^{ab\cdots} e^\alpha_a e^\beta_b \cdots$$
$$A_{\alpha\beta\cdots} e^\alpha_a e^\beta_b \cdots = A_{ab\cdots} = h_{am} h_{bn} \cdots A^{mn\cdots}$$

Projection tensor

$$h^{\alpha\beta} \equiv h^{ab} e^\alpha_a e^\beta_b = g^{\alpha\beta} - \epsilon n^\alpha n^\beta$$

Intrinsic covariant derivative

$$A_{a|b} \equiv A_{\alpha;\beta} e^\alpha_a e^\beta_b = A_{a,b} - \Gamma^c_{ab} A_c$$

Here, the connection Γ^c_{ab} is compatible with h_{ab}.

Extrinsic curvature

$$K_{ab} \equiv n_{\alpha;\beta} e^\alpha_a e^\beta_b$$
$$A^\alpha_{;\beta} e^\beta_b = A^a_{|b} e^\alpha_a - \epsilon A^a K_{ab} n^\alpha$$
$$e^\alpha_{a;\beta} e^\beta_b = \Gamma^c_{ab} e^\alpha_c - \epsilon K_{ab} n^\alpha$$

$$K_{ab} = \frac{1}{2}(\mathcal{L}_n g_{\alpha\beta}) e^\alpha_a e^\beta_b$$

$$K \equiv h^{ab} K_{ab} = n^\alpha_{;\alpha}$$

Theorem 6.2 Gauss-Codazzi theorem

1.

$$R^\mu_{\alpha\beta\gamma} e^\alpha_a e^\beta_b e^\gamma_c = R^m_{abc} e^\mu_m + \epsilon(K_{ab|c} - K_{ac|b}) n^\mu + \epsilon K_{ab} n^\mu_{;\gamma} e^\gamma_c - \epsilon K_{ac} n^\mu_{;\beta} e^\beta_b$$

2.

$$-2\epsilon G_{\alpha\beta} n^\alpha n^\beta = \ {}^3R + \epsilon(K^{ab}K_{ab} - K^2)$$

$$G_{\alpha\beta} e^\alpha_a n^\beta = K^b_{a|b} - K_{,a}$$

3.

$$R = \ {}^3R + \epsilon(K^2 - K^{ab}K_{ab}) + 2\epsilon(n^\alpha_{;\beta} n^\beta - n^\alpha n^\beta_{;\beta})_{;\alpha}$$

♣

Chapter 7
Formulation of General Relativity

Give the fields that generate mass-energy, and their time-rates of change, and give 3-geometry of space and its time-rate of change, all at one time, and solve for the 4-geometry of spacetime at that one time. Four of the ten components of Einstein's law connect the curvature of space here and now with the distribution of mass-energy here and now, and the other six equations tell how the geometry as thus determined then proceeds to evolve.

7.1 Basic assumptions of general relativity

1. Space-time is a four dimensional pseudo-Riemann manifold.
2. The metric of the manifold is governed by the Einstein field equation
$$\mathbf{G} = 8\pi\mathbf{T}.$$
3. All special relativistic laws of physics are valid in local Lorentz frames of metric.

7.2 Lagrangian formulation

7.2.1 Mechanics

$$S[q] = \int_{\tau_1}^{\tau_2} L(x^\alpha, \frac{dx^\alpha}{d\tau})d\tau$$

$$\delta x^\alpha(\tau_1) = 0, \delta x^\alpha(\tau_2) = 0$$

$$\delta S = 0 \Rightarrow \frac{d}{d\tau}\frac{\partial L}{\partial u^\alpha} - \frac{\partial L}{\partial x^\alpha} = 0$$

Example:

$$L = -m(-g_{\mu\nu}u^\mu u^\nu)^{-1/2} + eA_\mu u^\mu$$

$$\Downarrow$$

$$m(\frac{du_\alpha}{d\tau} - \frac{1}{2}\frac{\partial g_{uv}}{\partial x^\alpha}u^\mu u^\nu) = e(A_{\mu,\alpha} - A_{\alpha,\mu})u^\mu \Rightarrow ma_\alpha = eF_{\alpha\mu}u^\mu$$

7.2.2 Field Theory

$$S[q] = \int_{\mathcal{V}} \mathcal{L}(q, q_\alpha)\sqrt{-g}d^4x$$

$$\delta q|_{\partial\mathcal{V}} = 0$$

$$\delta S = 0 \Rightarrow \nabla_\alpha(\frac{\partial\mathcal{L}}{\partial q_{,\alpha}}) - \frac{\partial\mathcal{L}}{\partial q} = 0$$

Example:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + A_\mu j^\mu$$

$$\Downarrow$$

$$F^{\mu\nu}{}_{;\nu} = j^\mu$$

7.2.3 General relativity

$$S_H[g] = \frac{1}{16\pi}\int_{\mathcal{V}} R\sqrt{-g}d^4x$$

$$S_B[g] = \frac{1}{8\pi}\oint_{\partial\mathcal{V}} \epsilon K|h|^{\frac{1}{2}}d^3y$$

$$S_0 = \frac{1}{8\pi}\oint_{\partial\mathcal{V}} \epsilon K_0|h|^{\frac{1}{2}}d^3y$$

$$S_M[\phi; g] = \int_{\partial\mathcal{V}} \mathcal{L}(\phi, \phi_{,\alpha}; g_{\alpha\beta})\sqrt{-g}d^4x$$

Variation of Hilbert term

$$(16\pi)\delta S_H = \int_{\mathcal{V}} G_{\alpha\beta}\delta g^{\alpha\beta}\sqrt{-g}d^4x - \oint_{\partial\mathcal{V}} \epsilon h^{\alpha\beta}\delta g_{\alpha\beta,\mu}n^\mu|h|^{\frac{1}{2}}d^3y$$

Variation of boundary term

$$16\pi\delta S_B = \oint_{\partial\mathcal{V}} \epsilon h^{\alpha\beta}\delta g_{\alpha\beta,\mu}n^\mu|h|^{\frac{1}{2}}d^3y$$

Variation of matter action

$$\delta S_M = \int_{\mathcal{V}} \left(\frac{\partial \mathcal{L}}{\partial g^{\alpha\beta}} - \frac{1}{2}\mathcal{L} g_{\alpha\beta} \right) \delta g^{\alpha\beta} \sqrt{-g} d^4x$$

Define

$$T_{\alpha\beta} \equiv -2\frac{\partial \mathcal{L}}{\partial g^{\alpha\beta}} + \mathcal{L} g_{\alpha\beta}$$

Example:

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}$$

$$\Downarrow$$

$$T_{\alpha\beta} = F_{\mu\alpha} F^{\mu}{}_{\beta} - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} g_{\alpha\beta}$$

Nondynamical term

$K_0 =$ extrinsic curvature of $\partial\mathcal{L}$ embedded in flat space-time.

7.3 Hamiltonian formulation

7.3.1 3+1 decomposition

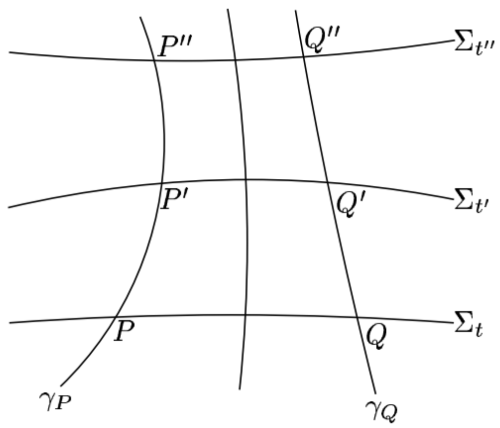


Figure 7.1: Foliation of space-time by spacelike hypersurfaces

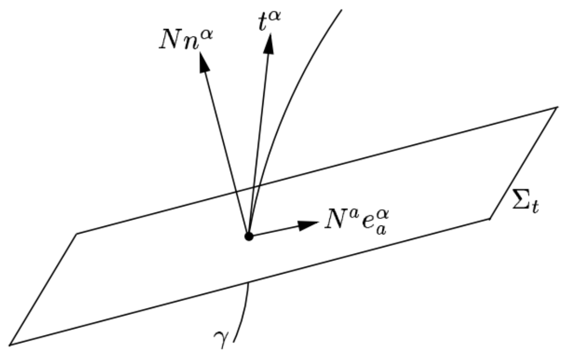


Figure 7.2: Decomposition of t^α into lapse and shift

The space-time is foliated by spacelike hypersurfaces Σ_t that is described by scalar function $t(x^\alpha)$. t is a single valued function and the unit normal to the hypersurfaces $n_\alpha \propto \partial_\alpha t$ is a future directed timelike vector field.
Consider a congruence of curves γ intersecting Σ_t. We use t as a parameter on the curves and the vector t^α is tangent to the congruence (so, $t^\alpha \partial_\alpha t = 1$).Install coordinates y^a on Σ_t and

impose $y^a(P'') = y^a(P') = y^a(P)$, so y^a is held constant on each member of the congruence. This construction defines a coordinate system (t, y^a) in $\mathcal{V}$.

base vector

$$t^\alpha = \left(\frac{\partial x^\alpha}{\partial t}\right)_{y^a}, \quad e^\alpha_a = \left(\frac{\partial x^\alpha}{\partial y^a}\right)_t, \quad \mathcal{L}_t e^\alpha_a = 0$$

Normal vector

$$n_\alpha = -N\partial_\alpha t, \quad n_\alpha e^\alpha_a = 0$$

Decomposition of t^α

$$t^\alpha = Nn^\alpha + N^a e^\alpha_a$$

Metric

$$\begin{aligned}
ds^2 &= g_{\alpha\beta}dx^\alpha dx^\beta \\
&= g_{\alpha\beta}(t^\alpha dt + e^\alpha_a dy^a)(t^\beta dt + e^\beta_b dy^b) \\
&= -N^2dt^2 + h_{ab}(dy^a + N^a dt)(dy^b + N^b dt) \\
\sqrt{-g} &= N\sqrt{h}
\end{aligned}$$

7.3.2 Field theory

$$\dot{q} = \frac{\partial q}{\partial t}, \quad p = \frac{\partial}{\partial \dot{q}}(\sqrt{-g}\mathcal{L})$$

$$\mathcal{H}(p, q, q_a) = p\dot{q} - \sqrt{-g}\mathcal{L}$$

$$H = \int_{\Sigma_t} \mathcal{H}(p, q, q_{,a})d^3y$$

$$S = \int_{t_1}^{t_2} dt \int_{\Sigma_t} (p\dot{q} - \mathcal{H})d^3y$$

$$\delta S = 0 \Rightarrow \dot{p} = -\frac{\partial \mathcal{H}}{\partial q} + \left(\frac{\partial \mathcal{H}}{\partial q_{,a}}\right)_{,a}, \quad \dot{q} = \frac{\partial \mathcal{H}}{\partial p}$$

Example: For electromagnetic field in 3+1 decomposition form, we define the electrical field as $E_a = F_{\alpha\beta}n^\beta e^\alpha_a$, the magnetic field as $\epsilon_{abc}B^c = F_{\alpha\beta}e^\alpha_a e^\beta_b$. In this definition, the equation of motion of particles in electromagnetic field can be written as

$$mA_a = \gamma e(E_a + \epsilon_{abc}v^b B^c)$$

Here, $A_a = u_{\alpha;\beta}u^\beta e^\alpha_a$, $\gamma = \frac{1}{\sqrt{1-v^2}}$. So, the three force

$$\vec{f} = \frac{d\vec{p}}{dt} = e(\vec{E} + \vec{v} \times \vec{B})$$

If we adopt the coordinates (t, y^a), it is easy to verify that

$$E^a = NF^{0a}, \quad B^a = \frac{1}{2}\epsilon^{abc}F_{bc}$$

We further define

$$\mathcal{E}^a = \sqrt{h}E^a, \quad \mathcal{B}^a = \sqrt{h}B^a, \quad \phi = -A_0, \quad \rho_e = -j^\alpha n_\alpha = Nj^0$$

If we notice that

$$F_{0a} = -h_{ab}N^2F^{0b} - F_{ab}N^b, \quad \tilde{\epsilon}_{abc}\tilde{\epsilon}_{ijk}h^{ai}h^{bj} = \frac{2h_{ck}}{h}$$

It is easy to verify that

$$\sqrt{-g}\mathcal{L} = -\mathcal{E}^a\dot{A}_a + \phi\mathcal{E}^a_{,a} - \frac{1}{2}Nh^{-\frac{1}{2}}h_{ab}(\mathcal{E}^a\mathcal{E}^b + \mathcal{B}^a\mathcal{B}^b) + \tilde{\epsilon}_{abc}N^a\mathcal{E}^b\mathcal{B}^c - \sqrt{h}\phi\rho_e + N\sqrt{h}A_aj^a$$

So,$\pi^a = -\mathcal{E}^a$, and we can get the Hamilton density

$$\mathcal{H} = \phi\pi^a_{,a} + \frac{1}{2}Nh^{-\frac{1}{2}}h_{ab}(\pi^a\pi^b + \mathcal{B}^a\mathcal{B}^b) + \tilde{\epsilon}_{abc}N^a\pi^b\mathcal{B}^c + \sqrt{h}\phi\rho_e - N\sqrt{h}A_aj^a$$

Then, the Hamilton equation can be written as

$$\dot{A}_a = -\phi_a + Nh^{-\frac{1}{2}}h_{ab}\pi^b + \tilde{\epsilon}_{abc}N^a\mathcal{B}^c$$

$$\dot{\pi}^a = -\tilde{\epsilon}^{jab}(Nh^{-\frac{1}{2}}h_{ij}\mathcal{B}^i)_{,b} - \tilde{\epsilon}^{cab}(\tilde{\epsilon}_{ijc}N^i\pi^j)_{,b} + N\sqrt{h}j^a$$

and also the constraint equation $\pi^a_{,a} + \sqrt{h}\rho_e = 0$. After simplification, the Maxwell equations are

$$\begin{aligned}
\frac{1}{\sqrt{h}}\frac{\partial}{\partial t}(\sqrt{h}\vec{E}) &= \nabla \times (N\vec{B} - \vec{N} \times \vec{E}) - N\vec{J} \\
\frac{1}{\sqrt{h}}\frac{\partial}{\partial t}(\sqrt{h}\vec{B}) &= -\nabla \times (N\vec{E} + \vec{N} \times \vec{B}) \\
\nabla \cdot \vec{E} &= \rho_e \\
\nabla \cdot \vec{B} &= 0
\end{aligned}$$

7.3.3 General relativity

$$S_G = \frac{1}{16\pi}\int_{t_1}^{t_2} dt \left\{ \int_{\Sigma_t} \left({}^3R + K^{ab}K_{ab} - K^2 \right) N\sqrt{h}d^3y + 2\oint_{\Sigma_t} (k - k_0)N\sqrt{\sigma}d^2\theta \right\}$$

$k_0 =$ extrinsic curvature of S_t embedded in flat space.

Gravitational Hamiltonian

$$\dot{h}_{ab} \equiv \mathcal{L}_t h_{ab} = \mathcal{L}_t(g_{\alpha\beta}e^\alpha_a e^\beta_b) = \mathcal{L}_t(g_{\alpha\beta}e^\alpha_a e^\beta_b = 2NK_{ab} + N_{a|b} + N_{b|a})$$

$$K_{ab} = \frac{1}{2N}(\dot{h}_{ab} - N_{a|b} - N_{b|a})$$

$$p^{ab} = \frac{\partial}{\partial \dot{h}_{ab}}(\sqrt{-g}\mathcal{L}_G) = \frac{\sqrt{h}}{16\pi}(K^{ab} - Kh^{ab})$$

$$\sqrt{h}K^{ab} = 16\pi(p^{ab} - \frac{1}{2}ph^{ab})$$

$$\mathcal{H}_G = p^{ab}\dot{h}_{ab} - \sqrt{-g}\mathcal{L}_G$$

$$\begin{aligned}16\pi H_G &= \int_{\Sigma_t}\left[N(K^{ab}K_{ab} - K^2 - \ {}^3R) - 2N_a(K^{ab} - Kh^{ab})_{|b}\right]\sqrt{h}d^3y \\ &\quad -2\oint_{S_t}\left[N(k - k_0) - N_a(K^{ab} - Kh^{ab})r_b\right]\sqrt{\sigma}d^2\theta\end{aligned}$$

Variation of gravitational Hamiltonian

$$\delta N = \delta N^a = \delta h_{ab} = 0 \text{ on } S_t$$

$$\delta H_G = \int_{\Sigma_t}(\mathcal{P}^{ab}\delta h_{ab} + \mathcal{H}_{ab}\delta p^{ab} - \mathcal{C}\delta N - 2\mathcal{C}_a\delta N^a)d^3y$$

$$\begin{aligned}(16\pi)\mathcal{P}^{ab} &= N\sqrt{h}G^{ab} - \sqrt{h}(N^{|ab} - h^{ab}N^{|c}{}_c) \\ &\quad +(16\pi)[2p^{c(a}N^{b)}{}_{|c} - \sqrt{h}(\frac{1}{\sqrt{h}}p^{ab}N^c)_{|c}] \\ &\quad +(16\pi)^2[\frac{2N}{\sqrt{h}}(p^a_c p^{bc} - \frac{1}{2}pp^{ab}) - \frac{N}{2\sqrt{h}}(p^{cd}p_{cd} - \frac{1}{2}p^2)h^{ab}] \\ \mathcal{H}_{ab} &= (16\pi)\frac{2N}{\sqrt{h}}(p_{ab} - \frac{1}{2}ph_{ab}) + 2N_{(a|b)} \\ \mathcal{C} &= \frac{\sqrt{h}}{16\pi}(\ {}^3R + K^2 - K^{ab}K_{ab}) \\ \mathcal{C}^a &= \frac{\sqrt{h}}{16\pi}(K_a{}^b - K\delta_a{}^b)_{|b}\end{aligned}$$

Variation of electromagnetic Hamiltonian

$$\delta H_E = \int_{\Sigma_t}(-\frac{1}{2}N\sqrt{h}\mathcal{I}^{ab}\delta h_{ab} + \sqrt{h}\rho\delta N - \sqrt{h}s_a\delta N^a)$$

$$\begin{aligned}\mathcal{I}^{ab} &= \frac{1}{2}(E^cE_c + B^cB_c)h^{ab} - E^aE^b - B^aB^b \\ \rho &= \frac{1}{2}(E^cE_c + B^cB_c) \\ s_a &= \epsilon_{abc}E^bB^c\end{aligned}$$

Hamilton's equations

$$\dot{h}_{ab} = \mathcal{H}_{ab}, \quad \dot{p}^{ab} = -\mathcal{P}^{ab} + \frac{1}{2}N\sqrt{h}\mathcal{I}^{ab}$$

$$^3R + K^2 - K^{ab}K_{ab} = 16\pi\rho$$

$$(K_a{}^b - K\delta_a{}^b)_{|b} = -8\pi s_a$$

Part IV

Quantum Mechanics

Chapter 8
Linear Algebra

8.1 Linear Vector Space

8.1.1 Definition

Definition 8.1 Linear vector space

A linear vector space is a set of elements, called vectors, which is closed under addition and multiplication by scalars. That is to say, if ϕ and ψ are vectors then so is $a\phi + b\psi$, where a and b are arbitrary scalars. If the scalars belong to the field of complex (real) numbers, we speak of a complex (real) linear vector space. Henceforth the scalars will be complex numbers unless otherwise stated.

Example:

1. Discrete vectors, which may be represented as columns of complex numbers.

2. Spaces of functions of some type, for example the space of all differentiable functions

8.1.2 Linear independence

Definition 8.2 Linear independence

A set of vectors $\{\phi_n\}$ is said to be linearly independent if no non-trivial linear combination of them sums to zero; that is to say, if the equation $\sum_n c_n\phi_n$ can hold only when $c_n = 0$ for all n. If this condition does not hold, the set of vectors is said to be linearly dependent, in which case it is possible to express a member of the set as a linear combination of the others.

Definition 8.3 Dimension

The maximum number of linearly independent vectors in a space is called the dimension of the space.

Definition 8.4 Base

A maximal set of linearly independent vectors is called a basis for the space. Any vector in the space can be expressed as a linear combination of the basis vectors.

8.1.3 Inner product

Definition 8.5 Inner product

An inner product (or scalar product) for a linear vector space associates a scalar (ϕ, ψ) with every ordered pair of vectors. It must satisfy the following properties:

1. $(\phi, \psi) =$ *a complex number*
2. $(\phi, \psi) = (\psi, \phi)^*$
3. $(\phi, c_1\phi_1 + c_2\psi_2) = c_1(\phi, \psi_1) + c_2(\phi, \psi_2)$
4. $(\phi, \phi) > 0$*,with equality holding if and only if* $\phi = 0$

Example:

1. If ψ is the column vector with elements a_1, a_2, $\cdots$, and ϕ is the column vector with elements b_1, b_2, $\cdots$, then

$$(\psi, \phi) = a_1^* b_1 + a_2^* b_2 + \cdots$$

2. If ψ and ϕ are functions of x, then

$$(\phi, \psi) = \int \psi^*(x)\phi(x)w(x)dx$$

where $w(x)$ is some non-negative weight function.

Definition 8.6 Norm

$$||\phi|| = (\phi, \phi)^{\frac{1}{2}}$$

Theorem 8.1 Schwarz's inequality

$$|(\psi, \phi)|^2 \leq (\psi, \psi)(\phi, \phi)$$

Theorem 8.2 triangle inequality

$$||(\psi + \phi)|| \leq ||\phi|| + ||\psi||$$

Definition 8.7 Orthonormal

A set of vectors $\{\phi_n\}$ *is said to be orthonormal if the vectors are pairwise orthogonal and of unit norm; that is to say, their inner products satisfy* $(\psi_m, \phi_n) = \delta_{mn}$.

8.1.4 Dual space

Definition 8.8 Dual vector

Corresponding to any linear vector space V *there exists the dual space of linear functionals on* V . *A linear functional* F *assigns a scalar* $F(\phi)$ *to each vector* ϕ, *such that*

$$F(a\phi + b\psi) = aF(\phi) + bF(\psi)$$

for any vectors for ϕ *and* ψ, *and any scalars* a *and* b. *The set of linear functionals may itself be regarded as forming a linear space* V' *if we define the sum of two functionals as*

$$(F_1 + F_2)(\phi) = F_1(\phi) + F_2(\phi)$$

Theorem 8.3 Riesz theorem

There is a one-to-one correspondence between linear functionals F in V' and vectors f in V, such that all linear functionals have the form

$$F(\phi) = (f, \phi)$$

f being a fixed vector, and ϕ being an arbitrary vector. Thus the spaces V and V' are essentially isomorphic.

8.1.5 Dirac's bra and ket notation

In Dirac' s notation, which is very popular in quantum mechanics, the vectors in V are called ket vectors, and are denoted as $|\phi\rangle$. The linear functionals in the dual space V' are called bra vectors, and are denoted as $\langle F|$. The numerical value of the functional is denoted as

$$F(\phi) = \langle F|\phi\rangle$$

According to the Riesz theorem, there is a one-to-one correspondence between bras and kets. Therefore we can use the same alphabetic character for the functional (a member of V') and the vector (in V) to which it corresponds, relying on the bra, $\langle F|$, or ket, $|F\rangle$, notation to determine which space is referred to.So

$$\langle F|\phi\rangle = (F, \phi)$$

Note that the Riesz theorem establishes, by construction, an antilinear correspondence between bras and kets. If $\langle F| \leftrightarrow |F\rangle$, then

$$c_1^*\langle F_1| + c_2^*\langle F_2| \leftrightarrow c_1|F_1\rangle + c_2|F_2\rangle$$

8.2 Linear Operators

Definition 8.9 Linear operators

An operator on a vector space maps vectors onto vectors. A linear operator satisfies

$$A(c_1\psi_1 + c_2\psi_2) = c_1A(\psi_1) + c_2A(\psi_2)$$

Define the sum and product of operators,

$$\begin{aligned}(A+B)\psi &= A\psi + B\psi \\ AB\psi &= A(B\psi)\end{aligned}$$

Define their action to the left on bra vectors as

$$(\langle\phi|A)\psi\rangle = \langle\phi|(A|\psi\rangle)$$

So we may define the operation of A on the bra space of functionals as

$$AF_\phi(\psi) = F_\phi(A\psi)$$

According to the Riesz theorem there must exist a ket vector χ such that

$$AF_\phi(\psi) = (\chi, \psi) = F_\chi(\psi)$$

Define operator $A^\dagger$ as

$$AF_\phi = F_{A^\dagger\chi}$$

Therefore,

$$(A^\dagger\phi, \psi) = (\phi, A\psi)$$

$$\langle\phi|A^\dagger|\psi\rangle^* = \langle\psi|A|\phi\rangle$$

Definition 8.10 Outer product

$$(|\psi\rangle\langle\phi|)|\lambda\rangle \equiv |\psi\rangle(\langle\phi|\lambda\rangle)$$

Definition 8.11 Trace

$$\mathrm{Tr} A \equiv \sum \langle u_j|A|u_j\rangle$$

where $\{u_j\}$ may be any orthonormal basis. It can be shown that the value of $\mathrm{Tr} A$ is independent of the particular orthonormal basis that is chosen for its evaluation.

Proposition 8.1

$$\begin{aligned}(cA)^\dagger &= c^* A^\dagger \\ (A+B)^\dagger &= A^\dagger + B^\dagger \\ (AB)^\dagger &= B^\dagger A^\dagger\end{aligned}$$

$$(|\psi\rangle\langle\phi|)^\dagger = |\phi\rangle\langle\psi|$$

8.3 Self-Adjoint operators

Definition 8.12 Self-Adjoint operators

An operator A that is equal to its adjoint $A^\dagger$ is called self-adjoint. This means that it satisfies

$$\langle\phi|A|\psi\rangle = \langle\psi|A|\phi\rangle^*$$

and that the domain of A coincides with the domain of $A^\dagger$. An operator that only satisfies above equation is called Hermitian.

Theorem 8.4

If $\langle\psi|A|\psi\rangle = \langle\psi|A|\psi\rangle^*$ for all $|\psi\rangle$, then it follows that $\langle\phi_1|A|\phi_2\rangle = \langle\phi_2|A|\phi_1\rangle^*$ for all $|\phi_1\rangle$ and $|\phi_2\rangle$, and hence that $A = A^\dagger$.
If an operator acting on a certain vector produces a scalar multiple of that same vector,

$$A|\phi\rangle = a|\phi\rangle$$

we call the vector $|\phi\rangle$ an eigenvector and the scalar a an eigenvalue of the operator A. The antilinear correspondence between bras and kets, and the definition of the adjoint operator $A^\dagger$, imply that the left-handed eigenvalue equation

$$\langle\phi|A^\dagger = a^*\langle\phi|$$

Theorem 8.5

If A is a Hermitian operator then all of its eigenvalues are real.

Theorem 8.6

Eigenvectors corresponding to distinct eigenvalues of a Hermitian operator must be orthogonal.

If the orthonormal set of vectors $\{\phi_i\}$ is complete, then we can expand an arbitrary vector $|v\rangle$ in terms of it:

$$|v\rangle = \sum|\phi_i\rangle(\langle\phi_i|v\rangle) = \left(\sum|\phi_i\rangle\langle\phi_i|\right)|v\rangle$$

So,

$$\sum|\phi_i\rangle\langle\phi_i| = I$$

If $A|\phi_i\rangle = a_i|\phi_i\rangle$ and the eigenvectors form a complete orthonormal set, then the operator can be reconstructed in a useful diagonal form in terms of its eigenvalues and eigenvectors:

$$A = \sum a_i|\phi_i\rangle\langle\phi_i|$$

We can define a function of an operator

$$f(A) = \sum f(a_i)|\phi_i\rangle\langle\phi_i|$$

The Hermitian operators in a finite N-dimensional vector space have complete sets of eigenvectors. But This statement does not carry over to infinite-dimensional spaces. A Hermitian operator in an infinite dimensional vector space may or may not possess a complete set of eigenvectors, depending upon the precise nature of the operator and the vector space. Instead, we have spectral theorem.

Theorem 8.7

To each self-adjoint operator A there corresponds a unique family of projection operators, $E(\lambda)$, for real λ, with the properties:

1. If $\lambda_1 < \lambda_2$ then $E(\lambda_1)E(\lambda_2) = E(\lambda_2)E(\lambda_1)E(\lambda_1)$
2. If $\epsilon > 0$, then $E(\lambda+\epsilon)|\psi\rangle \to E(\lambda)|\psi\rangle$ as $\epsilon \to 0$
3. $E(\lambda)|\psi\rangle \to 0$ as $\lambda \to -\infty$
4. $E(\lambda)|\psi\rangle \to |\psi\rangle$ as $\lambda \to \infty$
5. $\int_{-\infty}^{\infty} \lambda E(\lambda) = A$

We can define a function of an operator

$$f(A) = \int_{-\infty}^{\infty} f(\lambda)E(\lambda)$$

Following Dirac's pioneering formulation, it has become customary in quantum mechanics to write a formal eigenvalue equation for an operator such as Q that has a continuous spectrum,

$$Q|q\rangle = q|q\rangle$$

The orthonormality condition for the continuous case takes the form

$$\langle q'|q''\rangle = \delta(q-q')$$

Evidently the norm of these formal eigenvectors is infinite, since $\langle q|q\rangle \to \infty$. Instead of the spectral theorem for Q, Dirac would write

$$Q = \int_{-\infty}^{\infty} q|q\rangle\langle q|dq$$

Dirac's formulation does not fit into the mathematical theory of Hilbert space, which admits only vectors of finite norm. The projection operator formally given by

$$E(\lambda) = \int_{-\infty}^{\lambda} |q\rangle\langle q|dq$$

is is well defined in Hilbert space, but its derivative does not exist within the Hilbert space framework.

Theorem 8.8

If A and B are self-adjoint operators, each of which possesses a complete set of eigenvectors, and if $AB = BA$, then there exists a complete set of vectors which are eigenvectors of both A and B.

Let $(A, B, \cdots)$ be a set of mutually commutative operators that possess a complete set of common eigenvectors. Corresponding to a particular eigenvalue for each operator, there may be more than one eigenvector. If, however, there is no more than one eigenvector (apart from the arbitrary phase and normalization) for each set of eigenvalues $(a_n, b_m, \cdots)$, then the operators $(A, B, \cdots)$ are said to be a complete commuting set of operators.

Theorem 8.9

Any operator that commutes with all members of a complete commuting set must be a function of the operators in that set. ♣

8.4 Rigged Hilbert space

Definition 8.13 Rigged Hilbert spece

Formally, a rigged Hilbert space consists of a Hilbert space $\mathcal{H}$, together with a subspace Φ which carries a finer topology, that is one for which the natural inclusion $\Phi \subseteq \mathcal{H}$ is continuous. It is no loss to assume that Φ is dense in $\mathcal{H}$ for the Hilbert norm. We consider the inclusion of conjugate space $\mathcal{H}^X$ in Φ^X. Φ^X is the space of τ_Φ continuous antilinear functional on Φ.
For any $\phi \in \Phi$, $F \in \Phi^X$,we define

$$\langle \phi | F \rangle \equiv F(\phi)$$

$$\langle F | \phi \rangle \equiv F^*(\phi)$$

Now by applying the Riesz representation theorem we can identify $\mathcal{H}^X$ with $\mathcal{H}$. Therefore, the definition of rigged Hilbert space is in terms of a sandwich:

$$\Phi \subseteq \mathcal{H} \subseteq \Phi^X$$

♡

There may or may not exist any solutions to the eigenvalue equation $A|a_n\rangle = a_n|a_n\rangle$ for a self-adjoint operator A on an infinite-dimensional vector space. However, the generalized spectral theorem asserts that if A is self-adjoint in $\mathcal{H}$ then a complete set of eigenvectors exists in the extended space Φ^X. The precise conditions for the proof of this theorem are rather technical, so the interested reader is referred to *Gel'fand and Vilenkin (1964)* for further details.

There are many examples of rigged-Hilbert-space triplets. A Hilbert space $\mathcal{H}$ is formed by those functions that are square-integrable. That is, $\mathcal{H}$ consists of those functions $\psi(x)$ for which

$$\langle \psi | \psi \rangle = \int_{-\infty}^{\infty} |\psi(x)|^2 dx \text{ is finite}$$

A nuclear space Φ is made up of functions $\psi(x)$ which satisfy the infinite set of conditions,

$$\int_{-\infty}^{\infty} |\psi(x)|^2 (1 + |x|)^m dx \text{ is finite for } m = 0, 1, 2, \cdots$$

The functions $\psi(x))$ which make up Φ must vanish more rapidly than any inverse power of x in the limit $|x| \to \infty$. The extended space Φ^X, which is conjugate to Φ, consists of those functions $\chi(x)$ for which

$$\langle\chi|\psi\rangle = \int_{-\infty}^{\infty} \chi^*(x)\psi(x)dx \text{ is finite for any } \psi \text{ in } \Phi$$

In addition to the functions of finite norm, which also lie in $\mathcal{H}$, Φ^X will contain functions that are unbounded at infinity provided the divergence is no worse than a power of x. Hence Φ^X contains e^{ikx}, which is an eigenfunction of the operator $D = i\frac{d}{dx}$. It also contains the Dirac delta function, $\delta(x - \lambda)$, which is an eigenfunction of the operator X, defined by $X\psi(x) = x\psi(x)$. These two examples suffice to show that rigged Hilbert space seems to be a more natural mathematical setting for quantum mechanics than is Hilbert space.

8.5 Unitary operators

Definition 8.14 Unitary operator

A unitary operator is a bounded linear operator $U : H \to H$ on a Hilbert space H that satisfies $UU^\dagger = U^\dagger U = I$, where $U^\dagger$ is the adjoint of U, and $I : H \to H$ is the identity operator. ♡

Consider a family of unitary operators, $U(s)$, that depend on a single continuous parameter s. Let $U(0) = I$ be the identity operator, and let $U(s_1 + s_2) = U(s_1)U(s_2)$. We can demonstrate that

$$\left.\frac{dU}{ds}\right|_{s=0} = iK \text{ with } K = K^\dagger$$

The Hermitian operator K is called the generator of the family of unitary operators because it determines $U(s)$, not only for infinitesimal s, but for all s. This can be shown by differentiating

$$U(s_1 + s_2) = U(s_1)U(s_2)$$

with respect to s_2 and we can get

$$\left.\frac{dU}{ds}\right|_{s=s_1} = U(s_1)iK$$

This first order differential equation with initial condition $U(0) = I$ has the unique solution

$$U(s) = e^{iKs}$$

8.6 Antiunitary operators

Definition 8.15 Antiunitary operator

In mathematics, an antiunitary transformation, is a bijective antilinear map

$$U : H_1 \to H_2$$

between two complex Hilbert spaces such that

$$\langle Ux, Uy\rangle = \overline{\langle x, y\rangle}$$

for all x and y y in H_1, where the horizontal bar represents the complex conjugate. If additionally one has $H_1 = H_2$ then U is called an antiunitary operator.

Proposition 8.2

1. $\langle Ux, Uy\rangle = \overline{\langle x, y\rangle} = \langle y, x\rangle$ holds for all elements x, y of the Hilbert space and an antiunitary U.
2. When U is antiunitary then U^2 is unitary. This follows from
$$\langle U^2x, U^2y\rangle = \overline{\langle Ux, Uy\rangle} = \langle x, y\rangle$$
3. For unitary operator V the operator VK, where K is complex conjugate operator, is antiunitary. The reverse is also true, for antiunitary U the operator UK is unitary.
4. For antiunitary U the definition of the adjoint operator U^* is changed into
$$\langle U^*x, y\rangle = \overline{\langle x, Uy\rangle}$$
5. The adjoint of an antiunitary U is also antiunitary and $UU^* = U^*U = 1$.

Chapter 9
Formulation of quantum mechanics

9.1 Axioms of quantum mechanics

1. The properties of a quantum system are completely defined by specification of its state vector $|\psi\rangle$. The state vector is an element of a complex Hilbert space $\mathcal{H}$ called the space of states.

2. With every physical property A (energy, position, momentum, angular momentum, ...) there exists an associated linear, Hermitian operator A (usually called observable), which acts in the space of states. The eigenvalues of the operator are the possible values of the physical properties.

3.
   - If $|\psi\rangle$ is the vector representing the state of a system and if $|\phi\rangle$ represents another physical state, there exists a probability $P(|\psi\rangle, |\phi\rangle)$ of finding $|\psi\rangle$ in state $|\phi\rangle$, which is given by the squared modulus of the scalar product on $\mathcal{H}: \ P(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2$ (Born Rule)
   - If A is an observable with eigenvalues a_k and eigenvectors $|k\rangle$, given a system in the state $|\psi\rangle$, the probability of obtaining a_k as the outcome of the measurement of A is $|\langle k|\psi\rangle|^2$. After the measurement the system is left in the state projected on the subspace of the eigenvalue a_k (Wave function collapse).

4. The evolution of a closed system is unitary. The state vector $\psi(t)\rangle$ at time t is derived from the state vector $\psi(t_0)\rangle$ at time t_0 by applying a unitary operator $U(t, t_0)$, called the evolution operator: $\psi(t)\rangle = U(t, t_0)\psi(t)\rangle$.

9.2 Transformations of States

A transformation of states can be described by $|\psi\rangle \to U(\tau)|\psi\rangle \equiv |\psi'\rangle$. And we demand that

$$|\langle\phi|\psi\rangle| = |\langle\phi'|\psi'\rangle|$$

Theorem 9.1 Wigner Theorem

Any mapping of the vector space onto itself that preserves the value of $|\langle\phi|\psi\rangle|$ may be implemented by an operator U with U being either unitary (linear) or antiunitary (antilinear). ♣

Continuous transformation

Only linear operators can describe continuous transformations because every continuous transformation has a square root. Suppose, for example, that $U(l)$ describes a displacement through the distance l. This can be done by two displacements of $U(l/2)$, and hence $U(l) = U(l/2)U(l/2)$. The product of two antilinear operators is linear, since the second complex conjugation nullifies the effect of the first. Thus, regardless of the linear or antilinear character of $U(l/2)$, it must be the case that $U(l)$ is linear. A continuous operator cannot change discontinuously from linear to antilinear as a function of l, so the operator must be linear for all l.

Transformations of observables

For an observable Q,

$$\langle\phi'|Q|\phi'\rangle = \langle\phi|U^{-1}QU|\phi\rangle$$

If $U(\tau)^{-1}QU(\tau) = \tau Q$, we can prove that

$$U|q\rangle = |\tau q\rangle$$

Here, $|q\rangle$ is the eigenvector of Q with eigenvalue q.

9.3 Schrödinger equation

$U(t, t_0)$ is unitary and $U(t_2, t_0) = U(t_2, t_1)U(t_1, t_0)$. We can define $H(t_0)$ as

$$\frac{d}{dt}U(t, t_0)\bigg|_{t=t_0} = -iH(t_0) \text{ with } H(t_0) = H(t_0)^\dagger$$

We can demonstrate that

$$\frac{dU(t, t_0)}{dt}\bigg|_{t=t_1} = -iH(t_1)U(t_1, t_0)$$

The formal solution of the differential equation is

$$U(t, t_0) = I + (-i)^n \sum_{n=1}^{\infty} \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n H(t_1)H(t_2)\cdots H(t_n)$$

Suppose that T stands for time ordering, placing all operators evaluated at later times to the left, the above equation can be written as

$$U(t, t_0) = I + \frac{(-i)^n}{n!} \sum_{n=1}^{\infty} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n T\{H(t_1)H(t_2)\cdots H(t_n)\} \equiv \exp\left[-iT\left\{\int_{t_0}^{t} H(t')dt'\right\}\right]$$

If the Hamiltonian operator H is time-dependent but the H's at different times commute. The equation above can be simplified to

$$U(t, t_0) = \exp\left[-i\int_{t_0}^{t} H(t')dt'\right]$$

If the H is time-independent, then

$$U(t,t_0) = \exp\left[-iH(t-t_0)\right]$$

Since $|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle$, we can derive the Schrödinger equation

$$\frac{d|\psi(t)\rangle}{dt} = -iH(t)|\psi(t)\rangle$$

The expectation value of an observable Q is $\langle\psi|Q|\psi\rangle$, denoted by $\langle Q\rangle$. We can then derive that

$$\frac{d\langle Q\rangle}{dt} = -i\left\{\langle[Q,H]\rangle + \langle\frac{\partial Q}{\partial t}\rangle\right\}$$

This is called Ehrenfest's theorem.

9.4 Position operators

In three dimensional space, for a particle, we have three operators corresponding to the observations of its position in space, $\mathbf{X} = (X_1, X_2, X_3)$. If the particle has some other internal degrees of freedom, then $\mathbf{X}$ plus some other observables S's will form a complete commuting set of operators. The eigenstate state will be denoted by $|\mathbf{x}, s\rangle$, satisfying that

$$X_i|\mathbf{x}, s\rangle = x_i|\mathbf{x}, s\rangle$$

It describes a particle posited in $\mathbf{x}$ with internal state s. And we will normalize $|\mathbf{x}, s\rangle$ by

$$\langle\mathbf{x}, s'|\mathbf{x}, s\rangle = \delta_{ss'}\delta(\mathbf{x}-\mathbf{x}')$$

9.5 Momentum operators and canonical quantization

Since $\mathbf{X}$ plus some other observables S's form a complete commuting set of operators. So, the momentum operators can not be independent of them. Numerous experiments shows that the position and momentum of particles can not be measured simultaneously. So, we expect $[X,P]\neq 0$.

Guess For a system which has a classical correspondence, the classical equation of motion of a particle is

$$\begin{aligned} \dot{x} &= [x, H_C(x,p,t)]_C \\ \dot{p} &= [p, H_C(x,p,t)]_C \end{aligned}$$

$[\quad]_C$ is the Poisson bracket in classical mechanics. In quantum mechanics,

$$\begin{aligned} \frac{d\langle X\rangle}{dt} &= -i\langle[X,H]\rangle \\ \frac{d\langle P\rangle}{dt} &= -i\langle[P,H]\rangle \end{aligned}$$

If we assume that the classical equation of motion of a particle is an approximation of quantum mechanics, we may expect

$$[\quad] = i[\quad]_C$$

Since the Poisson bracket in classical mechanics and commutation bracket in quantum mechanics have the same algebra structure. To get the right classical equation of motion of the particle, we demand that

$$[X_i, X_j] = 0 \quad [X_i, X_j] = 0 \quad [X_i, P_j] = i\delta_{ij}$$

and

$$H = H_C(X, P, t)$$

For a general system, we formally define momentum operator $\mathbf{P}$ by

$$[X_i, P_j] = i\delta_{ij}$$

The form of H can not be given as a priori, which can be specified only by the hints from classical theory and experiments.

9.6 Momentum operators and translation of states

Theorem 9.2

$$\exp(iG\lambda)A\exp(-iG\lambda) = A + i\lambda[G, A] + \cdots + \frac{i^n\lambda^n}{n!}[G, [G, [G, \cdots [G, A]]]\cdots] + \cdots$$

Define $T(\mathbf{a}) \equiv e^{-i\mathbf{P}\cdot\mathbf{a}}$ We can get

$$T(\mathbf{a})^{-1}\mathbf{X}T(\mathbf{a}) = \mathbf{X} + \mathbf{a}$$

$$T(\mathbf{a})|\mathbf{x}\rangle = |\mathbf{x} + \mathbf{a}\rangle$$

So, $T(\mathbf{a})$ is the space translation operator. Now, we can also define the momentum operator as the generator of space translation.

9.7 Angular momentum operators and rotation of states

We define the angular momentum operators $\mathbf{J}$ as the generator of rotation.

$$R(\theta) \equiv e^{-i\mathbf{J}\cdot\mathbf{n}\theta}$$

If the operator $\mathbf{M} = (M_1, M_2, M_3)$ is a vector in configuration space and can be rotated by R, then we can demonstrate that

$$[J_i, M_j] = i\epsilon_{ijk}M_k$$

Especially,

$$[J_i, J_j] = i\epsilon_{ijk}J_k$$

Orbital angular momentum

Orbital angular momentum of a particle is defined as $\mathbf{L} = \mathbf{X} \times \mathbf{P}$. It is the generator of rotation of the position of the particle, since

$$[L_i, X_j] = i\epsilon_{ijk}X_k \quad [L_i, P_j] = i\epsilon_{ijk}P_k \quad [L_i, L_j] = i\epsilon_{ijk}L_k$$

Spin angular momentum

Experiments show that some microscopic particles possess a property called spin. The state of the spin is denoted by $|s\rangle$. The corresponding operators are $\mathbf{S} = [S_1, S_2, S_3]$, which measure the spin along the $\mathbf{1}, \mathbf{2}, \mathbf{3}$ direction. Spin operator is the generator of rotation of the spin of the particle, so we have

$$[S_i, S_j] = i\epsilon_{ijk}S_k$$

And the rotation of position and spin is independent, so

$$[S_i, L_j] = 0$$

Total angular momentum

The total angular momentum of the particle is

$$\mathbf{J} = \mathbf{L} + \mathbf{S}$$

It is the generator of the rotation of the entire system, which is equivalent to the rotation of the coordinates in opposite direction.

9.8 Heisenberg picture

Define

$$Q_H = U^\dagger(t, t_0)QU(t, t_0)$$

We can derive that

$$\frac{dQ_H(t)}{dt} = -i[Q_H(t), H_H(t)] + \left(\frac{\partial Q}{\partial t}\right)_H$$

Here, $H_H(t) = U^\dagger(t, t_0)H(t)U(t, t_0)$ If the state of the system at t_0 is $|\phi_0\rangle$, then

$$\langle Q\rangle_t = \langle\phi(t)|Q|\phi(t)\rangle = \langle\phi_0|Q_H(t)|\phi_0\rangle$$

If the state $|q\rangle$ is the eigenstate of the Q with the eigenvalue q, then $U^\dagger(t, t_0)|q\rangle$ is the eigenstate of the Q_H with eigenvalue q, which can be denoted by $|q_H(t)\rangle$, so we have

$$\langle q|\phi(t)\rangle = \langle q_H(t)|\phi_0\rangle$$

9.9 Symmetries and conservation laws

Let $U = e^{iKs}$ be a continuous unitary transformation with generator $K = K^\dagger$. To say that the Hamiltonian operator H is invariant under this transformation means that

$$U(s)^{-1}H(t)U(s) = H(t)$$

Then we can deduce that

$$[K, H(t)] = 0$$

Usually, K does not depend on time explicitly. If the above equation hold for all t, then in Heisenberg picture,

$$K_H(t) = K \quad |k_H(t)\rangle = |k\rangle$$

So,

$$\langle K\rangle_t = \langle K\rangle_{t_0} \quad \langle k|\phi(t)\rangle = \langle k|\phi_0\rangle$$

The probability distribution of the measurement of the observable K will not change with time for an arbitrary initial state. We can assume that the K is a constant of motion.

Note: *The concept of a constant of motion should not be confused with the concept of a stationary state. Suppose that the Hamiltonian operator H is independent of t, and that the initial state vector is an eigenvector of H, $|\phi_0\rangle = |E_n\rangle$ with $H|E_n\rangle = E_n|E_n\rangle$. This describes a state having a unique value of energy E_n. So*

$$|\phi(t)\rangle = e^{-iE_nt}|\phi_0\rangle$$

From this result it follows that the average of any dynamical variable R,

$$\langle\phi(t)|R|\phi(t)\rangle = \langle E_n|R|E_n\rangle$$

is independent of t for such a state. By considering functions of R we can further show that the probability distribution is independent of time. In a stationary state the averages and probabilities of all dynamical variables are independent of time, whereas a constant of motion has its average and probabilities independent of time for all states.

Chapter 10
Coordinate and Momentum Representation

10.1 Coordinate Representation

To form a representation of an abstract linear vector space, one chooses a complete orthonormal set of basis vectors $\{|u_i\rangle\}$ and represents an arbitrary vector $|\psi\rangle$ by its expansion coefficients $\{c_i\}$, where $|\psi\rangle = \sum c_i |u_i\rangle$. The array of coefficients $\langle u_i|\psi\rangle$ can be regarded as a column vector (possibly of infinite dimension), provided the basis set is discrete.

Coordinate representation is obtained by choosing as the basis set the eigenvectors $\{|\mathbf{x}\rangle\}$ of the position operator . Since this is a continuous set, the expansion coefficients define a function of a continuous variable,

$$\psi(\mathbf{x}) \equiv \langle \mathbf{x}|\psi\rangle$$

We can show that the inner product of the state vector in coordinate representation is

$$\langle \phi|\psi\rangle = \int \phi^*(\mathbf{x})\psi(\mathbf{x})d^3\mathbf{x}$$

It is a matter of taste whether one says that the set of functions forms a representation of the vector space, or that the vector space consists of the functions $\psi(\mathbf{x})$.

The action of an operator A on the function space is related to its action on the abstract vector space by the rule

$$A\psi(\mathbf{x}) \equiv \langle \mathbf{x}|A|\psi\rangle$$

The action of an position operator in coordinate representation is

$$\mathbf{X}\psi(\mathbf{x}) = \mathbf{x}\psi(\mathbf{x})$$

The action of an momentum operator in coordinate representation is

$$\mathbf{P}\psi(\mathbf{x}) = -i\nabla\psi(\mathbf{x})$$

For a spin-less particle in the scalar potential $W(\mathbf{x})$, $H = \frac{\mathbf{P}^2}{2m} + W(\mathbf{X})$. The equation of motion in the coordinate representation is

$$\left[-\frac{1}{2M}\nabla^2 + W(\mathbf{x})\right]\psi(\mathbf{x},t) = i\frac{\partial}{\partial t}\psi(\mathbf{x},t)$$

10.2 Galilei transformation of Schrödinger equation

For simplicity we shall treat only one spatial dimension. Let us consider two frames of reference: F with coordinates x and t, and F' with coordinates x' and t'. F' is moving uniformly with velocity v relative to F, so that

$$x = x' + vt' \quad t = t'$$

The potential energy is given by $W(x,t)$ in F , and by $W'(x',t')$ in F', with

$$W(x,t) = W'(x',t')$$

Because the requirement of invariance under Galilei transformation, we expect in F' the Schrödinger equation has the form

$$\left[-\frac{1}{2M}\frac{\partial^2}{\partial x'^2} + W'(x')\right]\psi'(x',t') = i\frac{\partial}{\partial t'}\psi'(x',t')$$

where $\psi'(x',t')$ is the wave function in F'. The probability density at a point in space–time must be the same in the two frames of reference

$$|\psi(x,t)|^2 = |\psi'(x',t')|^2$$

and hence we must have

$$\psi(x,t) = e^{if}\psi'(x',t')$$

where f is a real function of the coordinates. Put all the conditions above together, we can derive

$$f(x,t) = Mvx - \frac{1}{2}Mv^2 t$$

apart from an irrelevant constant term.

10.3 Probability flux and conditions on wave functions

Define the probability flux vector

$$\mathbf{J}(\mathbf{x},t) = \frac{1}{M}\mathrm{Im}(\psi^*\nabla\psi)$$

We can get a continuity equation

$$\frac{d}{dt}|\psi(\mathbf{x},t)|^2 + \nabla\cdot\mathbf{J}(\mathbf{x},t)$$

Applying the divergence theorem, we obtain

$$\frac{\partial}{\partial t}\int_\Omega |\psi(\mathbf{x},t)|^2 d^3x = -\oint_\sigma \mathbf{J}\cdot d\mathbf{s}$$

The equations of continuity require that the probability flux $\mathbf{J}(\mathbf{x},t)$ be continuous across any surface, since otherwise the surface would contain sources or sinks. Although this condition applies to all surfaces, implying that $\mathbf{J}(\mathbf{x},t)$ must be everywhere continuous, its practical applications are mainly to surfaces separating regions in which the potential has different analytic forms. Usually, we have the following conditions,

1.
$$\psi(x)|_{x+0} = \psi(x)|_{x-0} \quad \frac{d\psi}{dx}|_{x+0} = \frac{d\psi}{dx}|_{x-0}$$

2.
$$\psi(x)|_{x+0} = \psi(x)|_{x-0} = 0 \quad \frac{d\psi}{dx}|_{x+0} - \frac{d\psi}{dx}|_{x-0} \text{ is finite}$$

Consider next the behavior at a singular point, assumed for convenience to be the origin of coordinates. Let S be a small sphere of radius r surrounding the singularity. The probability that the particle is inside S must be finite. Suppose that $\psi = \frac{u}{r^\alpha}$, where u is a smooth function that does not vanish at $r = 0$. Then we must have $|\psi|^2 r^3$ convergent at the origin, which implies that $\alpha < \frac{3}{2}$.

The net outward flow through the surface S is $F = \oint_S J \cdot dS$. It must vanish in the limit $r \to 0$, since otherwise the origin would be a point source or sink. One has $\frac{\partial \psi}{\partial r} = r^{-\alpha}\frac{\partial u}{\partial r} - \alpha u r^{-\alpha-1}$. The second term does not contribute to the flux, so we obtain

$$F \propto r^{2-2\alpha}$$

where the integration is over solid angle. If the integral does not vanish, then we must have $\alpha < 1$ in order for F to vanish in the limit $r \to 0$. This is a stronger condition than that derived from the probability density.

Since $|\psi|^2$ is a probability density, it must vanish sufficiently rapidly at infinity so that its integral over all configuration space is convergent and equal to 1.

The conditions that we have discussed apply to wave functions $\psi(x)$ which represent physically realizable states, but they need not apply to the eigenfunctions of operators that represent observables. Those eigenfunctions, $\chi(x)$, which play the role of filter functions in computing probabilities, are only required to lie in the extended space, Φ^X, of the rigged-Hilbert-space triplet. It has been suggested that $\psi(x)$ be restricted to the nuclear space Φ, rather than merely to the Hilbert space $\mathcal{H}$. In many cases this would amount to requiring that $\psi(x)$ should vanish at infinity more rapidly than any inverse power of the distance.

10.4 Path integrals

Theorem 10.1 Gaussian integration

$$\int dx e^{-\frac{1}{2}ax^2+Jx} = \left(\frac{2\pi}{a}\right)^{\frac{1}{2}} e^{\frac{J^2}{2a}}$$

The time evolution of a quantum state vector, $|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle$, can be regarded as the propagation of an amplitude in configuration space,

$$\psi(x,t) = \int G(x,t;x',t_0)\psi(x',t_0)dx'$$

where

$$G(x,t;x',t_0) = \langle x,t|U(t,t_0)|x',t_0\rangle$$

is often called the propagator.

Making use of the multiplicative property of the time development operator, it follows that the propagator can be written as

$$G(x,t;x_0,t_0) = \int \cdots \int G(x,t;x_N,t_N)\cdots G(x_1,t_1;x_0,t_0)dx_N\cdots dx_1$$

The N-fold integration is equivalent to a sum over zigzag paths that connect the initial point (x_0,t_0) to the final point (x,t). If we now pass to the limit of $N\to\infty$ and $\Delta t = t_i - t_{i-1}\to 0$, we will have the propagator expressed as a sum (or, rather, as an integral) over all paths that connect the initial point to the final point. We can show that

$$\langle x|e^{-iH\Delta t}|x'\rangle = \sqrt{\frac{M}{2i\pi\Delta t}}\exp\left\{i\left[\frac{M(x-x')^2}{2\Delta t^2} - V(x')\right]\Delta t\right\} \qquad \Delta t\to 0$$

So,

$$G(x,t;x_0,t_0) = \lim_{N\to\infty}\int\cdots\int\left(\frac{M}{2i\pi\Delta t}\right)^{\frac{N+1}{2}}\exp\left\{i\sum_{j=0}^{N}\left[\frac{M(x_{j+1}-x_j)^2}{2\Delta t^2} - V(x_{j+1})\right]\Delta t\right\}dx_1\cdots dx_N$$

The final result can be expressed as

$$G(x,t;x_0,t_0) = \int \mathcal{D}[x(\tau)]e^{iS[x(\tau)]}$$

Here, $S[x(\tau)]$ is the action associated with the path

$$S[x(\tau)] = \int_{x(\tau)} L(x,\dot{x})d\tau$$

The integral is a functional integration over all paths $x(\tau)$ which connect the initial point (x_0,t_0) to the final point (x,t).

To conclude this section, let us generalize our path-integral formula to a more complicated systems. Consider a very general quantum system, described by arbitrary set of of coordinates q_i, conjugate momentum p^i, and Hamiltonian $H(q,p)$. We can show that

$$\langle q_{k+1}|e^{-i\epsilon H}|q_k\rangle = \left(\prod_i\int\frac{dp_k^i}{2\pi}\right)\exp\left[-i\epsilon H\left(\frac{q_{k+1}+q_k}{2},p_k\right)\right]\exp\left[i\sum_i p_k^i(q_{i,k+1}-q_{i,k})\right]$$

so,

$$\langle q_N|U(t,t_0)|q_0\rangle = \left(\prod_{i,k}\int\frac{dp_k^i dq_{i,k}}{2\pi}\right)\exp\left[i\sum_k\left(\sum_i p_k^i(q_{i,k+1}-q_{i,k}) - \epsilon H\left(\frac{q_{k+1}+q_k}{2},p_k\right)\right)\right]$$

There is one momentum integral for each k from 0 to N, and on coordinate integral for each k from 1 to N. The final result can be expressed as

$$\langle q_N|U(t,t_0)|q_0\rangle = \left(\prod_i\int\mathcal{D}q(t)\mathcal{D}p(t)\right)\exp\left[i\int_0^T dt\left(\sum_i p^i\dot{q}_i - H(q,p)\right)\right]$$

where the functions $q(t)$ are constrained at the endpoints, but $p(t)$ are not. The details of this generalization can be found in chapter 9.1 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*

Part V

Quantum Field Theory

Chapter 11
From classical field to quantum field

11.1 Heisenberg picture of fields

The state of the field is described by an element $|\psi\rangle$ in Hilbert space. The measurement of the field is described by an operator field $\phi_a(\vec{x}, t)$. In Heisenberg picture, the dynamic of the field satisfy the equation

$$\frac{d\phi_a(x)}{dt} = -i[\phi_a(x), H]$$

So, the mean value of the measurement of the field is described by Erenfest theorem

$$\frac{d\langle\psi|\phi_a|\psi\rangle}{dt} = -i\langle\psi|[\phi_a, H]|\psi\rangle$$

If $[\phi_a, H]_Q = i[\phi_a, H]_C$, we can reproduce the classical field equation. We also note that the bracket operation here $[A, B] = AB - BA$ has the same properties as the poission bracket in classical mechanics. So, what we need here is the canonical quantization

$$[\phi_a(\vec{x}, t), \phi_b(\vec{y}, t)] = 0 \quad [\pi^a(\vec{x}, t), \pi^b(\vec{y}, t)] = 0 \quad [\phi_a(\vec{x}, t), \pi^b(\vec{y}, t)] = i\delta_a^b\delta(\vec{x} - \vec{y})$$

and the definition of $\mathcal{L}$,π^a and H is the same as those in corresponding classical theory. Then we can recover the classical field theory.

11.2 Lorentz invariance in quantum field theory

$$|\bar{\psi}\rangle = U(\Lambda)|\psi\rangle$$

Scalar fields:

$$\langle\bar{\psi}|\phi(x)|\bar{\psi}\rangle = \langle\psi|\phi(\Lambda^{-1}x)|\psi\rangle$$

$$U^{-1}(\Lambda)\phi(x)U(\Lambda) = \phi(\Lambda^{-1}x)$$

Vector fields:

$$\langle\bar{\psi}|A^\mu(x)|\bar{\psi}\rangle = \langle\psi|\Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x)|\psi\rangle$$

$$U^{-1}(\Lambda)A^\mu(x)U(\Lambda) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x)$$

Lagrangian is a scalar, or more loosely, action is invariant under Lorentz transformation.

11.3 Momentum

The definition of momentum is the same as that in classical theory.

$$T^{\mu\nu} \equiv -\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \partial^\nu \phi_a + \eta^{\mu\nu} \mathcal{L} \quad \partial_\mu T^{\mu\nu} = 0$$

and

$$P^\mu \equiv \int T^{0\mu} d^3x \quad \frac{dP^\mu}{dt} = 0$$

$$P^0 = H, \quad P^i = \int -\pi^a \partial^i \phi_a d^3x$$

And we can get the commutation relationship that

$$\begin{aligned} [\phi_a, P^\mu] &= -i\partial^\mu \phi_a \\ [\pi^a, P^\mu] &= -i\partial^\mu \pi^a \\ [P^\mu, P^\nu] &= 0 \end{aligned}$$

We denote the translation operator as $T(s)$, so

$$T^{-1}(s)\phi_a(x)T(s) = \phi_a(x-s)$$

we can deduce that

$$T(\epsilon) = 1 - i\epsilon_\mu P^\mu \quad T(s) = e^{-iP^\mu s_\mu}$$

11.4 Angular Momentum

The definition of Angular momentum is the same as that in classical theory.

$$M^{\mu\nu\rho} \equiv x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} (\Sigma^{\nu\rho})_a{}^b \phi_b$$

and

$$M^{\nu\rho} \equiv \int M^{0\nu\rho} d^3x \quad \frac{dM^{\nu\rho}}{dt} = 0$$

$$M^{\mu\nu} = \int (x^\mu T^{0\nu} - x^\nu T^{0\mu} - \pi^a (\Sigma^{\mu\nu})_a{}^b \phi_b) d^3x$$

We denote that

$$M_L^{\mu\nu} \equiv \int (x^\mu T^{0\nu} - x^\nu T^{0\mu}) d^3x \quad M_S^{\mu\nu} \equiv \int (-\pi^a (\Sigma^{\mu\nu})_a{}^b \phi_b) d^3x$$

$$(L^{\mu\nu})_a{}^b \equiv -i(x^\mu \partial^\nu - x^\nu \partial^\mu)\delta_a{}^b \quad (S^{\mu\nu})_a{}^b \equiv -i(\Sigma^{\mu\nu})_a{}^b$$

And we have the commutation relationship that

$$M^{\mu\nu} = M_L^{\mu\nu} + M_S^{\mu\nu}$$

$$[\phi_a, M_L^{\mu\nu}] = (L^{\mu\nu})_a{}^b \phi_b \quad [\phi_a, M_S^{\mu\nu}] = (S^{\mu\nu})_a{}^b \phi_b$$

$$[\pi^a, M_L^{\mu\nu}] = (L^{\mu\nu})_b{}^a \pi^b \quad [\pi^a, M_S^{\mu\nu}] = -(S^{\mu\nu})_b{}^a \pi^b$$

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(-g^{\nu\rho}M^{\mu\sigma} + g^{\sigma\mu}M^{\rho\nu} + g^{\mu\rho}M^{\nu\sigma} - g^{\sigma\nu}M^{\rho\mu})$$

We again define $J_i \equiv \frac{1}{2}\epsilon_{ijk}M^{jk}$ and $K_i \equiv M^{i0}$, the communication relationship can be written as

$$\begin{aligned} [J_i, J_j] &= i\epsilon_{ijk}J_k \\ [J_i, K_j] &= i\epsilon_{ijk}K_k \\ [K_i, K_j] &= -i\epsilon_{ijk}J_k \end{aligned}$$

Further more,

$$[P^\mu, M^{\rho\sigma}] = i(g^{\mu\sigma}P^\mu - g^{\mu\rho}P^\sigma)$$

$$\begin{aligned} [J_i, H] &= 0 \\ [J_i, P_j] &= i\epsilon_{ijk}P_k \\ [K_i, H] &= iP_i \\ [K_i, P_j] &= i\delta_{ij}H \end{aligned}$$

At last, we define $L_i \equiv \frac{1}{2}\epsilon_{ijk}M_L^{jk}$ and $S_i \equiv \frac{1}{2}\epsilon_{ijk}M_S^{jk}$. So

$$\begin{aligned} [L_i, S_j] &= 0 \\ [S_i, P_j] &= 0 \\ [L_i, P_j] &= i\epsilon_{ijk}P_k \end{aligned}$$

We denote the rotation operator as $U(\Lambda)$, so

$$U^{-1}(\Lambda)\phi_a(x)U(\Lambda) = S_a{}^b\phi_b(\Lambda^{-1}x)$$

and

$$S_a{}^b = \delta_a{}^b + \frac{i}{2}\delta\omega_{\alpha\beta}(S^{\alpha\beta})_a{}^b$$

we can deduce that

$$U(1+\delta\omega) = 1 + \frac{i}{2}\delta\omega_{\mu\nu}M^{\mu\nu} \quad U(\Lambda) = e^{\frac{i}{2}\theta_{\mu\nu}M^{\mu\nu}}$$

$$U^{-1}(\Lambda)P^\mu U(\Lambda) = \Lambda^\mu{}_\nu P^\nu$$

$$U^{-1}(\Lambda)M^{\mu\nu}U(\Lambda) = \Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma M^{\rho\sigma}$$

Chapter 12
Spin 0 Fields

12.1 Klein-Gordon fields

Lagrangian

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2 + \Omega_0$$

Field equation

$$(\partial^\mu\partial_\mu - m^2)\phi = 0$$

Hamiltonian

$$\pi = \dot{\phi}$$

$$\mathcal{H} = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2 - \Omega_0$$

$$H = \int \mathcal{H} d^3x$$

Momentum and angular momentum

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}(\frac{1}{2}\partial^\mu\phi\partial_\mu\phi + \frac{1}{2}m^2\phi^2 - \Omega_0)$$

$$P^0 = H \quad P^i = \int -\pi\nabla^i\phi d^3x$$

$$J_k = \int -\pi\epsilon_{ijk}x^j\nabla^k\phi d^3x$$

12.2 Canonical quantization Formulation

Canonical quantization

$$\begin{aligned}
[\phi(\vec{x},t),\phi(\vec{y},t)] &= 0 \\
[\pi(\vec{x},t),\pi(\vec{y},t)] &= 0 \\
[\phi(\vec{x},t),\pi(\vec{y},t)] &= i\delta(\vec{x}-\vec{y})
\end{aligned}$$

Fourier expansion

$$\phi(\vec{x},t) = \int \widetilde{dk}\left[a(\vec{k})e^{ikx} + a^\dagger(\vec{k})e^{-ikx}\right]$$

$$\pi(\vec{x},t) = -i\int \widetilde{dk}\omega\left[a(\vec{k})e^{ikx} - a^\dagger(\vec{k})e^{-ikx}\right]$$

Here, $k^2 = \mathbf{k}^2 - \omega^2 = -m^2$, $kx = \mathbf{k}\cdot\mathbf{x} - \omega t$, $\widetilde{dk} = \frac{d^3}{(2\pi)^2 2\omega}$

$$a(\vec{k}) = \int d^3x e^{-ikx}(i\pi + \omega\phi)$$

$$a^\dagger(\vec{k}) = \int d^3x e^{ikx}(-i\pi + \omega\phi)$$

$$\begin{aligned}
[a(\vec{p}),a(\vec{q})] &= 0 \\
\left[a^\dagger(\vec{p}),a^\dagger(\vec{q})\right] &= 0 \\
\left[a(\vec{p}),a^\dagger(\vec{q})\right] &= (2\pi)^3 2\omega\delta(\vec{p}-\vec{q})
\end{aligned}$$

Operator represented by a and $a^\dagger$

$$H = \int \widetilde{dk}\,\omega\, a^\dagger(\vec{k})a(\vec{k}) + (\mathcal{E}_0 - \Omega_0)V \quad \mathcal{E}_0 = \frac{1}{2}(2\pi)^{-3}\int d^3k\,\omega$$

$$P^i = \int \widetilde{dk}\,k^i\,a^\dagger(\vec{k})a(\vec{k})$$

Particles

$$[H,a(\vec{k})] = -\omega a(\vec{k}) \quad [H,a^\dagger(\vec{k})] = \omega a^\dagger(\vec{k})$$

$$[P^i,a(\vec{k})] = -k^i a(\vec{k}) \quad [P^i,a^\dagger(\vec{k})] = k^i a^\dagger(\vec{k})$$

Let $|p\rangle = a^\dagger(\vec{p})|0\rangle$,so

$$H|p\rangle = \omega_p|p\rangle \quad P^i|p\rangle = p^i|p\rangle$$

So, we interpret the state $|\vec{p}\rangle$ as the momentum eigenstate of a single particle of mass m. We can also show that $J_i|\vec{p}=0\rangle = 0$, so the particle carries no internal angular momentum.

Causality

The amplitude for a particle to propagate from y to x is $\langle 0|\phi(x)\phi(y)|0\rangle$, denoted by $D(x-y)$.

$$D(x-y) = \int \widetilde{dk} e^{ik(x-y)}$$

$$[\phi(x), \phi(y)] = D(x-y) - D(y-x)$$

If $x-y$ is space-like, a continuous Lorentz transformation can take $(x-y)$ to $-(x-y)$. So $[\phi(x), \phi(y)] = 0$ for space-like $x-y$. A measurement performed at one point can not affect a measurement at another point whose separation is space-like.

The Klein-Gordon propagator

$$D_R(x-y) \equiv \theta(x^0 - y^0)\langle 0|\phi(x)\phi(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4} \frac{-i}{p^2+m^2} e^{ip(x-y)}$$

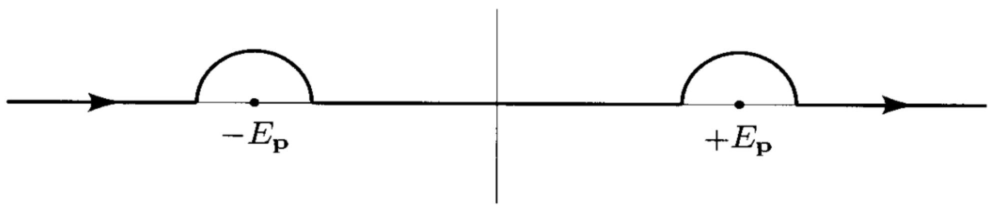


Figure 12.1: Retarded Green Function

$$(\partial^2 - m^2)D_R(x-y) = i\delta(x-y)$$

$$D_F(x-y) \equiv \langle 0|T\phi(x)\phi(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4} \frac{-i}{p^2+m^2-i\epsilon} e^{ip(x-y)}$$

Here, T stands for time ordering, placing all operators evaluated at later times to the left.

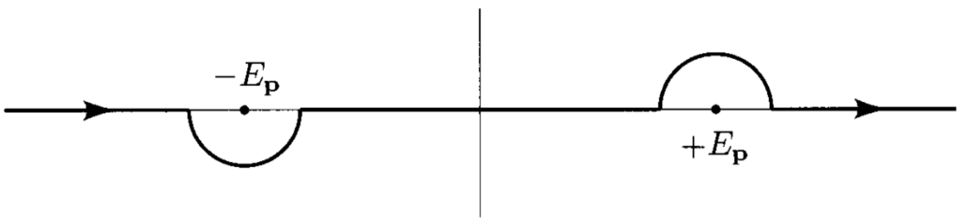


Figure 12.2: Feynman Green Function

12.3 Perturbation theory for canonical quantization

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4$$

$$H = H_0 + H_{int} \quad H_{int} = \int d^3x \frac{\lambda_0^4}{4!}\phi^4(\vec{x})$$

12.3.1 Perturbation expansion of correlation functions

Note: *The ground state of the interaction field theory is denoted by* $|\Omega\rangle$*, the ground state of the free field theory is denoted by* $|0\rangle$*. The zero of energy is defined by* $H_0|0\rangle = 0$ *and* $E_0 = \langle\Omega|H|\Omega\rangle$*.*

$$\phi(t_0,\vec{x}) = \int \frac{d^3p}{(2\pi)^3}(a(\vec{p})e^{i\vec{p}\cdot\vec{x}} + a^\dagger(\vec{p})e^{-i\vec{p}\cdot\vec{x}})$$
$$\phi(t,\vec{x}) = e^{iH(t-t_0)}\phi(t_0,\vec{x})e^{-iH(t-t_0)}$$
$$\phi_I(t,\vec{x}) \equiv e^{iH_0(t-t_0)}\phi(t_0,\vec{x})e^{-iH_0(t-t_0)}$$
$$H_I(x) = \int d^3x\frac{\lambda_0^4}{4!}\phi_I^4$$

The perturbation expansion of correlation functions is

$$\langle\Omega|T\{\phi(x)\phi(y)\}|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\langle 0|T\left\{\phi_I(x)\phi_I(y)\exp\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle}{\langle 0|T\left\{\exp\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle}$$

The proof can be found in chapter 4.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*

Theorem 12.1 Wick's theorem

$$T\{\phi(x_1)\phi(x_2)\cdots\phi(x_n)\} = N\{\phi(x_1)\phi(x_2)\cdots\phi(x_n) + \text{all possible contractions}\}$$

Normal order : all the a's are to the right of all the $a^\dagger$.

The proof can be found in chapter 4.3 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*

Example:

$$\begin{aligned}\langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)\}|0\rangle &= D_F(x_1-x_2)D_F(x_3-x_4)\\ &= D_F(x_1-x_3)D_F(x_2-x_4)\\ &= D_F(x_1-x_4)D_F(x_2-x_3)\end{aligned}$$

12.3.2 Feynman diagram

Expand $\langle 0|T\left\{\phi_I(x)\phi_I(y)\exp\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle$ to the first order of λ_0

$$\begin{aligned}&\langle 0|T\left\{\phi_I(x)\phi_I(y)\frac{-i\lambda_0}{4!}\int dz^4\phi_I(z)\phi_I(z)\phi_I(z)\phi_I(z)\right\}|0\rangle\\ =\;& 3\cdot(\frac{-i\lambda_0}{4!})D_F(x-y)\int d^4zD_F(z-z)D_F(z-z)\\ +\;& 12\cdot(\frac{-i\lambda_0}{4!})\int d^4zD_F(x-z)D_F(y-z)D_F(z-z)\end{aligned}$$

It can be represented by the so called Feynman diagram.

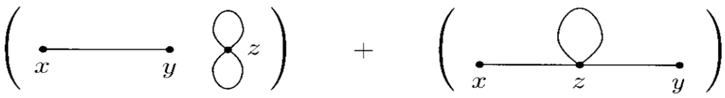


Figure 12.3: Feynman diagram representation of perturbation expansion

The symmetry factor of the first diagram is $S = \frac{4!}{3} = 8$. The symmetry factor of the second diagram is $S = \frac{4!}{12} = 2$. The Feynman rules for ϕ^4 theory are:

1. For each propagator, $P = D_F(x-y)$
2. For each vertex, $V = (-i\lambda_0) \int d^4 z$
3. For each external point, $E = 1$
4. Divided by the symmetry factor

At last, we can prove that

$\langle\Omega|T\{\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n)\}|\Omega\rangle =$ sum of all E-connected diagrams with n external points

Here, the "E-disconnected" means disconnected from all external points", being called "vacuum bubbles". They vacuum bubbles in $\langle 0|T\left\{\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n)\mathrm{exp}\left[-i\int_{-T}^{T} dt H_I\right]\right\}|0\rangle$ are all cancelled by the $\langle 0|T\left\{\mathrm{exp}\left[-i\int_{-T}^{T} dt H_I\right]\right\}|0\rangle$.

12.4 Path integral formulation

12.4.1 Basic formulation

Recall the path integrals formulation in quantum mechanics, we have

$$\langle\phi_b(\vec{x})|e^{-iHT}|\phi_a(\vec{x})\rangle = \int \mathcal{D}\phi\mathcal{D}\pi\mathrm{exp}\left[i\int_0^T d^4x(\pi\dot{\phi} - \frac{1}{2}\pi^2 - \frac{1}{2}(\nabla\phi)^2 - V(\phi))\right]$$

Here, $\langle\phi_b(\vec{x})|$ is the eigenstate of $\phi_S(\vec{x}) = \phi_H(\vec{x},0)$ with eigenvalue $\phi_b(\vec{x})$ at time $t=T$,$|\phi_a(\vec{x})\rangle$ is the eigenstate of $\phi_S(\vec{x})$ with eigenvalue $\phi_a(\vec{x})$ at time $t=0$.
Since the exponential is quadratic in π, we can complete the square and evaluate the $\mathcal{D}(\pi)$ integral to obtain

$$\langle\phi_b(\vec{x})|e^{-iHT}|\phi_a(\vec{x})\rangle = \int \mathcal{D}\phi\mathrm{exp}\left[i\int_0^T d^4x\mathcal{L}\right]$$

Now we can abandon the Hamiltonian formalism and take the the equation above to define the Hamiltonian dynamics.

Note: *We emphasize that in this subsection, ϕ_H denotes the Heisenberg picture of field whose value is operators,ϕ_S denotes the Schrödinger picture of field, $\phi(x)$ is classical field whose value is ordinary number.*

Correlation function

$$\langle\Omega|T\phi_H(x_1)\phi_H(x_2)|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\phi\phi(x_1)\phi(x_2)\text{exp}\left[i\int_T^T d^4x\mathcal{L}\right]}{\int \mathcal{D}\phi\text{exp}\left[i\int_T^T d^4x\mathcal{L}\right]}$$

The proof can be found in chapter 9.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder).*

Functional derivatives and the generating functional

We define the generating functional as

$$Z[J] \equiv \int \mathcal{D}\phi\text{exp}\left[i\int d^4x\mathcal{L} + J(x)\phi(x)\right]$$

We can prove that

$$\langle\Omega|T\phi_H(x_1)\cdots\phi_H(x_n)|\Omega\rangle = \frac{1}{Z_0}\left(-i\frac{\delta}{\delta J(x_1)}\right)\cdots\left(-i\frac{\delta}{\delta J(x_n)}\right)Z[J]|_{J=0}$$

Here, $Z_0 \equiv Z[J=0]$.

12.4.2 Free field theory

In Klein-Gordon field theory,

$$\int d^4x[\mathcal{L}_0(\phi) + J\phi] = \int d^4x[\frac{1}{2}\phi(\partial^2 - m^2 + i\epsilon)\phi + J\phi]$$

Define

$$\phi'(x) \equiv \phi(x) + \int d^4y(-iD_F(x-y))J(y)$$

Recall that $(\partial^2 - m^2)D_F(x-y) = i\delta(x-y)$, we can derive that

$$\int d^4x[\mathcal{L}_0 + J\phi] = \int d^4x[\frac{1}{2}\phi'(\partial^2 - m^2 + i\epsilon)\phi'] - \int d^4xd^4y\frac{1}{2}J(x)[-iD_F(x-y)]J(y)$$

After integration, we can know that

$$Z[J] = Z_0\text{exp}[-\frac{1}{2}\int d^4xd^4yJ(x)D_F(x-y)J(y)]$$

So,

$$\langle 0|T\phi_H(x_1)\phi_H(x_2)|0\rangle = -\frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}\text{exp}[-\frac{1}{2}\int d^4xd^4yJ(x)D_F(x-y)J(y)]|_{J=0} = D_F(x_1-x_2)$$

12.5 Perturbation theory for path integral quantization

$$
\begin{aligned}
\mathcal{L} &= -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4 \\
\mathcal{L} &= \mathcal{L}_0 + \mathcal{L}_1 \quad \mathcal{L}_1 = -\frac{\lambda_0}{4!}\phi^4(\vec{x}) \\
Z[J] &= \int \mathcal{D}\phi e^{i\int d^4x[\mathcal{L}_0+\mathcal{L}_1+J\phi]} \\
&= e^{i\int d^4y\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(y)})}\int \mathcal{D}\phi e^{i\int d^4x[\mathcal{L}_0+J\phi]} \\
&\propto e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)})}\exp[-\frac{1}{2}\int d^4yd^4zJ(y)D_F(y-z)J(z)] \\
&= \sum_{V=0}^{\infty}\frac{1}{V!}[\frac{-i\lambda_0}{4!}\int d^4x(\frac{1}{i}\frac{\delta}{\delta J(x)})^4]^V \times \sum_{P=0}^{\infty}\frac{1}{P!}[-\frac{1}{2}\int d^4yd^4zJ(y)D_F(y-z)J(z)]^P
\end{aligned}
\tag{12.1}
$$

If we focus on a term with particular values of V and P, then the number of surviving sources (after we take all the functional derivatives) is $E = 2P - 4V$. The $4V$ functional derivatives can act on the 2P sources in $\frac{(2P)!}{(2P-4V)!}$ different combinations. However, many of the resulting expressions are algebraically identical.

To organize them, we introduce Feynman diagrams similar to that in perturbation theory of canonical quantization. In these diagrams, a line segment stands for a propagator $D_F(x-y)$, a filled circle at one end of a line segment for a source $i\int d^4xJ(x)$, and a vertex joining four line segments for $-i\lambda_0\int d^4z$.

For each diagram, we can assign a symmetry factor S_P similar to that in perturbation theory for canonical quantization. Due to the fact that some external sources are identical here, usually $S_P \neq S_C$. But when calculating the correlation function, the exchange of the order of functional derivatives to identical sources can eliminate the difference.

We can demonstrate that

$$Z[J] \propto \exp(\sum_I C_I)$$

Here, C_I stands for a particular connected diagram, including its symmetry factor. We define $W[J]$ as

$$Z[J] \equiv Z_0\exp(-iW[J])$$

As, $W[0] = 0$, we know

$$-iW[J] = \sum_{I\neq\{0\}} C_I$$

The notation $I \neq \{0\}$ means that the vacuum diagrams are omitted from the sum.
The detailed discussion can be found in chapter 9 of *Quantum field theory (M. Srednicki)*.

12.6 LSZ reduction formula

12.6.1 Field strength renormalization

The completeness relation:

$$\mathbf{1} = |\Omega\rangle\langle\Omega| + \sum_\lambda \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}|\lambda_{\mathbf{p}}\rangle\langle\lambda_{\mathbf{p}}|$$

Here, $E_{\mathbf{p}} = \sqrt{m_\lambda^2 + \mathbf{p}^2}$

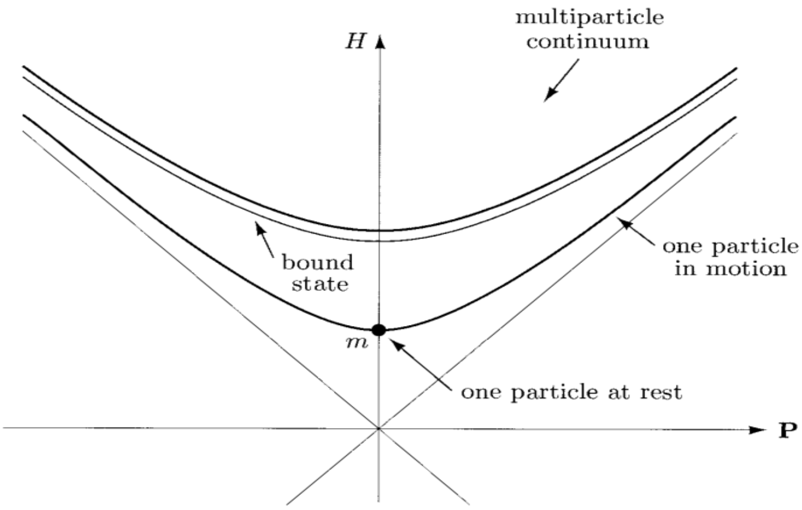


Figure 12.4: Particle's energy-momentum relation

Assume for now $x^0 > y^0$ and define $\langle\Omega|\phi(x)\phi(y)|\Omega\rangle_C = \langle\Omega|\phi(x)\phi(y)|\Omega\rangle - \langle\Omega|\phi(x)|\Omega\rangle\langle|\Omega\phi(y)|\Omega\rangle$ as connected two point function. (The term $\langle\Omega|\phi(x)|\Omega\rangle\langle|\Omega\phi(y)|\Omega\rangle$ is usually zero by symmetry; for higher spin fields, it is zero by Lorentz invariance.) The connected two point function is

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle_C = \sum_\lambda \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}\langle\Omega|\phi(x)|\lambda_{\mathbf{p}}\rangle\langle\lambda_{\mathbf{p}}|\phi(y)|\Omega\rangle$$

It can be verified that

$$\langle\Omega|\phi(x)|\lambda_{\mathbf{p}}\rangle = \langle\langle\Omega|\phi(0)|\lambda_0\rangle e^{ipx}|_{p^0=E_{\mathbf{p}}}$$

So,

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle_C = \sum_\lambda \int \frac{d^4p}{(2\pi)^4}\frac{-i}{p^2+m_\lambda^2-i\epsilon}e^{ip(x-y)}|\langle\Omega|\phi(0)|\lambda_0\rangle|^2$$

Analogous expressions hold for the case $y^0 > x^0$, and both cases can be summarized as

$$\langle\Omega|T\phi(x)\phi(y)|\Omega\rangle_C = \int_0^\infty \frac{dM^2}{2\pi}\rho(M^2)D_F(x-y;M^2)$$

and

$$\rho(M^2) = \sum_\lambda (2\pi)\delta(M^2-m^2)|\langle\Omega|\phi(0)|\lambda_0\rangle|^2$$

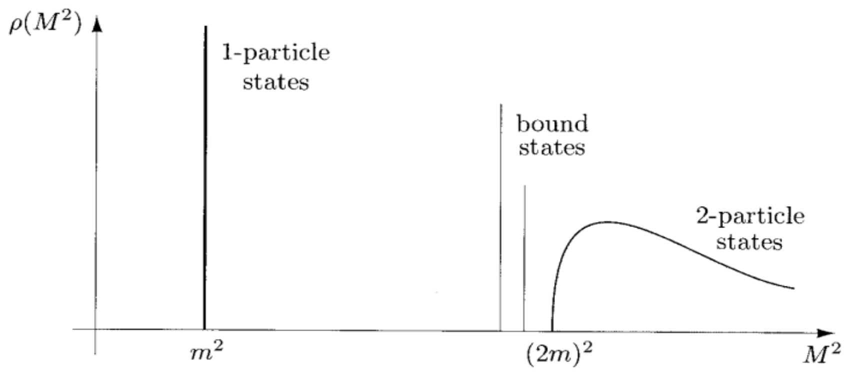


Figure 12.5: The structure of the spectral density function $\rho(M^2)$

The one-particle state contribute an isolated delta function to the spectral density function, so

$$\rho(M^2) = 2\pi\delta(M^2 - m^2) \cdot Z + \text{ (nothing else until } M^2 >\sim (2m)^2)$$

$Z = |\langle\Omega|\phi(0)|\lambda_0\rangle|^2$ is called field-strength renormalization. m is the physical mass of a single particle of the ϕ boson. The Fourier transformation of the two point function is

$$\begin{aligned}
&\int d^4x e^{-ipx}\langle\Omega|T\phi(x)\phi(0)|\Omega\rangle_C \\
= &\int_0^\infty \frac{dM^2}{2\pi}\rho(M^2)\frac{-i}{p^2+M^2-i\epsilon} = \frac{-iZ}{p^2+m^2-i\epsilon} + \int_{\sim 4m^2}^\infty \frac{dM^2}{2\pi}\rho(M^2)\frac{-i}{p^2+M^2-i\epsilon}
\end{aligned}$$

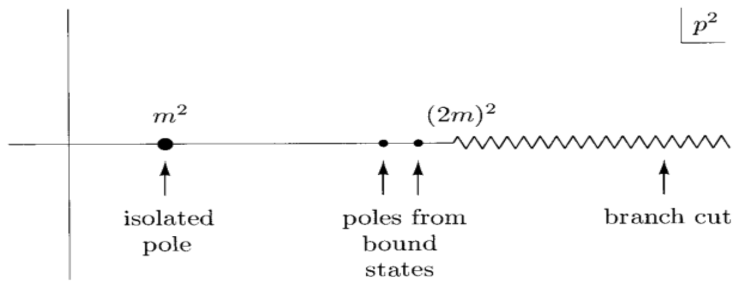


Figure 12.6: The structure of the two point function in Fourier space

12.6.2 LSZ reduction formula

Theorem 12.2 LSZ reduction formula

$$\prod_1^n \int d^4x_i e^{-ip_ix_i} \prod_1^m d^4y_i e^{ik_jy_j} \langle\omega|T\{\phi(x_1)\cdots\phi(x_n)\phi(y_1)\cdots\phi(y_m)\}|\Omega\rangle$$
$$\underset{p_i^0\to E_{\mathbf{p}_i}\, k_i^0\to E_{\mathbf{k}_i}}{\sim} \left(\prod_1^n \frac{-\sqrt{Z}i}{p_i^2+m^2-i\epsilon}\right)\left(\prod_1^m \frac{-\sqrt{Z}i}{k_i^2+m^2-i\epsilon}\right)\langle\mathbf{p}_1\cdots\mathbf{p}_n|S|\mathbf{k}_1\cdots\mathbf{k}_m\rangle$$

The $\sim$ means the two sides of the expression share the same singular structure around $p_i^0 \to E_{\mathbf{p}_i}, k_i^0 \to E_{\mathbf{k}_i}$. The proof can be found in chapter 7.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*. To express the formula above in the language of Feynman diagrams, we consider the S-matrix element for 2-particle $\to$ 2-particle for example. Note the disconnected diagram should be disregarded because they do not have the singularity structure with a product of four poles indicated by on the right side of the LSZ reduction formula. So, the exact four point function

$$\prod_1^2 \int d^4x_i e^{-ip_ix_i} \prod_1^2 d^4y_i e^{ik_jy_j} \langle\omega|T\{\phi(x_1)\phi(x_2)\phi(y_1)\phi(y_2)\}|\Omega\rangle$$

has the general form showed as below.

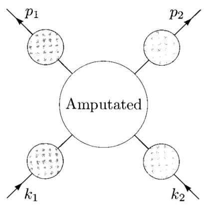


Figure 12.7: Amputated Feynman diagram

Note: *One-particle-irreducible, or 1PI for short, refers to diagrams that is still connected after one line is cut*

We can sum up the corrections to each external leg. Let $-iM^2(p^2)$ denote the sum of all 1PI insertions into the scalar propagator:

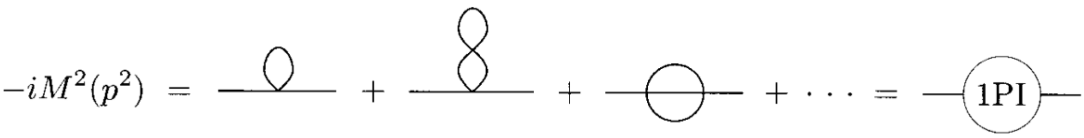


Figure 12.8: Diagram representation of 1PI propagator

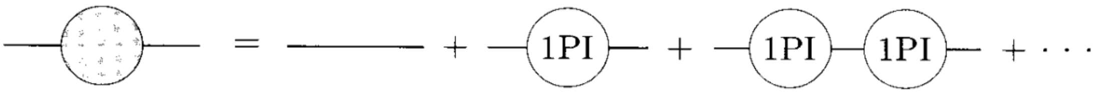


Figure 12.9: Diagram representation of exact propagator

Then the exact propagator can be written as a geometric series in Figure 2.9.
The result is $\frac{-i}{p^2+m_0^2+M^2}$. If we expand each resummed propagator about the physical particle pole, we see that each external leg of the four-point amplitude contributes

$$\frac{-i}{p^2+m_0^2+M^2} \underset{p^0\to E_\mathbf{p}}{\sim} \frac{-iZ}{p^2+m^2} + \text{(regular)}$$

Thus, the sum of diagrams contains a product of four point poles:

$$\frac{-iZ}{p_1^2+m^2}\frac{-iZ}{p_2^2+m^2}\frac{-iZ}{k_1^2+m^2}\frac{-iZ}{k_2^2+m^2}$$

So, the S matrix element can be represented by

$$\langle \mathbf{p}_1\mathbf{p}_2|\, S\, |\mathbf{k}_1\mathbf{k}_2\rangle = \left(\sqrt{Z}\right)^4$$

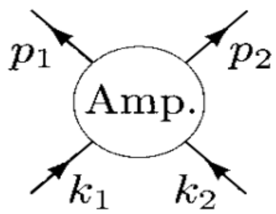


Figure 12.10: Feynman diagram representation of LSZ reduction formula

It is easy to be generalized to the more complicated scattering cases. After Fourier transforming the n-point function to momentum space and cutting off the external legs, the Feynman rules for S-matrix element can be stated as follows:

1. For each propagator, $P = \frac{-i}{p^2+m_0^2-i\epsilon}$;
2. For each vertex, $V = -i\lambda_0$;
3. For each external point, $E = 1$;
4. Impose momentum conservation at each vertex;
5. Integrate over each undetermined loop momentum: $\int \frac{d^4p}{(2\pi)^4}$;
6. Divided by the symmetry factor;
7. Multiply the total momentum conservation factor $(2\pi)^4\delta(\sum p_f - \sum p_i)$

We can write $\langle f|S|i\rangle = i\mathcal{M}(2\pi)^4\delta(\sum p_f - \sum p_i)$ for convenience.

12.7 Re-normalization

Re-normalization, the procedure in quantum field theory by which divergent parts of a calculation, leading to nonsensical infinite results, are absorbed by redefinition into a few measurable quantities, so yielding finite answers.

12.7.1 Counting of ultraviolet divergence

Consider a pure scalar theory in d dimensions with a ϕ^n interaction term

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{n!}\phi^n$$

Let N be the number of external lines in the diagram, P the number of propagators, V the number of vertices. The number of the loops in the diagram is $L = P - V + 1$. There are n lines meeting at each vertex, so $nV = 2P + N$. Loosely speaking, each loop has an integral $d^d p$, each propagator has a factor p^{-2}, so the superficial degrees of divergence is

$$D = dL - 2P = d + [n(\frac{d-2}{2}) - d)]V - (\frac{d-2}{2})N$$

According the superficial degrees of divergence of the diagram. These three possible types of ultraviolet behaviour of quantum field theories. We will refer to them as follows

1. Super-Re-normalizable theory: Only a finite number of Feynman diagrams superficially diverge.
2. Re-normalizable theory: Only a finite number of amplitudes superficially diverge; however, divergences occur at all orders in perturbation theory.
3. Non-Re-normalizable theory: All amplitudes are divergent at a sufficiently high order in perturbation theory.

So, for ϕ^4 theory in four dimension, $D = 4 - N$. It is a re-normalizable theory. For ϕ^3 theory in four dimension, $D = 4 - V - N$. It is a super-re-normalizable theory. For ϕ^6 theory in four dimension, $D = 4 + 2V - N$. It is a Non-re-normalizable theory.

The superficial degrees of freedom can also be derived from dimensional analysis. The dimension of λ is $d - \frac{n(d-2)}{2}$. Now consider an arbitrary diagram with N external lines. One way that such a diagram could arise is from an interaction term $\eta\phi^N$ in the Lagrangian. The dimension of η would then be $d - \frac{N(d-2)}{2}$, and therefore we conclude that any (amputated) diagram with N external lines has dimension $d - \frac{N(d-2)}{2}$. In our theory with only the $\lambda\phi^n$ vertex, if the diagram has V vertices, its divergent part is proportional to $\lambda^V\Lambda^D$, where Λ is a high momentum cut-off and D is the superficial degree of divergence. Applying dimensional analysis, we find

$$d - \frac{N(d-2)}{2} = V[d - \frac{n(d-2)}{2}] + D$$

Note that the quantity that multiplies V in this expression is just the dimension of the coupling constant λ. Thus we can characterize the three degrees of re-normalizability in a second way:

1. Super-Re-normalizable: Coupling constant has positive mass dimension.
2. Re-normalizable: Coupling constant is dimensionless.
3. Non-Re-normalizable: Coupling constant has negative mass dimension.

12.7.2 Renormalized perturbation theory

The Lagrangian of ϕ^4 theory is

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4$$

We write m_0 and λ_0, to emphasize that these are the bare values of the mass and coupling constant, not the values measured in experiments. Since the theory is invariant under $\phi \to -\phi$, all amplitudes with an odd number of external legs vanish. The only divergent amplitudes are therefore

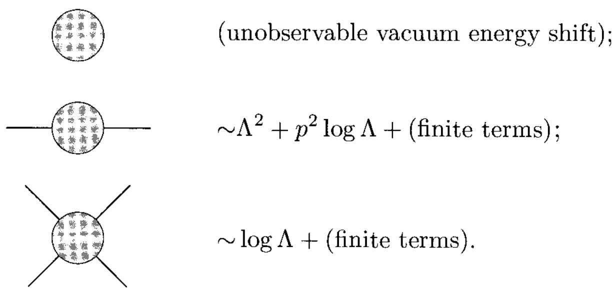


Figure 12.11: Divergence of ϕ^4 theory

Ignoring the vacuum diagram, these amplitudes contain three infinite constants. Our goal is to absorb these constants into the three unobservable parameters of the theory: the bare mass, the bare coupling constant, and the field strength. To accomplish this goal, it is convenient to reformulate the perturbation expansion so that these unobservable quantities do not appear explicitly in the Feynman rules. Recall that the exact two-point function has the form

$$\int d^4x\langle\Omega|\phi(x)\phi(0)|\Omega\rangle e^{-ipx} = \frac{-iZ}{p^2+m^2} + \text{ terms regular at } p^2 = m^2$$

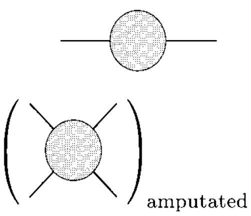
We can eliminate the Z from this equation by rescaling the field: $\phi = Z^{\frac{1}{2}}\phi_r$ We also define

$$\delta_Z = Z - 1 \quad \delta_m = Zm_0^2 - m^2 \quad \delta_\lambda = \lambda_0 Z^2 - \lambda$$

Then the Lagrangian becomes

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi_r\partial_\mu\phi_r - \frac{1}{2}m^2\phi_r^2 - \frac{\lambda}{4!}\phi_r^4 - \frac{1}{2}\delta_Z\partial^\mu\phi_r\partial_\mu\phi_r - \frac{1}{2}\delta_m\phi_r^2 - \frac{\delta\lambda}{4!}\phi_r^4$$

The last three terms, known as counter-terms, have absorbed the infinite but unobservable shifts between the bare parameters and the physical parameters.We give precise definitions of



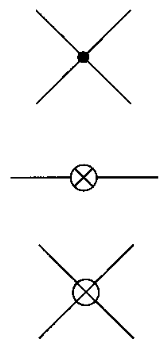
$$= \frac{-i}{p^2 + m^2} + (\text{terms regular at } p^2 = m^2);$$

$$= -i\lambda \qquad \text{at } s = 4m^2,\ t = u = 0.$$

Figure 12.12: Renormalization condition

the physical mass and coupling constant as Figure 2.12.

The re-normalization scheme here is called on-shell (OS) scheme. Other re-normalization scheme would be introduced later. These equations are called re-normalization conditions. Our new Lagrangian gives a new set of Feynman rules as Figure 2.13.



$$= \frac{-i}{p^2 + m^2 + i\epsilon}$$

$$= -i\lambda$$

$$= -i(p^2\delta_Z + \delta_m)$$

$$= -i\delta_\lambda$$

Figure 12.13: Feynman rules for renormalized perturbation theory

We can use these new Feynman rules to compute any amplitude in ϕ^4 theory. The procedure is as follows. Compute the desired amplitude as the sum of all possible diagrams created from the propagator and vertices shown above. The loop integrals in the diagrams will often diverge, so one must introduce a regulator. The result of this computation will be a function of the three unknown parameters δ_Z, δ_m, and δ_λ. Adjust (or "renormalise") these three parameters as necessary to maintain the re-normalization conditions. After this adjustment, the expression for the amplitude should be finite and independent of the regulator.

This procedure, using Feynman rules with counter-terms, is known as renormalized perturbation theory.

Mandelstam variable

In theoretical physics, the **Mandelstam variable** are numerical quantities that encode the energy, momentum, and angles of particles in a scattering process in a Lorentz-invariant fashion. They are used for scattering processes of two particles to two particles. The Mandelstam variables s, t, u are then defined by

$$
\begin{aligned}
s &= -(p_1 + p_2)^2 = -(p_3 + p_4)^2 \\
t &= -(p_1 - p_3)^2 = -(p_2 - p_4)^2 \\
u &= -(p_1 - p_4)^2 = -(p_2 - p_3)^2
\end{aligned}
$$

Where p_1 and p_2 are the four-momenta of the incoming particles and p_3 and p_4 are the four-momenta of the outgoing particles. s is also known as the square of the center-of-mass energy (invariant mass) and t is also known as the square of the four-momentum transfer. We can verify that

$$s + t + u = m_1^2 + m_3^2 + m_3^2 + m_4^2$$

12.7.3 Feynman's formula

Theorem 12.3 Feynman's formula

$$\frac{1}{A_1 \cdots A_n} = \int dF_n (x_1 A_1 + \cdots + x_n A_n)^{-n}$$

where the integration measure over the Feynman parameters x_i is

$$\int dF_n = (n-1)! \int_0^1 dx_1 \cdots dx_n \delta(x_1 + \cdots + x_n - 1)$$

This measure is normalized so that

$$\int dF_n = 1$$

A generalization of Feynman's formula is

$$\frac{1}{A_1^{\alpha_1} \cdots A_n^{\alpha_n}} = \frac{\Gamma(\sum_i \alpha_i)}{\prod_i \Gamma(\alpha_i)} \frac{1}{(n-1)!} \int dF_n \frac{\prod_i x_i^{\alpha_i - 1}}{(\sum_i x_i A_i)^{\sum_i \alpha_i}}$$

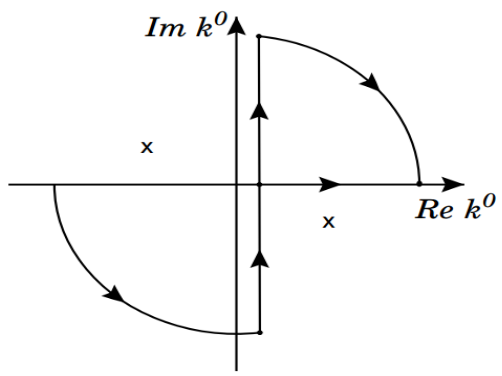


Figure 12.14: Wick rotation

12.7.4 Wick rotation

For an integral $\int d^d q f(q^2 - i\epsilon)$, if the integrand vanishes fast enough as $|q_0| \to \infty$, we can rotate this contour clockwise by $\frac{\pi}{2}$, so that it runs from $-i\infty$ to $i\infty$. In making this Wick rotation, the contour does not pass over any poles. (The $i\epsilon$ are needed to make this statement unambiguous.) Thus the value of the integral is unchanged. It is now convenient to define a Euclidean d-dimensional vector $\bar{q}$ via $q^0 = i\bar{q}_d$ and $q_j = \bar{q}_j$; then $q^2 = \bar{q}^2$, where

$$\bar{q}^2 = \bar{q}_1^2 + \cdots + \bar{q}_d^2$$

Also, $d^d q = i d^d \bar{q}$.Therefore, in general,

$$\int d^d q f(q^2 - i\epsilon) = i \int d^d \bar{q} f(\bar{q}^2)$$

12.7.5 Dimensional regularization

Dimensional regularization is a method for regularizing integrals in the evaluation of Feynman diagrams. For example, if one wishes to evaluate a loop integral which is logarithmically divergent in four dimensions, like

$$\int \frac{d^d p}{(2\pi)^d} \frac{1}{(p^2 + m^2)^2}$$

One first rewrites the integral in some way so that the number of variables integrated over does not depend on d, and then we formally vary the parameter d, to include non-integral values like $d = 4 - \epsilon$.

$$\int_0^\infty \frac{dp}{(2\pi)^{4-\varepsilon}} \frac{2\pi^{(4-\varepsilon)/2}}{\Gamma\left(\frac{4-\varepsilon}{2}\right)} \frac{p^{3-\varepsilon}}{(p^2+m^2)^2} = \frac{2^{\varepsilon-4}\pi^{\frac{\varepsilon}{2}-1}}{\sin(\frac{\pi\varepsilon}{2})\Gamma(1-\frac{\varepsilon}{2})} m^{-\varepsilon} = \frac{1}{8\pi^2\varepsilon} - \frac{1}{16\pi^2}\left(\ln\frac{m^2}{4\pi} + \gamma\right) + \mathcal{O}(\varepsilon)$$

There is a useful formula for calculating the integral

$$\int \frac{d^d \bar{q}}{(2\pi)^d} \frac{(\bar{q}^2)^a}{(\bar{q}^2 + D)^b} = \frac{\Gamma(b - a - \frac{1}{2}d)\Gamma(a + \frac{1}{2}d)}{(4\pi)^{d/2}\Gamma(b)\Gamma(\frac{1}{2}d)} D^{-(b-a-d/2)}$$

If $a = 0$, then the formula will be

$$\int \frac{d^d \bar{q}}{(2\pi)^d} \frac{1}{(\bar{q}^2 + D)^b} = \frac{\Gamma(b - \frac{1}{2}d)}{(4\pi)^{d/2}\Gamma(b)} D^{-(b-d/2)}$$

12.7.6 One loop structure of ϕ^4 theory

First consider the basic two-particle scattering amplitude, If we define $p = p_1 + p_2$, then the second diagram of Figure 2.15 is

$$\frac{(-i\lambda)^2}{2} \int \frac{d^4 k}{(2\pi)^4} \frac{-i}{k^2 + m^2} \frac{-i}{(k+m)^2 + m^2} \equiv (-i\lambda)^2 iV(-p^2)$$

So the entire amplitude is therefore

$$i\mathcal{M} = -i\lambda + (-i\lambda)^2[iV(s) + iV(t)iV(u)] - i\delta_\lambda + \mathcal{O}(\lambda^3)$$

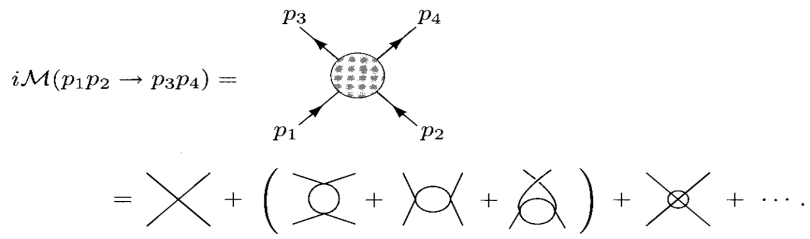


Figure 12.15: Feynman diagram representation of two-particle scattering to one loop

To keep λ dimensionless in dimensional regularization, we can make the transformation $\lambda \rightarrow \lambda\tilde{\mu}^{\epsilon}$. Here, μ is an arbitrary number with mass dimension 1 and $\epsilon \equiv 4 - d$.

We can calculate that

$$V(-p^2) = -\frac{1}{32\pi^2}\int_0^1 (\frac{2}{\epsilon} + \ln(\frac{\mu^2}{D(-p^2)}))$$

where $\mu \equiv \sqrt{4\pi}e^{-\gamma/2}\tilde{\mu}$, $D(-p^2) = x(1-x)p^2 + m^2$

The re-normalization condition implies that

$$\delta_\lambda = -\lambda^2[V(4m^2) + 2V(0)] + \mathcal{O}(\lambda^3)$$

So,

$$i\mathcal{M} = -i\lambda - \frac{i\lambda^2}{32\pi^2}\int_0^1 dx\left[\ln(\frac{D(s)}{D(4m^2)}) + \ln(\frac{D(t)}{D(0)}) + \ln(\frac{D(u)}{D(0)})\right] + \mathcal{O}(\lambda^3)$$

To determine δ_Z and δ_m we must compute the two-point function. Define $-iM(p^2)$ as the sum of all one-particle-irreducible insertions into the propagator.The full two-point function is given by

$$\frac{-i}{p^2 + m^2 + M^2}$$

The re-normalization conditions require that the pole in this full propagator occur at $p^2 = -m^2$ and have residue 1. These two conditions are equivalent, respectively, to

$$M^2(p^2)|_{p^2=-m^2} = 0 \quad \frac{d}{dp^2}M^2(p^2)|_{p^2=-m^2} = 0$$

We can calculate that

$$-iM^2(p^2) = \frac{i\lambda}{32\pi^2}(\frac{2}{\epsilon} + \ln(\frac{\mu^2}{m^2}) + 1)m^2 - i(p^2\delta_Z + \delta_m)$$

So, to the order of λ,

$$\delta_Z = \mathcal{O}(\lambda^2) \quad \delta_m = \frac{\lambda}{32\pi^2}(\frac{2}{\epsilon} + \ln(\frac{\mu^2}{m^2}) + 1)m^2 + \mathcal{O}(\lambda^2) \quad M^2(p^2) = \mathcal{O}(\lambda^2)$$

The detailed calculation can be found in chapter 10.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)* and will be eliminated here.

12.7.7 Modified minimal-subtraction scheme

The Lagrangian of ϕ^4 theory is

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 - \frac{1}{2}\delta_Z\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}\delta_m\phi_r - \frac{\delta\lambda}{4!}\phi^4$$

For minimal-subtraction scheme, we do not demand that m be the physics mass of the field and ϕ create a normalized one-particle state. The physical meaning of λ is not expressed directly as well. Instead we choose δ_Z, δ_m and δ_λ to cancel the infinities, and nothing more; we say that δ_Z, δ_m and δ_λ have no finite parts. It is called the modified minimal-subtraction or $\overline{\text{MS}}$ scheme. ("modified" because we introduced μ via $\lambda \to \lambda\tilde{\mu}^\epsilon$, with $\mu \equiv \sqrt{4\pi}e^{-\gamma/2}\tilde{\mu}$; had we set $\mu = \tilde{\mu}$ instead, the scheme would be just plain minimal subtraction or MS.)

For loop corrections to propagator,

$$\delta_Z = \mathcal{O}(\lambda^2) \quad \delta_m = \left[\frac{\lambda}{16\pi^2} + \mathcal{O}(\lambda^2)\right]\frac{1}{\epsilon}m^2 \quad M^2(p^2) = \frac{\lambda}{32\pi^2}(\ln(\frac{m^2}{\mu^2}) - 1)m^2 + \mathcal{O}(\lambda^2)$$

Firstly, in the $\overline{\text{MS}}$ scheme, the propagator will no longer have a pole at $k^2 = -m^2$. The pole will be somewhere else. However, by definition, the actual physical mass m_{ph} of the particle is determined by the location of this pole: $k^2 = -m^2_{ph}$. Thus, the Lagrangian parameter m is no longer the same as m_{ph}. The relation of m and m_{ph} is

$$m^2_{ph} = M^2(-m^2_{ph}) + m^2$$

To the lowest order,

$$m^2_{ph} = \left[1 + \frac{\lambda}{32\pi^2}(\ln(\frac{m^2}{\mu^2}) - 1)\right]m^2$$

Because m_{ph} is a independent of μ, according to $\frac{d}{d\mu}m_{ph} = 0$, it can be derived that

$$\frac{dm}{d\ln\mu} = \left[\frac{\lambda}{32\pi^2} + \mathcal{O}(\lambda^2)\right]m$$

Furthermore, the residue of this pole is no longer one. Let us call the residue R. So, in the LSZ formula, we get a net factor of $\sqrt{R}$ for each external line when using the $\overline{\text{MS}}$ scheme. And in ϕ^4 theory,

$$R = 1 + \mathcal{O}(\lambda^2)$$

For loop corrections to vertex,

$$\delta_\lambda = \left[\frac{3\lambda^2}{16\pi^2} + \mathcal{O}(\lambda^3)\right]\frac{1}{\epsilon}$$

$$i\mathcal{M} = -i\lambda - \frac{i\lambda^2}{32\pi^2}\int_0^1 dx\left[\ln(\frac{D(s)}{\mu^2}) + \ln(\frac{D(t)}{\mu^2}) + \ln(\frac{D(u)}{\mu^2})\right] + \mathcal{O}(\lambda^3)$$

12.7.8 The re-normalization group

The Lagrangian of ϕ^4 theory is

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\phi_0\partial_\mu\phi_0 - \frac{1}{2}m_0^2\phi_0^2 - \frac{\lambda_0}{4!}\phi_0^4$$

It can be written as

$$\mathcal{L} = -\frac{1}{2}Z_\phi\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}Z_m m^2\phi^2 - Z_\lambda\mu^\epsilon\frac{\lambda}{4!}\phi^4$$

So,

$$\phi_0 = Z_\phi^{1/2}\phi \quad m_0 = Z_\phi^{-1/2}Z_m^{1/2}m \quad \lambda = Z_\phi^{-2}Z_\lambda\lambda\tilde{\mu}^\epsilon$$

After using dimensional regularization, the infinities coming from loop integrals take the form of inverse powers of ϵ. In the $\overline{\text{MS}}$ re-normalization scheme, we choose the Zs to cancel off these powers of $1/\epsilon$, and nothing more. Therefore the Zs can be written as

$$Z_\phi = 1 + \sum_{n=1}^{\infty}\frac{a_n(\lambda)}{\epsilon^n}$$
$$Z_m = 1 + \sum_{n=1}^{\infty}\frac{b_n(\lambda)}{\epsilon^n}$$
$$Z_\lambda = 1 + \sum_{n=1}^{\infty}\frac{c_n(\lambda)}{\epsilon^n}$$

In ϕ^4 theory, $a_1 = \mathcal{O}(\lambda^2)$, $b_1 = \frac{\lambda}{16\pi^2} + \mathcal{O}(\lambda^2)$,$c_1 = \frac{3\lambda}{16\pi^2} + \mathcal{O}(\lambda^2)$
Remember that bare fields and parameters must be independent of μ.Define

$$G(\lambda,\epsilon) \equiv \ln(Z_\phi^{-2}Z_\lambda) = \sum_{n=1}^{\infty}\frac{G_n(\lambda)}{\epsilon^n}$$

We can calculate $G_1 = c_1 - 2a_1 = \frac{3\lambda}{16\pi^2} + \mathcal{O}(\lambda^2)$. As $\ln\lambda_0 = G + \ln\lambda + \epsilon\ln\tilde{\mu}$. From the independence of λ_0, we can derive

$$\left(1 + \frac{\lambda G_1'}{\epsilon} + \cdots\right)\frac{d\lambda}{d\ln\mu} + \epsilon\lambda = 0$$

In a re-normalizable theory, we should have

$$\frac{d\lambda}{d\ln\mu} = -\epsilon\lambda + \beta(\lambda)$$

So

$$\beta(\lambda) = \lambda^2 G_1'(\lambda)$$

In ϕ^4 theory, we have $\beta(\lambda) = \frac{3\lambda^2}{16\pi^2} + \mathcal{O}(\lambda^3)$.Define

$$M(\lambda,\epsilon) \equiv \ln(Z_m^{1/2}Z_\phi^{-1/2}) = \sum_{n=1}^{\infty}\frac{M_n(\lambda)}{\epsilon^n}$$

We can calculate $M_1 = \frac{1}{2}b_1 - \frac{1}{2}a_1 = \frac{\lambda}{32\pi^2} + \mathcal{O}(\lambda^2)$. As $\ln m_0 = M + \ln m$, define the anomalous dimension of the mass

$$\gamma_m(\lambda) \equiv \frac{1}{m}\frac{dm}{d\ln\mu}$$

From the independence of m_0, we can derive

$$\gamma_m(\lambda) = \lambda M_1'$$

In ϕ^4 theory, we have $\gamma_m(\lambda) = \frac{\lambda}{32\pi^2} + \mathcal{O}(\lambda^2)$. We can expand $\ln Z_\phi$ as

$$\ln Z_\phi = \frac{a_1}{\epsilon} + \frac{a_2 - \frac{1}{2}a_1^2}{\epsilon^2}$$

Define the anomalous dimension of the field

$$\gamma_\phi(\lambda) = \frac{1}{2}\frac{d\ln Z_\phi}{d\ln\mu}$$

We can derive

$$\gamma_\phi(\lambda) = -\frac{1}{2}\lambda a_1'$$

In ϕ^4 theory, we have $\gamma_m(\lambda) = \mathcal{O}(\lambda^2)$.

Callen-Symanzik equation

$$G^{(n)}(x_1, \cdots, x_n) \equiv \langle\Omega|T\phi(x_1)\cdots\phi(x_n)|\Omega\rangle_C$$

As $G_0^{(n)} = Z_\phi^{n/2}G^{(n)}$, from the independence of bare Green's function, we have

$$\left(\frac{\partial}{\partial\ln\mu} + \beta(\lambda)\frac{\partial}{\partial\lambda} + \gamma_m(\lambda)m\frac{\partial}{\partial m} + n\gamma_\phi(\lambda)\right)G^n(x_1, \cdots, x_n; \lambda, m, \mu) = 0$$

12.8 Spontaneous symmetry breaking

12.8.1 Effective action

$$Z[J] = e^{-iE[J]} = \int \mathcal{D}\phi \exp\left[i\int d^4x(\mathcal{L}[\phi] + J\phi)\right]$$

Define

$$\phi_{\rm cl}(x) \equiv \langle\Omega|\phi(x)|\Omega\rangle_J$$

So, we can derive

$$\frac{\delta}{\delta J(x)}E[J] = -\phi_{\rm cl}(x)$$

Define

$$\Gamma[\phi_{\rm cl}] \equiv -E[J] - \int d^4y J(y)\phi_{\rm cl}(y)$$

We can verify that

$$\frac{\delta}{\delta\phi_{\text{cl}}(x)}\Gamma[\phi_{\text{cl}}] = -J(x)$$

If the external source is set to zero, the effective action satisfy the equation

$$\frac{\delta}{\delta\phi_{\text{cl}}(x)}\Gamma[\phi_{\text{cl}}] = 0$$

The solution to this equation are the values of $\langle\phi(x)\rangle$ in the stable quantum states of the theory. For a translational-invariant vacuum state, we will find a solution in which ϕ_{cl} is independent of x. For the field theory that the possible vacuum states are invariant under translations and Lorentz transformations, for each possible vacuum states, the corresponding solution ϕ_{cl} will be a constant, independent of x. If T is the time extent of the region and V is its three dimensional volume, we can write

$$\Gamma[\phi_{\text{cl}}] = -(VT)\cdot V_{\text{eff}}(\phi_{\text{cl}})$$

The coefficient V_{eff} is called effective potential. The condition that $\Gamma[\phi_{\text{cl}}]$ has an extreme then reduces to the simple equation

$$\frac{\partial}{\partial\phi_{\text{cl}}}V_{\text{eff}}(\phi_{\text{cl}}) = 0$$

A system with spontaneously broken symmetry will have several minimum of V_{eff}, all with the same energy by virtue of the symmetry. The choice of one among these vacuum is the spontaneous symmetry breaking.

12.8.2 Computation of the effective action

Decompose the Lagrangian into a piece depending on renormalized parameters and one containing the counter-terms

$$\mathcal{L} = \mathcal{L}_1 + \delta\mathcal{L}$$

Define J_1 by

$$\frac{\delta\mathcal{L}_1}{\delta\phi}|_{\phi=\phi_{\text{cl}}} + J_1(x) = 0$$

Define δJ by

$$J(x) = J_1(x) + \delta J(x)$$

So, we have

$$e^{-iE[J]} = \int \mathcal{D}\phi e^{i\int d^4x(\mathcal{L}_1+J_1\phi)}e^{i\int d^4x(\delta\mathcal{L}+\delta J\phi)}$$

Replace ϕ by $\phi_{\text{cl}} + \eta$,

$$\begin{aligned}\int d^4x\,(\mathcal{L}_1 + J_1\phi) &= \int d^4x\,(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}}) + \int d^4x\,\eta(x)\left(\frac{\delta\mathcal{L}_1}{\delta\phi} + J_1\right) \\ &+ \frac{1}{2}\int d^4x\,d^4y\,\eta(x)\eta(y)\frac{\delta^2\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)} \\ &+ \frac{1}{3!}\int d^4x\,d^4y\,d^4z\,\eta(x)\eta(y)\eta(z)\frac{\delta^3\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)\delta\phi(z)} + \cdots\end{aligned}$$

The term linear in η vanishes by definition of J_1. Keeping only the term up to quadratic order in η and still neglecting the counter-terms, we have a pure Gaussian integral, which can be evaluated in terms of a functional determinant:

$$\begin{aligned}&\int \mathcal{D}\eta \exp\left[i\left(\int(\mathcal{L}_1[\phi_{\rm cl}]+J_1\phi_{\rm cl})+\frac{1}{2}\int \eta\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\eta\right)\right]\\ =\ &\exp\left[i\int(\mathcal{L}_1[\phi_{\rm cl}]+J_1\phi_{\rm cl})\right]\left(\det\left[\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right]\right)^{-\frac{1}{2}}\end{aligned}$$

Finally, put back the effects of the counter-term Lagrangian,writing it as

$$(\delta\mathcal{L}[\phi_{\rm cl}]+\delta J\phi_{\rm cl})+(\delta\mathcal{L}[\phi_{\rm cl}+\eta]-\delta\mathcal{L}[\phi_{\rm cl}]+\delta J\eta)$$

Define

$$\mathcal{L}_2=\left(\frac{1}{3!}\int d^4x\,d^4y\,d^4z\,\eta(x)\eta(y)\eta(z)\frac{\delta^3\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)\delta\phi(z)}+\cdots\right)+(\delta\mathcal{L}[\phi_{\rm cl}+\eta]-\delta\mathcal{L}[\phi_{\rm cl}]+\delta J\eta)$$

So

$$e^{-iE[J]}=\exp\left[i\int(\mathcal{L}_1[\phi_{\rm cl}]+J_1\phi_{\rm cl}+\delta\mathcal{L}[\phi_{\rm cl}]+\delta J\phi_{\rm cl})\right]e^{i\int\mathcal{L}_2(\frac{1}{i}\frac{\delta}{\delta I})}\int\mathcal{D}\eta e^{i\int\left(\frac{1}{2}\eta\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\eta+I\eta\right)}$$

Therefore, define propagator as

$$D_F\equiv i\left(\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right)^{-1}$$

We have

$$e^{-iE[J]}=\exp\left[i\int(\mathcal{L}_1[\phi_{\rm cl}]+J_1\phi_{\rm cl}+\delta\mathcal{L}[\phi_{\rm cl}]+\delta J\phi_{\rm cl})\right]\left(\det\left[\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right]\right)^{-\frac{1}{2}}e^{i\int\mathcal{L}_2(\frac{1}{i}\frac{\delta}{\delta I})}\int\mathcal{D}\eta e^{i\int(-\frac{1}{2}ID_FI)}|_{I=0}$$

Similar to the procedure in the perturbation theory for path integral quantization, we can get a perturbation expansion for $iE[J]$ using connected Feynman diagram,

$$-iE[J]=i\int(\mathcal{L}_1[\phi_{\rm cl}]+J_1\phi_{\rm cl}+\delta\mathcal{L}[\phi_{\rm cl}]+\delta J\phi_{\rm cl})-\frac{1}{2}\log\det\left[\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right]+\text{connected diagrams}$$

From this equation, Γ follows directly:

$$\Gamma[\phi_{\rm cl}]=\int d^4x\mathcal{L}_1[\phi_{\rm cl}]+\frac{i}{2}\log\det\left[\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right]-i\,\text{connected diagrams}+\int d^4x\delta\mathcal{L}[\phi_{\rm cl}]$$

Notice that there are no terms remaining that depend explicitly on J; thus, Γ is expressed as a function of $\phi_{\rm cl}$, as it should be. The Feynman diagrams contributing to $\Gamma[\phi_{\rm cl}]$ have no external lines, and the simplest ones turn out to have two loops. The lowest-order quantum correction to Γ is given by the functional determinant.

The last term provides a set of counter-terms that can be used to satisfy the re-normalization conditions on Γ and, in the process, to cancel divergences that appear in the evaluation of the functional determinant and the diagrams. The re-normalization conditions will determine all of the counter-terms in $\delta\mathcal{L}$. However, the formalism we have constructed contains a new counter-term δJ. That coefficient is determined by $\langle\eta\rangle=0$. In practice, we will satisfy this condition by simply ignoring any one-particle-irreducible one-point diagram, since any such diagram will be cancelled by adjustment of δJ.

12.8.3 The effective action as a generating functional

$E[J]$ is called the generating of connected correlation functions,

$$\frac{\delta^n E[J]}{\delta J(x_1)\cdots\delta J(x_n)} = i^{n+1}\langle\phi(x_1)\cdots\phi(x_n)\rangle\text{conn}$$

The effective action $\Gamma[\phi_{\text{cl}}]$ is the generating functional of one-particle-irreducible correlation functional,

$$\frac{\delta\Gamma[\phi_{\text{cl}}]}{\delta\phi_{\text{cl}}(x)} = 0$$

$$\frac{\delta^2\Gamma[\phi_{\text{cl}}]}{\delta\phi_{\text{cl}}(x)\delta\phi_{\text{cl}}(y)} = iD^{-1}(x,y)$$

Here, $D(x,y) = \langle\phi(x)\phi(y)\rangle_{\text{conn}}$. When $n \geq 3$,

$$\frac{\delta^n\Gamma[\phi_{\text{cl}}]}{\delta\phi_{\text{cl}}(x_1)\cdots\delta\phi_{\text{cl}}(x_n)} = -i\langle\phi(x_1)\cdots\phi(x_n)\rangle_{\text{1PI}}$$

The proof of statements above can be found in chapter 10.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*

The chapter 21 of *Quantum field theory (M. Srednicki)* gives an constructive way to define the effective action.

$$\Gamma[\phi] \equiv \frac{1}{2}\int\frac{d^dk}{(2\pi)^d}\tilde{\phi}(-k)(-k^2-m^2-M^2(k^2))\tilde{\phi}(k)+\frac{1}{n!}\int\frac{d^dk_1}{(2\pi)^d}\cdots\frac{d^dk_n}{(2\pi)^d}(2\pi)^d\delta(k_1+\cdots+k_n)V_n(k_1,\cdots,k$$

Here $\tilde{\phi}(k) = \int d^dxe^{-ikx}\phi(x)$, and $iV_n(k_1,\cdots,k_n)$ equals the value of 1PI Feynman diagram in momentum space. The effective action has the property that the tree-level Feynman diagrams it generates give the complete scattering amplitude of the original theory. The author also proved that this definition is equivalent to the definition from *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*.

12.8.4 Re-normalization and symmetry

Consider first the computation of the effective potential for constant classical fields, in a field theory with an arbitrary number of fields ϕ^i. The effective potential has mass dimension 4, so we expect that $V_{\text{eff}}(\phi_{\text{cl}})$ will have divergent terms up to Λ^4. To understand these divergences, expand $V_{\text{eff}}(\phi_{\text{cl}})$ in a Taylor series:

$$V_{\text{eff}}(\phi_{\text{cl}}) = A_0 + A_2^{ij}\phi_{\text{cl}}^i\phi_{\text{cl}}^j + A_4^{ijkl}\phi_{\text{cl}}^i\phi_{\text{cl}}^j\phi_{\text{cl}}^k\phi_{\text{cl}}^l + \cdots$$

In theories without a symmetry of $\phi \to \phi$, there might also be terms linear and cubic in ϕ^i; we omit these for simplicity. The coefficients A_0, A_2, A_4 have mass dimension, respectively, 4, 2, and 0; thus we expect them to contain Λ^4, Λ^2, and $\log\Lambda$ divergences, respectively. The power-counting analysis predicts that all higher terms in the Taylor series expansion should be finite.

The constant term A_0 is independent of $\phi_{\rm cl}$; it has no physical significance. However, the divergences in A_2 and A_4 appear in physical quantities, since these coefficients enter the inverse propagator and the irreducible four-point function and therefore appear in the computation of S-matrix elements. There is one further coefficient in the effective action that has non-negative mass dimension by power counting; this is the coefficient of the term quadratic in $\partial_\mu \phi_{\rm cl}$, which appears when the effective action is evaluated for a non-constant background field:

$$\Delta\Gamma[\phi_{\rm cl}] = \int d^4x B_2^{ij} \partial_\mu \phi^i_{\rm cl} \partial^\mu \phi^j_{\rm cl}$$

All other coefficients in the Taylor expansion of the effective action in powers of $\phi_{\rm cl}$ are finite by power counting.

We can now argue that the counter-terms of the original Lagrangian suffice to remove the divergences that might appear in the computation of $\Gamma[\phi_{\rm cl}]$. The argument proceeds in two steps. We first use the BPHZ theorem to argue that the divergences of Green's functions can be removed by adjusting a set of counter-terms corresponding to the possible operators that can be added to the Lagrangian with coefficients of mass dimension greater than or equal to zero. The coefficients of these counter-terms are in 1-to-1 correspondence with the coefficients A_2, A_4, and B_2 of the effective action. Next, we use the fact that the effective action is manifestly invariant to the original Symmetry group of the model. This is true even if the vacuum state of the model has spontaneous symmetry breaking, since the method we presented for computing the effective action is manifestly invariant to the original symmetry of the Lagrangian. Combining these two results, we conclude that the effective action can always be made finite by adjusting the set of counter-terms that are invariant to the original symmetry of the theory, even if this symmetry is spontaneously broken.

12.8.5 Goldstone's theorem

Theorem 12.4 Goldstone's theorem

Goldstone's theorem examines a generic continuous symmetry which is spontaneously broken; i.e., its currents are conserved, but the ground state is not invariant under the action of the corresponding charges. Then, necessarily, new massless (or light, if the symmetry is not exact) scalar particles appear in the spectrum of possible excitations. There is one scalar particle - called a Nambu-Goldstone boson — for each generator of the symmetry that is broken, i.e., that does not preserve the ground state.

A general continuous symmetry transformation has the form

$$\phi^a \to \phi^a + \alpha \Delta^a(\phi)$$

where α is an infinitesimal parameter and Δ^a is some function of all the ϕ's. Specialize to constant fields; then the derivative terms in $\mathcal{L}$ vanish and the potential alone must be invariant.

This condition can be written

$$V(\phi^a) = V(\phi^a + \alpha\Delta^a(\phi)) \quad \text{or} \quad \Delta^a(\phi)\frac{\partial}{\partial\phi^a}V(\phi) = 0$$

The effective potential V_{eff} encapsulates the full solution to the theory, including all orders of quantum corrections. At the same time, it satisfies the general properties of the classical potential: It is invariant to the symmetries of the theory, and its minimum gives the vacuum expectation value of ϕ_{cl}. So

$$\Delta^a(\phi)\frac{\partial}{\partial\phi^a}V_{\text{eff}}(\phi) = 0$$

Now differentiate with respect to ϕ^b, and set $\phi = \phi_{\text{cl}}$

$$0 = \left(\frac{\partial\Delta^a}{\partial\phi^b}\right)_{\phi_{\text{cl}}}\left(\frac{\partial V_{\text{eff}}}{\partial\phi^a}\right)_{\phi_{\text{cl}}} + \Delta^a(\phi_{\text{cl}})\left(\frac{\partial^2}{\partial\phi^a\partial\phi^b}V_{\text{eff}}\right)_{\phi_{\text{cl}}}$$

The first term vanishes since ϕ_{cl} is a minimum of V_{eff}, so the second term must also vanish. If the transformation leaves ϕ_{cl} unchanged (i.e., if the symmetry is respected by the ground state), then $\Delta^a(\phi_{\text{cl}}) = 0$ and this relation is trivial. A spontaneously broken symmetry is precisely one for which $\Delta^a(\phi_{\text{cl}}) \neq 0$; in this case $\Delta^a(\phi_{\text{cl}})$ is the vector with eigenvalue zero.

We now argue that the presence of such a zero eigenvalue implies the existence of a massless scalar particle. Effective action's second functional derivative is the inverse propagator

$$i\tilde{D}^{-1}_{ij}(p^2) = \int d^4x e^{-ip(x-y)}\frac{\delta\Gamma}{\delta\phi^i\delta\phi^j}(x,y)$$

A particle of mass m corresponds to a zero eigenvalue of this matrix equation at $p^2 = -m^2$. Now set $p = 0$. This implies that we differentiate $\Gamma[\phi_{\text{cl}}]$ with respect to constant fields. Thus, we can replace $\Gamma[\phi_{\text{cl}}]$ by its value with constant classical fields, which is just the effective potential. We find that the quantum field theory contains a scalar particle of zero mass when the matrix of second derivatives,

$$\frac{\partial^2}{\partial\phi^i_{\text{cl}}\partial\phi^j_{\text{cl}}}V_{\text{eff}}$$

has a zero eigenvalue. This completes the proof of Goldstone's theorem.

12.9 Linear sigma model

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi^i\partial^\mu\phi^i + \frac{1}{2}\mu^2(\phi^i)^2 - \frac{\lambda}{4}[(\phi^i)]$$

The Lagrangian is invariant under the symmetry

$$\phi^i \to R^{ij}\phi^j$$

for any $N \times N$ orthogonal group in N dimensions, also called the N-dimensional orthogonal group or simply $O(N)$.

12.10 Cross section and the S-matrix