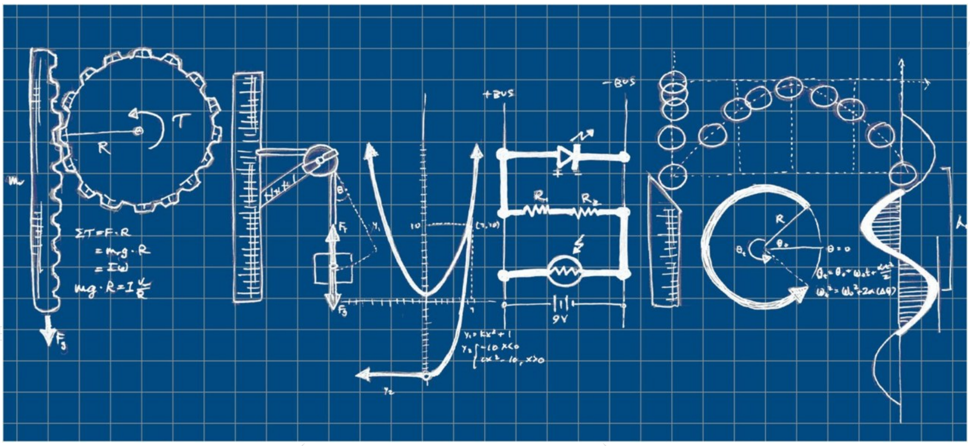
Physics

物理



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

Editor: Yuyang Songsheng

Date: November 8, 2016

Email: songshengyuyang@gmail.com

Version: 1.00

Contents

Part I

Classical Mechanics

Chapter 1
The formulation of Classical Mechanics

1.1 Lagrangian Formulation

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

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

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:

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

   The equation of motion is

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

   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 $\boldsymbol{V}(t)$ and rotating with angular velocity $\boldsymbol{\Omega}$ relative to the inertial reference frame. We use the

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

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

The equation of motion will be

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

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 $\delta \boldsymbol{r}_i \to \boldsymbol{r}_i + \delta\boldsymbol{a}$, so

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

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

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

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,

$$\boldsymbol{p}_i = m_i \boldsymbol{v}_i$$

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

$$\dot{\boldsymbol{p}}_i = -\nabla_{\boldsymbol{r}_i} U \quad \dot{r}_i = \frac{\boldsymbol{p}_i}{m_i}$$

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.
$$\frac{d}{dt}(F + \sum_i P^i Q_i - \sum_i p^i q_i) = -\sum_i q_i \dot{p}^i + \sum_i Q_i \dot{P}^i + (H' - H)$$

Assume $\Phi(p^i, P^i, t) = F + \sum_i P^i Q_i - \sum_i p^i q_i$, 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

$$dS = p_t dq_t + p_{t+\tau} dq_{t+\tau} - (H_{t+\tau} - H_t)dt$$

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

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

The invariant symbol of the vector representation of Lorentz transformation is $\eta^{\mu\nu}$

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

where,

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

The inverse matrix of $\eta^{\mu\nu}$ is

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

We can use $\eta^{\mu\nu}$ and and its inverse $\eta^{\mu\nu}$ to raise and lower vector indices,

$$x_\mu \equiv \eta_{\mu\nu} x^\nu$$

And we can verify the following equations,

$$\Lambda^\rho{}_\mu \Lambda_\rho{}^\nu = \delta^\nu_\mu$$

$$x^\mu = \eta^{\mu\nu} x_\nu$$

$$\bar{x}_\mu = \Lambda_\mu{}^\nu x_\nu$$

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

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

$$\bar{x}^\mu = x^\mu + x_\nu \delta\Omega^{\mu\nu}$$

$$\delta S = \sum m u^\mu x^\nu \delta\Omega_{\mu\nu}\big|_a^b = 0$$

$\sum M^{\mu\nu}$ 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 $\phi(\boldsymbol{x}, t)$ directly to $\phi(\boldsymbol{y}, t)$ with $\boldsymbol{x} \neq \boldsymbol{y}$. The closet we get for the $\boldsymbol{x}$ label is coupling between $\phi(\boldsymbol{x}, t)$ and $\phi(\boldsymbol{x} + \delta\boldsymbol{x}, t)$ 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 $\delta\mathcal{L} = \partial_\mu K^\mu$, 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 $P^\mu \equiv \int T^{0\mu}d^3x$,then we have the law of momentum conservation:

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

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

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

In the limit of infinitesimal Lorentz transformation, we have

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

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

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

If we define $M^{\nu\rho} \equiv \int M^{0\nu\rho} d^3x$, then we have the law of angular momentum conservation:

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

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

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

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

$$\frac{\delta F}{\delta \rho}(y) = \lim_{\epsilon \to \infty} \frac{1}{\epsilon}\{F[\rho(x) + \epsilon\delta(x-y)] - F[\rho(x)]\}$$
$$\frac{\delta f(x)}{\delta f(y)} = \delta(x-y)$$
$$\frac{\delta}{\delta f(y)} \int g(f(x))\, dx = g'(f(y))$$
$$\frac{\delta f'(x)}{\delta f(y)} = \frac{d}{dx}\delta(x-y)$$
$$\frac{\delta}{\delta f(y)} \int g(f'(x))\, dx = -\frac{d}{dy} g'(f'(y))$$

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 $\delta S = 0$,

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

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

3.4.1 Poission bracket

First, we define that

$$[\phi_a(\boldsymbol{x}), \phi_b(\boldsymbol{y})] \equiv [\pi^a(\boldsymbol{x}), \phi_b(\boldsymbol{y})] \equiv 0$$
$$[\phi_a(\boldsymbol{x}), \pi^b(\boldsymbol{y})] \equiv \delta_a^b \delta(\boldsymbol{x} - \boldsymbol{y})$$

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

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

and

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

As a result, we can verify that

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

Especially,

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

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 that

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

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 Poisson bracket

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

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 define that

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

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

So, we can get the Poisson bracket

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

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

Because $\frac{dM^{\mu\nu}}{dt} = 0$, $M^{\mu\nu}$ can commutate with $\frac{d}{dt}$, so

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

At last, we can get the Poisson bracket

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

Since we can prove that

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

if we demand that

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

We will get get the Poisson bracket of the $M^{\mu\nu}$,

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

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 we have

$$M^{\mu\nu} = \begin{bmatrix} 0 & -K_1 & -K_2 & -K_3 \\ K_1 & 0 & J_3 & -J_2 \\ K_2 & -J_3 & 0 & J_1 \\ K_3 & J_2 & -J_1 & 0 \end{bmatrix} \quad \left(\delta\omega_{\mu\nu} = \begin{bmatrix} 0 & -\beta_1 & -\beta_2 & -\beta_3 \\ \beta_1 & 0 & -\theta_3 & \theta_2 \\ \beta_2 & \theta_3 & 0 & -\theta_1 \\ \beta_3 & -\theta_2 & \theta_1 & 0 \end{bmatrix} \right)$$

the Poisson bracket 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

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

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

$$V^r_s = V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$$

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

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

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

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

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) $\forall f \in C^\infty(M), \alpha \in \boldsymbol{R}, v(\alpha f) = \alpha \cdot v(f)$;
(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 $\phi : \boldsymbol{R} \times M \to M$ a smooth mapping, and $\forall (t, p) \in \boldsymbol{R} \times M$,denote $\phi_t(p) = \phi(t, p)$. If ϕ satisfy that
(1) $\phi_0 = \mathrm{id} : M \to M$;
(2) $\forall s, t \in \boldsymbol{R}, \phi_s \circ \phi_t = \phi_{s+t}$;
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 $X_1, \dots, X_n$ such that for any point $y \in N_x$, $X_1(y), \dots, X_n(y)$ 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 $\{X_1, \dots, X_n\}$ 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 $\Delta|_V = \mathrm{Span}\{\frac{\partial}{\partial y^1}, \cdots, \frac{\partial}{\partial y^h}\}$.

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

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

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

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

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

$f_i^\alpha(x, y)$ 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.

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

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

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

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

♡

Definition 4.28 Integral of differential form with compact support

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

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

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

♡

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 $R_{ijkl} = R^h_{ikl} g_{hj}$, 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 $\boldsymbol{e}_i = \frac{\partial}{\partial x^i}$; 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,$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}$.

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

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

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,

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

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 $\boldsymbol{d}t$ is covariant constant, i.e.,

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

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

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

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

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

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

5. The Ricci curvature tensor has the form

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

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 $\boldsymbol{w}$ and $\boldsymbol{v}$, and for any $\boldsymbol{u}$ whatsoever,

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

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

7. The Jacobi curvature operator $\mathcal{J}(\boldsymbol{u},\boldsymbol{v})$, defined for any vectors $\boldsymbol{u},\boldsymbol{n},\boldsymbol{p}$ by

$$\mathcal{J}(\boldsymbol{u},\boldsymbol{n})\boldsymbol{p}=\frac{1}{2}[\mathcal{R}(\boldsymbol{p},\boldsymbol{n})\boldsymbol{u}+\mathcal{R}(\boldsymbol{p},\boldsymbol{u})\boldsymbol{n}]$$

is ”self-ad-joint” when operating on spatial vectors,i.e.,

$$\boldsymbol{v}\cdot[\mathcal{R}(\boldsymbol{u},\boldsymbol{n})\boldsymbol{w}]=\boldsymbol{w}\cdot[\mathcal{R}(\boldsymbol{u},\boldsymbol{n})\boldsymbol{v}] \text{ for all spactial } \boldsymbol{v},\boldsymbol{w}; \text{ and for any } \boldsymbol{u},\boldsymbol{n}.$$

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

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

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

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

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 $\{\boldsymbol{e}_{\hat{\alpha}}\}$ with

$$\boldsymbol{e}_{\hat{0}} = \boldsymbol{u} = \frac{d\mathcal{P}_0}{d\tau}$$

and with

$$\boldsymbol{e}_{\hat{\alpha}} \cdot \boldsymbol{e}_{\hat{\beta}} = \eta_{\alpha\beta}$$

3. The tetrad changes from point to point along the observer's world line, relative to parallel transport:

$$\nabla_{\boldsymbol{u}} \boldsymbol{e}_{\hat{\alpha}} = -\boldsymbol{\Omega} \cdot \boldsymbol{e}_{\hat{\alpha}}$$

$$\Omega^{\mu\nu} = a^\mu u^\nu - u^\mu a^\nu + u_\alpha \omega_\beta \epsilon^{\alpha\beta\mu\nu}$$

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).

$$\boldsymbol{a} = \nabla_{\boldsymbol{u}} \boldsymbol{u}$$

$$\boldsymbol{u} \cdot \boldsymbol{a} = \boldsymbol{u} \cdot \boldsymbol{\omega} = 0$$

If ω were zero, the observer would be Fermi-Walker-transporting his tetrad (gyroscope-type transport). If both $\boldsymbol{a}$ and $\boldsymbol{\omega}$ 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 $\boldsymbol{u}$), 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 $\mathcal{G}[\tau, \boldsymbol{n}, s]$. [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 $\boldsymbol{n} = n^{\hat{j}} \boldsymbol{e}_{\hat{j}}$; 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:

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

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

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

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

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

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

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

$$\boldsymbol{G} = 8\pi \boldsymbol{T}.$$

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

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

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

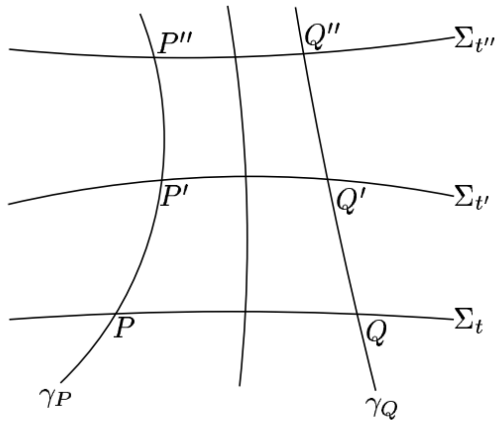


Figure 7.1: Foliation of space-time by spacelike hypersurfaces

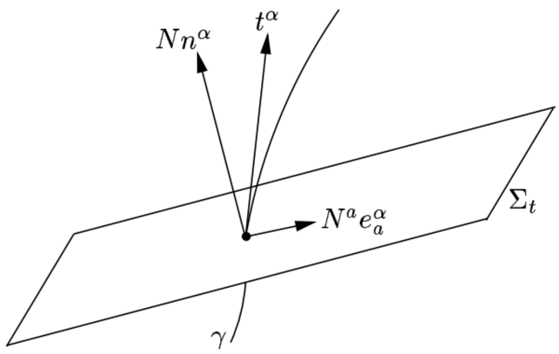


Figure 7.2: Decomposition of t^α into lapse and shift

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

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

Metric

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

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

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

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

$$
\begin{gathered}
\dot{h}_{ab} \equiv \mathcal{L}_t h_{ab} = \mathcal{L}_t(g_{\alpha\beta} e_a^\alpha e_b^\beta) = \mathcal{L}_t(g_{\alpha\beta} e_a^\alpha e_b^\beta = 2NK_{ab} + N_{a|b} + N_{b|a}) \\
K_{ab} = \frac{1}{2N}(\dot{h}_{ab} - N_{a|b} - N_{b|a}) \\
p^{ab} = \frac{\partial}{\partial \dot{h}_{ab}}(\sqrt{-g}\mathcal{L}_G) = \frac{\sqrt{h}}{16\pi}(K^{ab} - Kh^{ab}) \\
\sqrt{h}K^{ab} = 16\pi(p^{ab} - \frac{1}{2}ph^{ab}) \\
\mathcal{H}_G = p^{ab}\dot{h}_{ab} - \sqrt{-g}\mathcal{L}_G
\end{gathered}
$$

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

Variation of gravitational Hamiltonian

$$
\begin{gathered}
\delta N = \delta N^a = \delta h_{ab} = 0 \text{ on } S_t \\
\delta H_G = \int_{\Sigma_t}(\mathcal{P}^{ab}\delta h_{ab} + \mathcal{H}_{ab}\delta p^{ab} - \mathcal{C}\delta N - 2\mathcal{C}_a\delta N^a)d^3y
\end{gathered}
$$

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

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
$$\langle U^2 x, U^2 y \rangle = \overline{\langle Ux, Uy \rangle} = \langle x, y \rangle$$
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, $\boldsymbol{X} = (X_1, X_2, X_3)$. If the particle has some other internal degrees of freedom, then $\boldsymbol{X}$ plus some other observables S's will form a complete commuting set of operators. The eigenstate state will be denoted by $|\boldsymbol{x}, s\rangle$, satisfying that

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

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

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

9.5 Momentum operators and canonical quantization

Since $\boldsymbol{X}$ 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 $\boldsymbol{P}$ 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 $T(\boldsymbol{a}) \equiv e^{-i\boldsymbol{P}\cdot\boldsymbol{a}}$ We can get

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

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

So, $T(\boldsymbol{a})$ 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 $\boldsymbol{J}$ as the generator of rotation.

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

If the operator $\boldsymbol{M} = (M_1, M_2, M_3)$ 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 $\boldsymbol{L} = \boldsymbol{X} \times \boldsymbol{P}$. 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 $\boldsymbol{S} = [S_1, S_2, S_3]$, 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

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

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 $\{|\boldsymbol{x}\rangle\}$ of the position operator . Since this is a continuous set, the expansion coefficients define a function of a continuous variable,

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

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

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

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 $\psi(\boldsymbol{x})$.

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

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

The action of an position operator in coordinate representation is

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

The action of an momentum operator in coordinate representation is

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

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

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

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

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

apart from an irrelevant constant term.

10.3 Probability flux and conditions on wave functions

Define the probability flux vector

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

We can get a continuity equation

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

Applying the divergence theorem, we obtain

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

The equations of continuity require that the probability flux $\boldsymbol{J}(\boldsymbol{x},t)$ be continuous across any surface, since otherwise the surface would contain sources or sinks. Although this condition applies to all surfaces, implying that $\boldsymbol{J}(\boldsymbol{x},t)$ 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

$$\langle q_N|U(t,t_0)|q_0\rangle=\left(\prod_i\int\mathcal{D}q(t)\mathcal{D}p(t)\right)\exp\left[i\int_0^T dt\left(\sum_i p^i\dot{q}_i-H\left(q,p\right)\right)\right]$$

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 Motivation

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 $\phi_a(\boldsymbol{x})$. The mean value of the measurement of the field is described by Erenfest's theorem

$$\frac{d\langle\psi|\phi_a(\boldsymbol{x})|\psi\rangle}{dt} = -i\langle\psi|[\phi_a(\boldsymbol{x}), H_S]|\psi\rangle$$

If $[\phi_a(\boldsymbol{x}), H_S]_Q = i[\phi_a(\boldsymbol{x}), H_S]_C$, we can reproduce the classical field theory as an average effect of quantum field theory. We also note that the commutation relation here $[A, B] = AB - BA$ obeys the same operation laws as the Poisson bracket in classical field theory. So, what we need here is

1. the canonical quantization

$$[\phi_a(\boldsymbol{x}), \phi_b(\boldsymbol{y})] = 0 \quad [\pi^a(\boldsymbol{x}), \pi^b(\boldsymbol{y})] = 0 \quad [\phi_a(\boldsymbol{x}), \pi^b(\boldsymbol{y})] = i\delta_a^b\delta(\boldsymbol{x} - \boldsymbol{y})$$

2. the same definition of $\mathcal{L}$,π^a and H as those in corresponding classical theory.

In Heisenberg picture, we define

$$A_H \equiv U^{-1}(t, t_0)A_S U(t, t_0)$$

Especially,

$$\phi_a(x) \equiv U^{-1}(t, t_0)\phi_a(\boldsymbol{x})U(t, t_0)$$
$$\pi^a(x) \equiv U^{-1}(t, t_0)\pi^a(\boldsymbol{x})U(t, t_0)$$

If $A = f(\phi_a(\boldsymbol{x}), \pi^a(\boldsymbol{x}), t)$, then we have $A_H = f(\phi_a(x), \pi^a(x), t)$.
The canonical quantization can be generalized to the field operator in any time

$$[\phi_a(\boldsymbol{x}, t), \phi_b(\boldsymbol{y}, t)] = 0 \quad [\pi^a(\boldsymbol{x}, t), \pi^b(\boldsymbol{y}, t)] = 0 \quad [\phi_a(\boldsymbol{x}, t), \pi^b(\boldsymbol{y}, t)] = i\delta_a^b\delta(\boldsymbol{x} - \boldsymbol{y})$$

The dynamics of the field can be describe by Heisenberg equation

$$\frac{d\phi_a(x)}{dt} = -i[\phi_a(x), H_H]$$
$$\frac{d\pi^a(x)}{dt} = -i[\pi^a(x), H_H]$$

whose form will be equivalent to the classical field equation.

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:

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

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

Lorentz invariance means Lagrangian must be a scalar, or more loosely, action must be 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

$$P^\mu \equiv \int T^{0\mu}d^3x \quad \frac{dP^\mu}{dt} = 0$$

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

We can get the commutation relation

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

We define the translation operator $T(s)$ by

$$T^{-1}(s)\phi_a(x)T(s) = \phi_a(x-s)$$

then we can derive 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 also define that

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

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

We can get the commutation relation

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

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

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

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(-\eta^{\nu\rho} M^{\mu\sigma} + \eta^{\sigma\mu} M^{\rho\nu} + \eta^{\mu\rho} M^{\nu\sigma} - \eta^{\sigma\nu} M^{\rho\mu})$$

We again define $J_i \equiv \frac{1}{2}\epsilon_{ijk} M^{jk}$ and $K_i \equiv M^{i0}$, the communication relation 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,

$$[P^\mu, M^{\rho\sigma}] = i(\eta^{\mu\sigma} P^\mu - \eta^{\mu\rho} P^\sigma)$$

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

At last, we also 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 define the rotation operator $U(\Lambda)$ by

$$U^{-1}(\Lambda)\phi_a(x)U(\Lambda) = \mathcal{S}_a{}^b\phi_b(\Lambda^{-1}x)$$

where

$$\mathcal{S}_a{}^b = \delta_a{}^b + \frac{i}{2}\delta\omega_{\alpha\beta}(S^{\alpha\beta})_a{}^b$$

We can derive that

$$U(1+\delta\omega) = 1 + \frac{i}{2}\delta\omega_{\mu\nu}M^{\mu\nu} \quad U(\Lambda) = e^{\frac{i}{2}\theta_{\mu\nu}M^{\mu\nu}}$$

$$U^{-1}(\Lambda)P^\mu U(\Lambda) = \Lambda^\mu{}_\nu P^\nu$$

$$U^{-1}(\Lambda)M^{\mu\nu}U(\Lambda) = \Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma M^{\rho\sigma}$$

11.5 Anticommutation relation

We define the anticommutation relation of operators as $\{A, B\} \equiv AB + BA$. Suppose that the field operator and its canonical momentum operator has the following anticommutation relation

$$\{\phi_a(\boldsymbol{x}, t), \phi_b(\boldsymbol{y}, t)\} = 0 \quad \{\pi^a(\boldsymbol{x}, t), \pi^b(\boldsymbol{y}, t)\} = 0 \quad \{\phi_a(\boldsymbol{x}, t), \pi^b(\boldsymbol{y}, t)\} = i\delta_a^b\delta(\boldsymbol{x} - \boldsymbol{y})$$

If the operator A composes of the terms like $\pi^a\mathcal{E}_a{}^b\phi_b$, we can show that the value of $[\phi_a, A]$ and $[\pi^a, A]$ is the same as those in the theory quantized with commutation relation. It is easy to verify that the operator P^i and $M_S^{\mu\nu}$ has the required form. The form of H is determined by the specific theory. As we can see later, the Hamiltonian of Dirac field has the required form. When it is quantized with anticommutation relation, the commutation relation between field operator and momentum, angular momentum discussed in previous section will be hold automatically.

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

$$
\begin{aligned}
[\phi(\boldsymbol{x},t),\phi(\boldsymbol{y},t)] &= 0\\
[\pi(\boldsymbol{x},t),\pi(\boldsymbol{y},t)] &= 0\\
[\phi(\boldsymbol{x},t),\pi(\boldsymbol{y},t)] &= i\delta(\boldsymbol{x}-\boldsymbol{y})
\end{aligned}
$$

Fourier expansion

$$
\phi(\boldsymbol{x},t) = \int \widetilde{dk}\left[a(\boldsymbol{k})e^{ikx} + a^\dagger(\boldsymbol{k})e^{-ikx}\right]
$$

$$
\pi(\boldsymbol{x},t) = -i\int \widetilde{dk}\omega\left[a(\boldsymbol{k})e^{ikx} - a^\dagger(\boldsymbol{k})e^{-ikx}\right]
$$

Here, $k^2 = \boldsymbol{k}^2 - \omega^2 = -m^2$, $kx = \boldsymbol{k}\cdot\boldsymbol{x} - \omega t$, $\widetilde{dk} = \frac{d^3}{(2\pi)^2 2\omega}$

$$
a(\boldsymbol{k}) = \int d^3x e^{-ikx}(i\pi + \omega\phi)
$$

$$
a^\dagger(\boldsymbol{k}) = \int d^3x e^{ikx}(-i\pi + \omega\phi)
$$

$$
\begin{aligned}
[a(\boldsymbol{p}),a(\boldsymbol{q})] &= 0\\
\left[a^\dagger(\boldsymbol{p}),a^\dagger(\boldsymbol{q})\right] &= 0\\
\left[a(\boldsymbol{p}),a^\dagger(\boldsymbol{q})\right] &= (2\pi)^3 2\omega\delta(\boldsymbol{p}-\boldsymbol{q})
\end{aligned}
$$

Operator represented by a and $a^\dagger$

$$
H = \int \widetilde{dk}\,\omega\, a^\dagger(\boldsymbol{k})a(\boldsymbol{k}) + (\mathcal{E}_0 - \Omega_0)V \quad \mathcal{E}_0 = \frac{1}{2}(2\pi)^{-3}\int d^3k\,\omega
$$

$$
P^i = \int \widetilde{dk}\,k^i\,a^\dagger(\boldsymbol{k})a(\boldsymbol{k})
$$

Particles

$$
[H,a(\boldsymbol{k})] = -\omega a(\boldsymbol{k}) \quad [H,a^\dagger(\boldsymbol{k})] = \omega a^\dagger(\boldsymbol{k})
$$

$$
[P^i,a(\boldsymbol{k})] = -k^i a(\boldsymbol{k}) \quad [P^i,a^\dagger(\boldsymbol{k})] = k^i a^\dagger(\boldsymbol{k})
$$

Let $|p\rangle = a^\dagger(\boldsymbol{p})|0\rangle$,so

$$
H|p\rangle = \omega_p|p\rangle \quad P^i|p\rangle = p^i|p\rangle
$$

So, we interpret the state $|\boldsymbol{p}\rangle$ as the momentum eigenstate of a single particle of mass m. We can also show that $J_i|\boldsymbol{p}=0\rangle = 0$, 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

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4$$

$$H = H_0 + H_{int} \quad H_{int} = \int d^3x \frac{\lambda_0^4}{4!}\phi^4(\boldsymbol{x})$$

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* $E_0 = \langle\Omega|H|\Omega\rangle$*.*

$$\phi(t_0, \boldsymbol{x}) = \int \frac{d^3p}{(2\pi)^3}(a(\boldsymbol{p})e^{i\boldsymbol{p}\cdot\boldsymbol{x}} + a^\dagger(\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}})$$
$$\phi(t, \boldsymbol{x}) = e^{iH(t-t_0)}\phi(t_0, \boldsymbol{x})e^{-iH(t-t_0)}$$
$$\phi_I(t, \boldsymbol{x}) \equiv e^{iH_0(t-t_0)}\phi(t_0, \boldsymbol{x})e^{-iH_0(t-t_0)}$$
$$H_I(x) = \int d^3x \frac{\lambda_0^4}{4!}\phi_I^4$$

The perturbation expansion of correlation functions is

$$\langle\Omega|T\{\phi(x)\phi(y)\}|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\langle 0|T\left\{\phi_I(x)\phi_I(y)\exp\left[-i\int_{-T}^{T} dt H_I\right]\right\}|0\rangle}{\langle 0|T\left\{\exp\left[-i\int_{-T}^{T} dt H_I\right]\right\}|0\rangle}$$

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 $\langle 0|T\left\{\phi_I(x)\phi_I(y)\exp\left[-i\int_{-T}^{T} dt H_I\right]\right\}|0\rangle$ to the first order of λ_0

$$\begin{aligned}&\langle 0|T\left\{\phi_I(x)\phi_I(y)\frac{-i\lambda_0}{4!}\int dz^4\phi_I(z)\phi_I(z)\phi_I(z)\phi_I(z)\right\}|0\rangle\\ =\ & 3\cdot(\frac{-i\lambda_0}{4!})D_F(x-y)\int d^4z D_F(z-z)D_F(z-z)\\ +\ & 12\cdot(\frac{-i\lambda_0}{4!})\int d^4z D_F(x-z)D_F(y-z)D_F(z-z)\end{aligned}$$

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 $\langle 0 | T \left\{ \phi_I(x_1)\phi_I(x_2) \cdots \phi_I(x_n) \text{exp} \left[-i \int_{-T}^{T} dt H_I \right] \right\} | 0 \rangle$ are all cancelled by the $\langle 0 | T \left\{ \text{exp} \left[-i \int_{-T}^{T} dt H_I \right] \right\} | 0 \rangle$.

12.4 Path integral formulation

12.4.1 Basic formulation

Recall the path integrals formulation in quantum mechanics, we have

$$\langle \phi_b(\boldsymbol{x}) | e^{-iHT} | \phi_a(\boldsymbol{x}) \rangle = \int \mathcal{D}\phi \mathcal{D}\pi \text{exp} \left[i \int_0^T d^4 x (\pi \dot{\phi} - \frac{1}{2}\pi^2 - \frac{1}{2}(\nabla \phi)^2 - V(\phi)) \right]$$

Here, $\langle \phi_b(\boldsymbol{x}) |$ is the eigenstate of $\phi_S(\boldsymbol{x}) = \phi_H(\boldsymbol{x}, 0)$ with eigenvalue $\phi_b(\boldsymbol{x})$ at time $t = T$,$| \phi_a(\boldsymbol{x}) \rangle$ is the eigenstate of $\phi_S(\boldsymbol{x})$ with eigenvalue $\phi_a(\boldsymbol{x})$ 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

$$\langle \phi_b(\boldsymbol{x}) | e^{-iHT} | \phi_a(\boldsymbol{x}) \rangle = \int \mathcal{D}\phi \text{exp} \left[i \int_0^T d^4 x \mathcal{L} \right]$$

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,

$$\langle 0|T\phi_H(x_1)\phi_H(x_2)|0\rangle = \frac{1}{Z_0}-\frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}\text{exp}[-\frac{1}{2}\int d^4xd^4yJ(x)D_F(x-y)J(y)]|_{J=0} = D_F(x_1-x_2)$$

12.5 Perturbation theory for path integral quantization

$$
\begin{aligned}
\mathcal{L} &= -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4 \\
\mathcal{L} &= \mathcal{L}_0 + \mathcal{L}_1 \quad \mathcal{L}_1 = -\frac{\lambda_0}{4!}\phi^4(\boldsymbol{x}) \\
Z[J] &= \int \mathcal{D}\phi e^{i\int d^4x[\mathcal{L}_0+\mathcal{L}_1+J\phi]} \\
&= e^{i\int d^4y\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(y)})}\int \mathcal{D}\phi e^{i\int d^4x[\mathcal{L}_0+J\phi]} \\
&\propto e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)})}\exp[-\frac{1}{2}\int d^4yd^4zJ(y)D_F(y-z)J(z)] \\
&= \sum_{V=0}^{\infty}\frac{1}{V!}[\frac{-i\lambda_0}{4!}\int d^4x(\frac{1}{i}\frac{\delta}{\delta J(x)})^4]^V \times \sum_{P=0}^{\infty}\frac{1}{P!}[-\frac{1}{2}\int d^4yd^4zJ(y)D_F(y-z)J(z)]^P
\end{aligned}
$$

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 Symmetries in the functional formalism

Equations of motion

The correlation function of the field theory is given by

$$\langle\Omega|T\phi_H(x_1)\cdots\phi_H(x_n)|\Omega\rangle = Z_0^{-1}\int\mathcal{D}\phi e^{iS}\phi(x_1)\cdots\phi(x_n)$$

The equation of motion of classical field theory will be give by

$$\frac{\delta S}{\delta\phi(x)} = 0$$

In quantum field theory, we derive the equation of motion by claim that the path integral will be invariant under the infinitesimal change of field, i.e. $\phi(x)\to\phi(x)+\epsilon(x)$. Define

$$Z[\phi(x_1),\cdots,\phi(x_n)] \equiv \int\mathcal{D}\phi e^{iS}\phi(x_1)\cdots\phi(x_n)$$

We know

$$\delta Z = \int\mathcal{D}\phi e^{iS}\left\{\int d^4x\epsilon(x)\left[i\frac{\delta S}{\delta\phi(x)}\phi(x_1)\cdots\phi(x_n)+\delta(x-x_1)\cdots\phi(x_n)+\cdots+\phi(x_1)\cdots\delta(x-x_n)\right]\right\}$$

so

$$\langle\frac{\delta S}{\delta\phi(x)}\phi(x_1)\cdots\phi(x_n)\rangle = \sum_{i=1}^{n}\langle\phi(x_1)\cdots(i\delta(x-x_1))\cdots\phi(x_n)\rangle$$

Example: For Klein-Gordon field,

$$\mathcal{L} = -\frac{1}{2}\phi^\mu\phi\partial_\mu\phi-\frac{1}{2}m\phi^2$$

The variation of S gives

$$\frac{\delta S}{\delta\phi(x)} = (\partial^2-m^2)\phi(x)$$

So, we can get

$$(\partial^2-m^2)\langle T\phi(x)\phi(x_1)\rangle = i\delta(x-x_1)$$

Conservation law

Consider a local field theory of a set of fields $\phi_a(x)$, governed by a Lagrangian $\mathcal{L}(\phi)$. An infinitesimal symmetry transformation on the fields ϕ_a will be of the general form

$$\phi_a(x)\to\phi_a(x)+\epsilon\Delta\phi_a(x)$$

We assume that when ϵ is a constant, the action is invariant under this transformation. Then the Lagrangian must be invariant up to a total divergence:

$$\mathcal{L}[\phi]\to\mathcal{L}[\phi]+\epsilon\partial_\mu K^\mu$$

If the symmetry parameter ϵ depend on x, the variation of Lagrangian will be

$$L[\phi] \to \mathcal{L}[\phi] + (\partial_\mu \epsilon)\Delta\phi_a \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} + \epsilon \partial_\mu K^\mu$$

So,

$$\frac{\delta S}{\delta \epsilon(x)} = \partial_\mu j^\mu \quad j^\mu = -\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\Delta\phi_a + K^\mu$$

If the measure $\mathcal{D}\phi$ is invariant under the transformation, we can derive similarly that

$$\langle \partial_\mu j^\mu(x)\phi(x_1)\cdots\phi(x_n)\rangle = \sum_{i=1}^{n} \langle \phi(x_1)\cdots(i\Delta\phi(x_i)\delta(x-x_i))\cdots\phi(x_n)\rangle$$

12.7 Cross section and the S-matrix

Recall the definition of cross section

$$\sigma \equiv \frac{\text{Number of Events}}{\text{Time} \times \text{Incident Flux}}$$

Consider a $2 \to n$ process:

$$p_1 + p_2 \to \{p_j\}$$

Suppose the volume of the space in which the scattering process takes place is V, the duration of the scattering process is T. So, the incident flux is

$$\Phi = \frac{|\boldsymbol{v}_1 - \boldsymbol{v}_2|}{V},$$

and the number of events is

$$N = \frac{|\langle i|S|f\rangle|^2}{|\langle i|i\rangle||\langle f|f\rangle|} d\Pi.$$

Here, $d\Pi$ is the region of final state momenta at which we are looking

$$d\Pi = \prod_j \frac{V}{(2\pi)^3} d^3 p_j$$

Recall that

$$\delta^{(3)}(0) = \frac{V}{(2\pi)^3} \quad \delta^{(4)}(0) = \frac{VT}{(2\pi)^4} \quad \langle p|p\rangle = (2\pi)^3 2\omega\delta^{(3)}(0)$$

We can calculate that

$$|\langle i|S|f\rangle|^2 = |\mathcal{M}|^2 VT(2\pi)^4 \delta(\sum p_j - p_1 - p_2) \quad \langle i|i\rangle = (2E_1 V)(2E_2 V) \quad \langle f|f\rangle = \prod_j (2E_j V)$$

At last, putting everything together, we have

$$d\sigma = \frac{1}{(2E_1)(2E_2)|\boldsymbol{v}_1 - \boldsymbol{v}_2|} |\mathcal{M}|^2 d\Pi_{\text{LIPS}}$$

where,

$$d\Pi_{\text{LIPS}} = \prod_j \frac{d^3 p_j}{(2\pi)^3} \frac{1}{2E_j} (2\pi)^4 \delta(\sum p_j - p_1 - p_2)$$

is called the Lorentz-invariant phase space (LIPS).
All the factors of V and T have dropped out, so now it is trivial to take $V \to \infty$ and $T \to \infty$.

Decay rates

The definition of decay rates is

$$\Gamma \equiv \frac{\text{Number of Events}}{\text{Time}}$$

Consider a $1 \to n$ process:

$$p_1 \to \{p_j\}$$

We can derive the decay rates by the similar method,

$$d\Gamma = \frac{1}{2E_1}|\mathcal{M}|^2 d\Pi_{\text{LIPS}}$$

Special cases

For $2 \to 2$ scattering in the center-of-mass frame

$$p_1 + p_2 \to p_3 + p_4$$

Then

$$d\Pi_{\text{LIPS}} = \frac{d^3p_3}{(2\pi)^3}\frac{1}{2E_3}\frac{d^3p_4}{(2\pi)^3}\frac{1}{2E_4}(2\pi)^4\delta(p_3 + p_4 - p_1 - p_2)$$

We can now integrate over $\boldsymbol{p}_4$ to give

$$d\Pi_{\text{LIPS}} = \frac{1}{16\pi^2}d\Omega\int dp_f \frac{p_f^2}{E_3E_4}\delta(E_3 + E_4 - E_{CM})$$

where $p_f = |\boldsymbol{p}_3| = |\boldsymbol{p}_4|$, $E_3 = \sqrt{p_f^2 + m_3^2}$,$E_4 = \sqrt{p_f^2 + m_4^2}$ and $E_{CM} = E_1 + E_2$. Define $x(p_f) \equiv E_3 + E_4 - E_{CM}$, we can get

$$\frac{dx}{dp_f} = \frac{E_3 + E_4}{E_3E_4}p_f$$

So,

$$d\Pi_{\text{LIPS}} = \frac{1}{16\pi^2}d\Omega\frac{p_f}{E_{CM}}\theta(E_{CM} - m_3 - m_4)$$

Plugging this into the general equation for cross section, we can get

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2E_{CM}^2}\frac{|\boldsymbol{p}_f|}{|\boldsymbol{p}_i|}|\mathcal{M}|^2\theta(E_{CM} - m_3 - m_4)$$

If all the masses are equal then this formula simplifies further

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2E_{CM}^2}|\mathcal{M}|^2$$

12.8 LSZ reduction formula

12.8.1 Field strength renormalization

The completeness relation:

$$\mathbf{1} = |\Omega\rangle\langle\Omega| + \sum_\lambda \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\boldsymbol{p}}}|\lambda_{\boldsymbol{p}}\rangle\langle\lambda_{\boldsymbol{p}}|$$

Here, $E_{\boldsymbol{p}} = \sqrt{m_\lambda^2 + \boldsymbol{p}^2}$

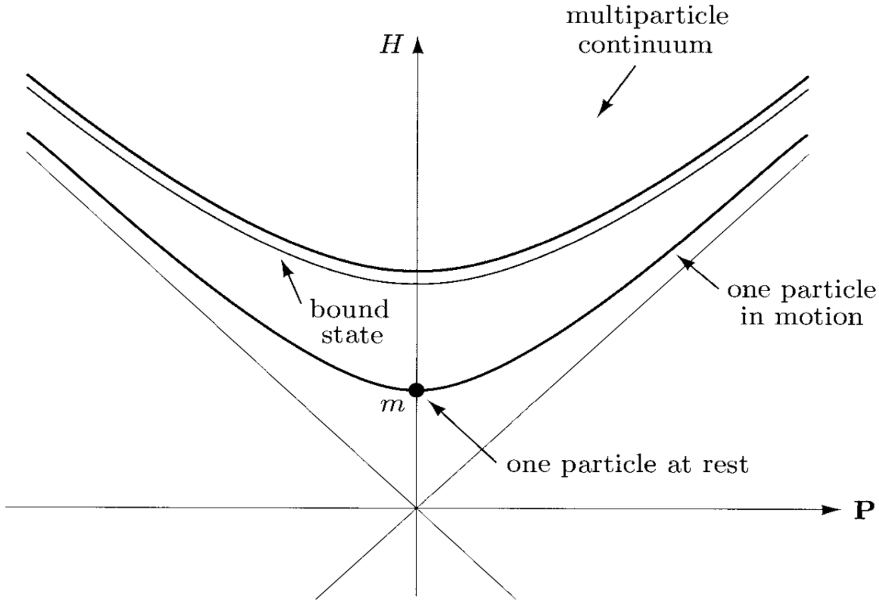


Figure 12.4: Particle's energy-momentum relation

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

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle_C = \sum_\lambda \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\boldsymbol{p}}}\langle\Omega|\phi(x)|\lambda_{\boldsymbol{p}}\rangle\langle\lambda_{\boldsymbol{p}}|\phi(y)|\Omega\rangle$$

It can be verified that

$$\langle\Omega|\phi(x)|\lambda_{\boldsymbol{p}}\rangle = \langle\langle\Omega|\phi(0)|\lambda_0\rangle e^{ipx}|_{p^0=E_{\boldsymbol{p}}}$$

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

$$\rho(M^2) = \sum_{\lambda}(2\pi)\delta(M^2 - m^2)|\langle\Omega|\phi(0)|\lambda_0\rangle|^2$$

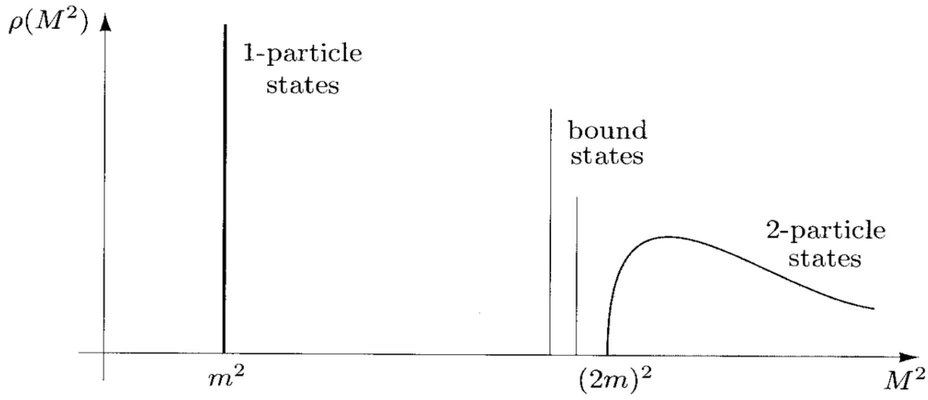


Figure 12.5: The structure of the spectral density function $\rho(M^2)$

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

$$\begin{aligned}&\int d^4x e^{-ipx}\langle\Omega|T\phi(x)\phi(0)|\Omega\rangle_C \\ =\ &\int_0^\infty \frac{dM^2}{2\pi}\rho(M^2)\frac{-i}{p^2+M^2-i\epsilon} = \frac{-iZ}{p^2+m^2-i\epsilon} + \int_{\sim 4m^2}^\infty \frac{dM^2}{2\pi}\rho(M^2)\frac{-i}{p^2+M^2-i\epsilon}\end{aligned}$$

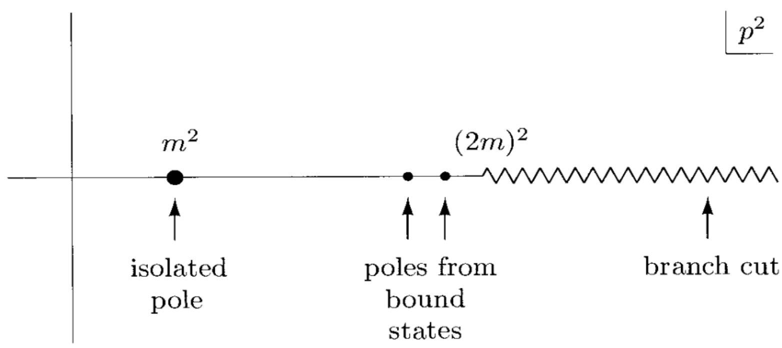


Figure 12.6: The structure of the two point function in Fourier space

12.8.2 LSZ reduction formula

Theorem 12.2 LSZ reduction formula

$$\prod_1^n \int d^4x_i e^{-ip_ix_i} \prod_1^m d^4y_j e^{ik_jy_j} \langle\Omega|T\{\phi(x_1)\cdots\phi(x_n)\phi(y_1)\cdots\phi(y_m)\}|\Omega\rangle$$
$$\underset{p_i^0\to E_{\boldsymbol{p}_i}\, k_i^0\to E_{\boldsymbol{k}_i}}{\sim} \left(\prod_1^n \frac{-\sqrt{Z}i}{p_i^2+m^2-i\epsilon}\right)\left(\prod_1^m \frac{-\sqrt{Z}i}{k_i^2+m^2-i\epsilon}\right)\langle \boldsymbol{p}_1\cdots\boldsymbol{p}_n|S|\boldsymbol{k}_1\cdots\boldsymbol{k}_m\rangle$$

The $\sim$ means the two sides of the expression share the same singular structure around $p_i^0 \to E_{\boldsymbol{p}_i}$, $k_i^0 \to E_{\boldsymbol{k}_i}$. 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.

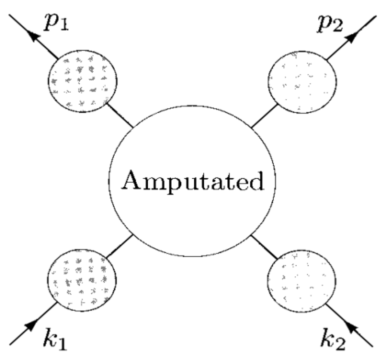


Figure 12.7: Amputated Feynman diagram

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:

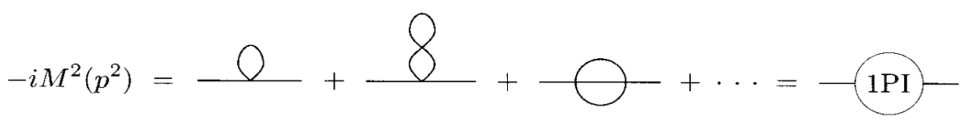


Figure 12.8: Diagram representation of 1PI propagator

Note: *One-particle-irreducible, or 1PI for short, refers to diagrams that is still connected after one line is cut*

Then the exact propagator can be written as a geometric series in Figure 2.9.

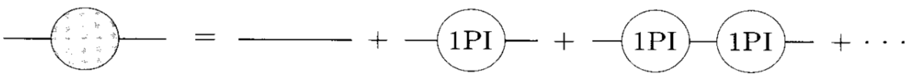


Figure 12.9: Diagram representation of exact propagator

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

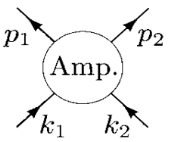
$$\frac{-i}{p^2+m_0^2+M^2} \underset{p^0 \to E_p}{\sim} \frac{-iZ}{p^2+m^2} + \text{(regular)}$$

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 $\langle f|S|i\rangle = (Z_1)^{\frac{n_s}{2}} i\mathcal{M}(2\pi)^4\delta(\sum p_f - \sum p_i)$ for convenience.

12.9 Renormalization

Renormalization, 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.

12.9.1 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^dp, 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-renormalizable theory: Only a finite number of Feynman diagrams superficially diverge.

2. renormalizable theory: Only a finite number of amplitudes superficially diverge; however, divergences occur at all orders in perturbation theory.

3. Non-renormalizable 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 renormalizable theory. For ϕ^3 theory in four dimension, $D = 4 - V - N$. It is a super-renormalizable theory. For ϕ^6 theory in four dimension, $D = 4 + 2V - N$. It is a Non-renormalizable 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 renormalizability in a second way:

1. Super-renormalizable: Coupling constant has positive mass dimension.
2. renormalizable: Coupling constant is dimensionless.
3. Non-renormalizable: Coupling constant has negative mass dimension.

12.9.2 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

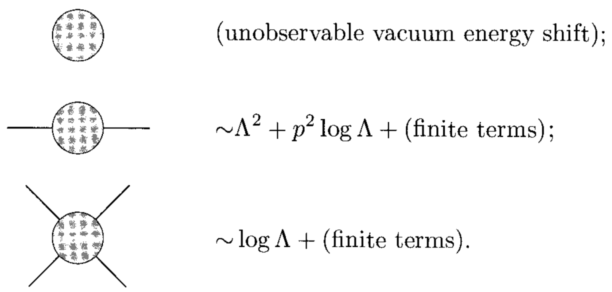


Figure 12.11: Divergence of ϕ^4 theory

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

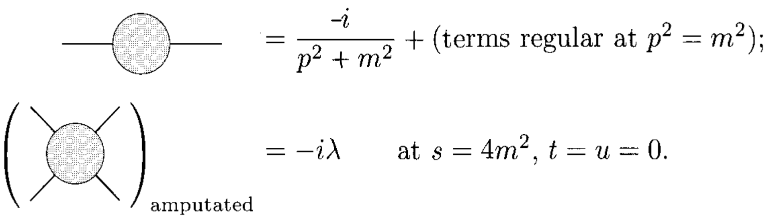


Figure 12.12: Renormalization condition

the physical mass and coupling constant as Figure 2.12.

The renormalization scheme here is called on-shell (OS) scheme. Other renormalization scheme would be introduced later. These equations are called renormalization conditions. Our new Lagrangian gives a new set of Feynman rules as Figure 2.13.

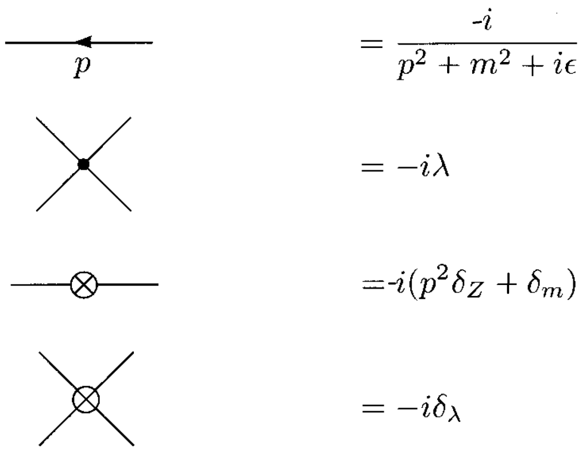


Figure 12.13: Feynman rules for renormalized perturbation theory

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 renormalization 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

$$s = -(p_1 + p_2)^2 = -(p_3 + p_4)^2$$
$$t = -(p_1 - p_3)^2 = -(p_2 - p_4)^2$$
$$u = -(p_1 - p_4)^2 = -(p_2 - p_3)^2$$

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

$$s + t + u = m_1^2 + m_3^2 + m_3^2 + m_4^2$$

12.9.3 Techniques for renormalization

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

$$\frac{1}{A_1^{\alpha_1} \cdots A_n^{\alpha_n}} = \frac{\Gamma(\sum_i \alpha_i)}{\prod_i \Gamma(\alpha_i)} \frac{1}{(n-1)!} \int dF_n \frac{\prod_i x_i^{\alpha_i - 1}}{(\sum_i x_i A_i)^{\sum_i \alpha_i}}$$

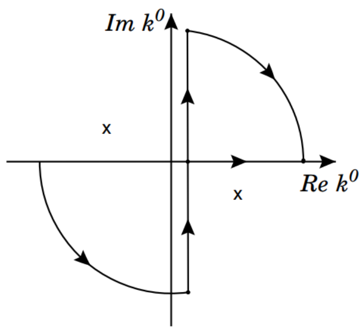


Figure 12.14: Wick rotation

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,

$$\int d^d q f(q^2 - i\epsilon) = i \int d^d \bar{q} f(\bar{q}^2)$$

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

$$\int \frac{d^d \bar{q}}{(2\pi)^d} \frac{1}{(\bar{q}^2 + D)^b} = \frac{\Gamma(b - \frac{1}{2}d)}{(4\pi)^{d/2}\Gamma(b)} D^{-(b-d/2)}$$

12.9.4 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

$$\frac{(-i\lambda)^2}{2} \int \frac{d^4 k}{(2\pi)^4} \frac{-i}{k^2 + m^2} \frac{-i}{(k+m)^2 + m^2} \equiv (-i\lambda)^2 i V(-p^2)$$

So the entire amplitude is therefore

$$i\mathcal{M} = -i\lambda + (-i\lambda)^2[iV(s) + iV(t)iV(u)] - i\delta_\lambda + \mathcal{O}(\lambda^3)$$

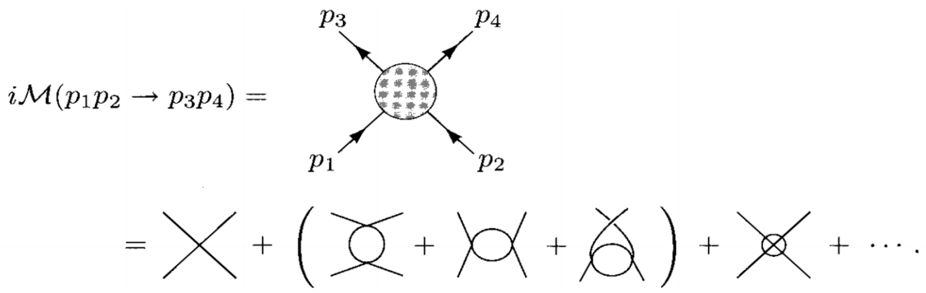


Figure 12.15: Feynman diagram representation of two-particle scattering to one loop

To keep λ dimensionless in dimensional regularization, we can make the transformation $\lambda \to \lambda\tilde{\mu}^{\epsilon}$. 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 renormalization 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 renormalization 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.

Perturbation theory to all orders

We begin by summing all one-particle irreducible diagrams with two external lines; this gives us the self-energy. We next sum all 1PI diagrams with four external lines; this gives us the four-point vertex function $V_4(k_1, k_2, k_3, k_4)$. Order by order in λ, we must adjust the value of the lagrangian coefficients δ_Z, δ_m, and δ_λ to maintain the conditions $M^2(-m^2) = 0$, $\frac{dM^2}{dp^2}(-m^2) = 0$, and $V_4(s = 4m^2) = 0$.

Next we will construct the n-point vertex functions V_n with $4 < n \leq E$, where E is the number of external lines in the process of interest. We compute these using a skeleton expansion. This means that we draw all the contributing 1PI diagrams, but omit diagrams that include either propagator or three-point vertex corrections. That is, we include only diagrams that are not only 1PI, but also 2PI and 3PI: they remain connected when any one, two, or three lines are cut. (Cutting three lines may isolate a single treelevel vertex, but nothing more complicated.) Then we take the propagators and vertices in these diagrams to be given by the exact propagator $\frac{-i}{p^2+m^2+M^2(p^2)}$ and vertex $V_4(k_1, k_2, k_3, k_4)$, rather than by the tree-level propagator and vertex. We then sum these skeleton diagrams to get V_n for $4 < n \leq E$. Order by order in λ, this procedure is equivalent to computing V_n by summing the usual set of contributing 1PI diagrams.

Next we draw all tree-level diagrams that contribute to the process of interest (which has E external lines), including not only four-point vertices, but also n-point vertices. Then we evaluate these diagrams using the exact propagator for internal lines, and the exact 1PI vertices V_n; external lines are assigned a factor of one. We sum these tree diagrams to get the scattering amplitude. Order by order in λ, this procedure is equivalent to computing the scattering amplitude by summing the usual set of contributing diagrams.

Thus we now know how to compute an arbitrary scattering amplitude to arbitrarily high order. The procedure is the same in any quantum field theory; only the form of the propagators and vertices change, depending on the spins of the fields.

12.9.5 General renormalization theory

Recall some of the major results and methods of renormalization theory:

1. In perturbation theory, bare and physical quantities are related by ultraviolet-divergent expressions

$$m_{phys} = m_0 + \Delta_m$$

   where m_{phys} is finite, Δ_m is ultraviolet-divergent, and so m_0 is necessarily ultraviolet-divergent.

2. We express the Lagrangian in terms of physical quantites, and separate it into

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I + \mathcal{L}_{CT}$$

   where $\mathcal{L}_0$ is the canonically normalized free Lagrangian for physical fields and masses, $\mathcal{L}_I$ contains the interaction, again in terms of physical parameters, and $\mathcal{L}_{CT}$ contains the counterterms with ultraviolet divergent coefficients. From $\mathcal{L}_0$, we obtain the propagators of the physical fields. $\mathcal{L}_I$ and $\mathcal{L}_{CT}$ give interaction vertices.

3. At the one-loop level, the self-energy is given by the effective two-point vertices: the 1PI two-point vertex of the interaction and the counter-term two-point vertex. The counterterms absorb ultraviolet divergences, and the finite parts of the counterterms are determined by renormalization conditions, which ensure the quantities in $\mathcal{L}_0 + \mathcal{L}_I$ are physical. The conditions constrain the self-energy and the effective vertices, and give a finite, uniquely-determined value for the counterterms.

Degrees of divergences

Now, for a general theory in d spacetime dimensions, the field content is given by ϕ_f, $f = 1, 2, \cdots$, where f labels the field type. $[\phi_f] = \Delta f$ and Δf in all physical theories. We have interaction vertices of type i, $i = 1, 2, \cdots$ contributing a term of the form

$$\lambda_i \partial^{n_i} \prod_f \phi_f^{n_{if}}$$

Here, λ_i is the coupling constant, with dimension

$$[\lambda_i] = \kappa_i = d - n_i - \sum_f n_{if} \Delta f$$

Now consider a 1PI diagram in such a theory:

$$E_f \equiv \text{ number of external lines of } \phi_f$$

$$V_i \equiv \text{ number of vertices of type i .}$$

Then $M \sim \Lambda^D \prod_i \lambda_i^{V_i}$ and so $D = [M] - \sum_i V_i \kappa_i$. Again

$$[M] - d = \sum E_f \Delta f$$

and so the general expression for the superficial degree of divergence is given by

$$D = d - \sum_f E_f \Delta f - \sum_i V_i \kappa_i$$

Diagrams which are ultraviolet divergent satisfy

$$\sum_f E_f \Delta f + \sum_i V_i \kappa_i < d$$

We can now divide all theories into

1. All $\kappa_i > 0$: superrenormalizable theories.
2. All $\kappa_i \geq 0$: renormalizable theories.
3. There exists at least one $\delta_i < 0$: non-renormalizable theories.

These terms also apply to individual interactions for a vertex of type i:

1. $\kappa_i > 0$: super-renormalizable, relevant interaction.
2. $\kappa_i \geq 0$: renormalizable, marginal interaction.
3. $\kappa_i < 0$: non-renormalizable, irrelevant interaction.

Cancellation of divergences

Consider a generic divergent diagram M of degree D; that is,

$$M = \int^{\Lambda} dss^{D-1}$$

if all loop momenta are taken proportional to s. Generically, internal propagators have the form

$$\frac{1}{(as+p)^{\alpha}\cdots} \sim \frac{1}{s^{\alpha}}$$

for large s, where a is a numerical constant and p is a combination of the external momenta. Differentiating M, with respect to p, n times gives a term proportional to

$$\frac{1}{(as+p)^{\alpha+n}} \sim \frac{1}{p^{\alpha+n}}$$

and so $D+1$ derivatives with respect to the external momenta will make M finite. This means that

$$M(p) = M_0 + M_1 p + \cdots + M_D p^D + \text{ finite terms}$$

where the argument p of the function represents the collection of external momenta, we have suppressed the index structure, and $M_0, M_1, \cdots, M_D$ are potentially divergent constants. Suppose that M has E_f external lines of the field ϕ_f. Then divergences of $M(p)$ can be cancelled by counterterms of the form

$$\sum_{j=0}^{D} A_j (\partial)^j \prod_f \phi_f^{E_f}$$

where the A_j are divergent coefficients in order to cancel divergences in M_j. The index structure in $A_j \partial^j$ should match the suppressed index structure of $M(p)$.

12.10 Renormalization group

12.10.1 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: $k^2 = -m_{ph}^2$. Thus, the Lagrangian parameter m is no longer the same as m_{ph}. The relation of m and m_{ph} is

$$m_{ph}^2 = M^2(-m_{ph}^2) + m^2$$

To the lowest order,

$$m_{ph}^2 = \left[1 + \frac{\lambda}{32\pi^2}(\ln(\frac{m^2}{\mu^2}) - 1)\right] m^2$$

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,

$$\delta_\lambda = \left[\frac{3\lambda^2}{16\pi^2} + \mathcal{O}(\lambda^3)\right] \frac{1}{\epsilon}$$

$$i\mathcal{M} = -i\lambda - \frac{i\lambda^2}{32\pi^2}\int_0^1 dx \left[\ln(\frac{D(s)}{\mu^2}) + \ln(\frac{D(t)}{\mu^2}) + \ln(\frac{D(u)}{\mu^2})\right] + \mathcal{O}(\lambda^3)$$

For a process with $p^2 \gg m^2$, we have

$$D \approx x(1-x)p^2$$

In OS renormalization scheme, the one-loop correction to propagator or vertex generally includes a factor

$$\ln\left(\frac{D}{D_0}\right) \sim \ln\frac{p^2}{m^2}$$

so perturbation theory is no longer a good approximation when $p^2 \gg m^2$. In $\overline{MS}$ renormalization scheme, introducing μ allows us to address this problem: if we choose $\mu \sim p$, no such logarithm arises. If we choose μ appropriately, that is, to be comparable to the momentum scale of the physical process, we can improve our perturbation expansion. So $\lambda(\mu)$ and $m(\mu)$ can be considered as the scale-dependent coupling constants. And the reason we get large logarithmic terms in the on-shell scheme is that we are trying to use coupling defined at one scale to describe physics at very different scales.

12.10.2 Beta and gamma function

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 $\overline{\text{MS}}$ renormalization 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

$$\left(1 + \frac{\lambda G_1'}{\epsilon} + \cdots\right) \frac{d\lambda}{d\ln \mu} + \epsilon\lambda = 0$$

$\frac{d\lambda}{d\ln\mu}$ is the rate at which λ must change to compensate for a small change in $\ln \mu$. If compensation is possible at all, this rate should be finite in the $\epsilon \to 0$ limit. In a renormalizable 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

$$\ln Z_\phi = \frac{a_1}{\epsilon} + \cdots$$

Define the anomalous dimension of the field

$$\gamma_\phi(\lambda) \equiv \frac{1}{2}\frac{d\ln Z_\phi}{d\ln\mu}$$

We can derive

$$\gamma_\phi(\lambda) = -\frac{1}{2}\lambda a_1'$$

In ϕ^4 theory, we have

$$\gamma_\phi(\lambda) = \mathcal{O}(\lambda^2)$$

12.10.3 Callen-Symanzik equation

Consider the n point green function,

$$G^{(n)}(x_1, \cdots, x_n) \equiv \langle\Omega|T\phi(x_1)\cdots\phi(x_n)|\Omega\rangle_C$$

As $G_0^{(n)} = Z_\phi^{n/2}G^{(n)}$, from the independence of bare Green's function, we have

$$\left(\frac{\partial}{\partial\ln\mu} + \beta(\lambda)\frac{\partial}{\partial\lambda} + \gamma_m(\lambda)m\frac{\partial}{\partial m} + n\gamma_\phi(\lambda)\right)G^n(x_1, \cdots, x_n; \lambda, m, \mu) = 0$$

From now on, we will focus on the ϕ^4 theory in massless limit. Firstly, consider the two point Green's function in momentum space $G^{(2)}(p)$, we can express its dependence on p and μ as

$$G^{(2)} = \frac{-i}{p^2}g(-p^2/\mu^2, \lambda(\mu))$$

Then the C-S equation can be written as

$$\left[p\frac{\partial}{\partial p} - \beta(\lambda)\frac{\partial}{\partial\lambda} + 2 - 2\gamma_\phi\right]G^{(2)}(p, \lambda(\mu), \mu) = 0$$

Here, $p \equiv (-p^2)^{1/2}$ In the free theory, β and γ vanish and we recover the trivial result

$$G^{(2)}(p) = \frac{-i}{p^2}$$

In an interacting theory, the solution to the C-S equation can be expressed as

$$G^{(2)}(p, \lambda_0, \mu_0) = G^{(2)}(\mu_0, \lambda_p, \mu_0)\exp\left(-\int_{p'=\mu_0}^{p'=p} d\ln(p'/\mu_0)\cdot 2[1-\gamma_\phi(\lambda_{p'})]\right)$$

Here, $\lambda_0 = \lambda(\mu_0)$, λ_p satisfy the following equation

$$\frac{\partial}{\partial \ln p}\lambda_p(p,\lambda_0) = \beta(\lambda_p) \quad \lambda_p(\mu_0,\lambda_0) = \lambda_0$$

The solution can be checked directly by noticing that

$$\left(p\frac{\partial}{\partial p} - \beta(\lambda_0)\frac{\partial}{\partial \lambda_0}\right)\lambda_p(p,\lambda_0) = 0$$

A convenient way of writing the solution is

$$G^{(2)}(p,\lambda_0,\mu_0) = \frac{-i}{p^2}g^{(2)}(\mu_0,\lambda_p,\mu_0)\exp\left(2\int_{\mu_0}^{p} d\ln(p'/\mu_0)\gamma_\phi(\lambda_{p'})\right)$$

And $g(\lambda_p) = 1 + O(\lambda_p^2)$

Now consider the connected four-point function of ϕ^4 theory evaluated at spacelike momenta p_i such that $p_i^2 = P^2$, $p_i \cdot p_j = 0$, so that s, t, and u are of order $-P^2$. To leading order in perturbation theory, this function is given by

$$G^{(4)}(P) = \left(\frac{i}{P^2}\right)^4(-i\lambda)$$

The solution of C-S equation is

$$G^{(2)}(p,\lambda_0,\mu_0) = \frac{1}{P^8}g^{(4)}(\mu_0,\lambda_p,\mu_0)\exp\left(4\int_{\mu_0}^{p} d\ln(p'/\mu_0)\gamma_\phi(\lambda_{p'})\right)$$

And $g(\lambda_p) = -i\lambda_p + O(\lambda_p^2)$.

The ordinary Feynman perturbation series for a Green's function depends both on the coupling constant λ and on the dimensionless parameter $\ln(p^2/\mu_0^2)$. The perturbation theory can be badly behaved even when λ is small if the ratio p^2/μ_0^2 is large. The solutions of C-S equation reorganize this dependence into a function of the running coupling constant and an exponential scale factor.

The first factor of the solution is a function of the running coupling constant, evaluated at the momentum scale p. If p were of order μ_0, this function would essentially be the ordinary perturbation evaluation of the Green's function. We can make use of this same expression at the scale p, but to replace λ_0 with a new coupling constant λ_p appropriate to that scale. The exponential factor is the accumulated field strength resealing of the correlation function from the reference point μ_0 to the actual momentum p at which the Green's function is evaluated. This factor receives a multiplicative contribution from each intermediate scale between μ and p.

In ϕ^4 theory, we have

$$\frac{d}{d\ln p}\lambda_p = \frac{3\lambda_p^2}{16\pi^2}$$

The solution is

$$\lambda_p = \frac{\lambda_0}{1 - \frac{3\lambda_0}{16\pi^2}\ln(p/\mu_0)}$$

This expression for the running coupling constant goes to zero at a logarithmic rate as $p \to 0$. If we expand the running coupling constant λ_p in powers of λ_0, we find that the successive powers of the coupling constant are multiplied by powers of logarithms,

$$\lambda_0^{n+1}(\ln p/\mu_0)^n$$

which become large and invalidate a simple perturbation expansion for p much greater or much less than μ_0. If the running coupling constant becomes large, as happens in ϕ^4 theory for $p \to \infty$, the perturbation expansion will break down anyway, and we will need more advanced methods. However, if the running coupling constant becomes small, as for ϕ^4 theory as $p \to 0$, we will have successfully organized the powers of logarithms into a meaningful and controlled expression.

12.10.4 Running of coupling constants

In the limit of $\epsilon \to 0$, the coupling constant satisfies the differential equation

$$\frac{\partial \lambda}{\partial \ln \mu} = \beta(\lambda)$$

Three behaviours are possible in the region of small λ:

1. $\beta(\lambda) > 0$
2. $\beta(\lambda) = 0$
3. $\beta(\lambda) < 0$

In theories of the first class, the running coupling constant goes to zero in the infra-red, leading to definite predictions about the small-momentum behaviour of the theory. However, the running coupling constant becomes large in the region of high momenta. Thus the short-distance behaviour of the theory cannot be computed using Feynman diagram perturbation theory.A Feynman diagram analysis is useful in such theories if one is mainly interested in large-distance or macroscopic behaviour.

In theories of the second class, the coupling constant does not flow. In these theories, the running coupling constant is independent of the momentum scale, and thus equal to the bare coupling. This means that there can be no ultraviolet divergences in the relation of coupling constants. The only possible ultraviolet divergences in such theories are those associated with field resealing, which automatically cancel in the computation of S-matrix elements. Such theories are called finite quantum field theories. Before the emergence of our modern understanding of renormalization, these theories would have been embraced as the solution to the problem of ultraviolet infinities. But in fact the known finite field theories in four dimensions are very special constructions the so-called gauge theories with extended supersymmetry with no known physical application.

In theories of the third class, the running coupling constant becomes large in the large-distance regime and becomes small at large momenta or short distances. Such theories are called asymptotically free. In theories of this class, the short-distance behaviour is completely solvable by Feynman diagram methods. Though ultraviolet divergences appear in every order of perturbation theory, the renormalization group tells us that the sum of these divergences is completely harmless.

In the region of strong coupling, the approximation we have made, ignoring the higher-order terms in the β function is no longer valid. It is a logical possibility that the leading order term is positive while the higher terms of the β function are negative, so that the β function has a zero at a non-zero value λ_*. When λ_* approaches this value, the renormalization group flow slows to a halt; thus $\lambda = \lambda_*$ would be a non-trivial fixed point of the renormalization group.

If the β function behaves in the vicinity of the fixed point as

$$\beta \approx -B(\lambda - \lambda_*)$$

where B is a positive constant. For λ near λ_*

$$\frac{d}{d\ln\mu}\lambda \approx -B(\lambda - \lambda_*)$$

The solution of this equation is

$$\lambda(\mu) = \lambda_* + C(\frac{\mu_0}{\mu})^B$$

Thus, λ indeed tends to λ_* as $\mu \to \infty$, and the rate of approach is governed by the slope of the β function at the fixed point. The fixed point here is called ultraviolet-stable fixed point.

If p is sufficiently large, $\lambda(p)$ is close to λ_*.In the massless limit, the solution of C-S equation for two point Green function in momentum space becomes

$$G^{(2)}(p, \lambda_0, \mu_0) \approx G^{(2)}(\mu_0, \lambda_*, \mu_0) \exp\left(-\ln(p/\mu_0) \cdot 2[1 - \gamma_\phi(\lambda_*)]\right)$$

Thus the naive scaling law $G(p) \sim\sim k^{-2}$ is changed to $G(p) \sim k^{-2[1-\gamma_\phi(\lambda_*)]}$. This has applications in the theory of critical phenomena.

A similar behaviour is possible in an asymptotically free theory. If the β function behaves in the vicinity of the fixed point as

$$\beta \approx B(\lambda - \lambda_*)$$

where B is a positive constant. the running coupling constant will tend to a fixed point as $\mu \to 0$. The fixed point is called infrared-stable fixed points.

12.11 Spontaneous symmetry breaking

12.11.1 Effective action

$$Z[J] = e^{-iE[J]} = \int \mathcal{D}\phi \exp\left[i\int d^4x(\mathcal{L}[\phi] + J\phi)\right]$$

Define

$$\phi_{\text{cl}}(x) \equiv \langle\Omega|\phi(x)|\Omega\rangle_J$$

So, we can derive

$$\frac{\delta}{\delta J(x)}E[J] = -\phi_{\text{cl}}(x)$$

Define

$$\Gamma[\phi_{\text{cl}}] \equiv -E[J] - \int d^4y J(y)\phi_{\text{cl}}(y)$$

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.

12.11.2 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:

$$\begin{aligned}&\int \mathcal{D}\eta \exp\left[i\left(\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}}) + \frac{1}{2}\int \eta\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\eta\right)\right]\\ =\;& \exp\left[i\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}})\right]\left(\det\left[\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right]\right)^{-\frac{1}{2}}\end{aligned}$$

Finally, put back the effects of the counter-term Lagrangian,writing it as

$$(\delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\phi_{\text{cl}}) + (\delta\mathcal{L}[\phi_{\text{cl}} + \eta] - \delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\eta)$$

Define

$$\mathcal{L}_2 = \left(\frac{1}{3!}\int d^4x\,d^4y\,d^4z\,\eta(x)\eta(y)\eta(z)\frac{\delta^3\mathcal{L}_1}{\delta\phi(x)\delta\phi(y)\delta\phi(z)} + \cdots\right) + (\delta\mathcal{L}[\phi_{\text{cl}}+\eta] - \delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\eta)$$

So

$$e^{-iE[J]} = \exp\left[i\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}} + \delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\phi_{\text{cl}})\right] e^{i\int\mathcal{L}_2(\frac{1}{i}\frac{\delta}{\delta I})}\int\mathcal{D}\eta e^{i\int\left(\frac{1}{2}\eta\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\eta + I\eta\right)}$$

Therefore, define propagator as

$$D_F \equiv i\left(\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right)^{-1}$$

We have

$$e^{-iE[J]} = \exp\left[i\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}} + \delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\phi_{\text{cl}})\right]\left(\det\left[\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right]\right)^{-\frac{1}{2}} e^{i\int\mathcal{L}_2(\frac{1}{i}\frac{\delta}{\delta I})}\int\mathcal{D}\eta e^{i\int(-\frac{1}{2}ID_FI)}|_{I=0}$$

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,

$$-iE[J] = i\int(\mathcal{L}_1[\phi_{\text{cl}}] + J_1\phi_{\text{cl}} + \delta\mathcal{L}[\phi_{\text{cl}}] + \delta J\phi_{\text{cl}}) - \frac{1}{2}\log\det\left[\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right] + \text{connected diagrams}$$

From this equation, Γ follows directly:

$$\Gamma[\phi_{\text{cl}}] = \int d^4x\mathcal{L}_1[\phi_{\text{cl}}] + \frac{i}{2}\log\det\left[\frac{\delta^2\mathcal{L}_1}{\delta\phi\delta\phi}\right] - i\,\text{connected diagrams} + \int d^4x\delta\mathcal{L}[\phi_{\text{cl}}]$$

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 renormalization conditions on Γ and, in the process, to cancel divergences that appear in the evaluation of the functional determinant and the diagrams. The renormalization 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.

12.11.3 The effective action as a generating functional

$E[J]$ is called the generating of connected correlation functions,

$$\frac{\delta^n E[J]}{\delta J(x_1)\cdots\delta J(x_n)} = i^{n+1}\langle\phi(x_1)\cdots\phi(x_n)\rangle_{\rm conn}$$

The effective action $\Gamma[\phi_{\rm cl}]$ is the generating functional of one-particle-irreducible correlation functional,

$$\frac{\delta\Gamma[\phi_{\rm cl}]}{\delta\phi_{\rm cl}(x)} = 0$$

$$\frac{\delta^2\Gamma[\phi_{\rm cl}]}{\delta\phi_{\rm cl}(x)\delta\phi_{\rm cl}(y)} = iD^{-1}(x,y)$$

Here, $D(x,y) = \langle\phi(x)\phi(y)\rangle_{\rm conn}$. When $n \geq 3$,

$$\frac{\delta^n\Gamma[\phi_{\rm cl}]}{\delta\phi_{\rm cl}(x_1)\cdots\delta\phi_{\rm cl}(x_n)} = -i\langle\phi(x_1)\cdots\phi(x_n)\rangle_{\rm 1PI}$$

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 $\tilde\phi(k) = \int d^dx e^{-ikx}\phi(x)$, 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)*.

12.11.4 Renormalization 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 ϕ_{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 $\partial_\mu\phi_{\text{cl}}$, which appears when the effective action is evaluated for a non-constant background field:

$$\Delta\Gamma[\phi_{\text{cl}}] = \int d^4x B_2^{ij}\partial_\mu\phi_{\text{cl}}^i\partial^\mu\phi_{\text{cl}}^j$$

All other coefficients in the Taylor expansion of the effective action in powers of ϕ_{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 $\Gamma[\phi_{\text{cl}}]$. 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.

12.11.5 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

$$i\tilde{D}_{ij}^{-1}(p^2) = \int d^4x e^{-ip(x-y)}\frac{\delta\Gamma}{\delta\phi^i\delta\phi^j}(x,y)$$

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 $\Gamma[\phi_{\rm cl}]$ 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,

$$\frac{\partial^2}{\partial\phi^i_{\rm cl}\partial\phi^j_{\rm cl}}V_{\rm eff}$$

has a zero eigenvalue. This completes the proof of Goldstone's theorem.

12.12 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 $O(N)$.

Chapter 13
Spin 1/2 Field

13.1 Representation of the Lorentz group

Recall the rotation of the field

$$U^{-1}(\Lambda)\phi_a(x)U(\Lambda) = \mathcal{S}_a{}^b\phi_b(\Lambda^{-1}x)$$

For infinitesimal rotation,

$$\mathcal{S}_a{}^b = \delta_a{}^b + \frac{i}{2}\delta\omega_{\alpha\beta}(S^{\alpha\beta})_a{}^b$$

The matrix $S^{\alpha\beta}$ satisfy the following commutation relation

$$[S^{\mu\nu}, S^{\rho\sigma}] = -\eta^{\nu\rho}S^{\mu\sigma} + \eta^{\sigma\mu}S^{\rho\nu} + \eta^{\mu\rho}S^{\nu\sigma} - \eta^{\sigma\nu}S^{\rho\mu}$$

Define $C_i \equiv \frac{1}{2}\epsilon_{ijk}S^{jk}$,$D_i \equiv S^{i0}$, so we can get

$$[C_i, C_j] = i\epsilon_{ijk}C_k \quad [C_i, D_j] = i\epsilon_{ijk}D_k \quad [D_i, D_j] = -i\epsilon_{ijk}C_k$$

We further define $N_i \equiv \frac{1}{2}(C_i - iD_i)$ and $N_i^\dagger \equiv \frac{1}{2}(C_i + iD_i)$, then the commutation relation will be

$$[N_i, N_j] = i\epsilon_{ijk}N_k \quad [N_i^\dagger, N_j^\dagger] = i\epsilon_{ijk}N_k^\dagger \quad [N_i, N_j^\dagger] = 0$$

We see that we have two different $SU(2)$ Lie algebras that are exchanged by hermitian conjugation. As we just discussed, a representation of the $SU(2)$ Lie algebra is specified by an integer or half integer; we therefore conclude that a representation of the Lie algebra of the Lorentz group in four spacetime dimensions is specified by two integers or half-integers n and n'.

We will label these representations as $(2n+1, 2n'+1)$; the number of components of a representation is then $(2n+1)(2n'+1)$. Different components within a representation can also be labeled by their angular momentum representations. We have $C_i = N_i + N_i^\dagger$. Thus, deducing the allowed values of j given n and n' becomes a standard problem in the addition of angular momenta. The general result is that the allowed values of j are $|n - n'|, |n - n'| + 1, \cdots, n + n'$, and each of these values appears exactly once.

13.2 Spin–statistics theorem

Theorem 13.1 Spin–statistics theorem

States with identical particles of integer spin are symmetric under the interchange of the particles, while states with identical particles of half-integer spin are antisymmetric under the interchange of the particles.
This is equivalent to the statement that the creation and annihilation operators for integer spin particles satisfy canonical commutation relations, while creation and annihilation operators for halfinteger spin particles satisfy canonical anticommutation relations.
Particles quantized with canonical commutation relations are called bosons, and satisfy Bose–Einstein statistics, and particles quantized with canonical anticommutation relations are called fermions, and satisfy Fermi–Dirac statistics.

Roughly speaking, one way to interchange two particles is to rotate them around their midpoint by π. For a particle of spin s, this rotation will introduce a phase factor of $e^{i\pi s}$. Thus, a two-particle state with identical particles both of spin s will pick up a factor of $e^{i2\pi s}$. For s a half-integer, this will give a factor of -1; for s an integer, it will give a factor of $+1$. So, the creation and annihilation operators for integer spin particles satisfy canonical commutation relations, while creation and annihilation operators for halfinteger spin particles satisfy canonical anticommutation relations. The detailed proof can be found in chapter 12.1 and 12.2 from *Quantum Field Theory and the Standard Model (Matthew D. Schwartz)*.

13.3 Spinor field

Consider a left-handed spinor field $\psi_a(x)$, also known as a left-handed Weyl field, which is in the $(2,1)$ representation of the Lie algebra of the Lorentz group. Here the index a is a left-handed spinor index that takes on two possible values. Under a Lorentz transformation, we have

$$U(\Lambda)^{-1}\psi_a(x)U(\Lambda) = L_a{}^b(\Lambda)\psi_b(\Lambda^{-1}x)$$

For an infinitesimal transformation, we can write

$$L_a{}^b(1+\delta\omega) = \delta_a{}^b + \frac{i}{2}\delta\omega_{\mu\nu}(S_L^{\mu\nu})_a{}^b$$

$n=1, \quad n'=0$ implies that

$$(S_L^{ij}) = \frac{1}{2}\epsilon^{ijk}\sigma_k \quad (S_L^{k0}) = \frac{1}{2}i\sigma_k$$

where σ_k is Pauli matrix. Then the infinitesimal transformation can be written as

$$L(1+\delta\omega) = I - \frac{i}{2}\theta_i\sigma_i - \beta_i\sigma_i$$

Recall that hermitian conjugation swaps the two $SU(2)$ Lie algebras that comprise the Lie algebra of the Lorentz group. Therefore, the hermitian conjugate of a field in the $(2,1)$ representation should be a field in the $(1,2)$ representation; such a field is called a right-handed spinor field or a right-handed Weyl field. We will distinguish the indices of the $(1,2)$ representation from those of the $(2,1)$ representation by putting dots over them. Thus, we write

$$[\psi_a(x)]^\dagger = \psi^\dagger_{\dot{a}}(x)$$

Under a Lorentz transformation, we have

$$U(\Lambda)^{-1}\psi^\dagger_{\dot{a}}(x)U(\Lambda) = R_a{}^b(\Lambda)\psi^\dagger_{\dot{b}}(x)(\Lambda^{-1}x)$$

For an infinitesimal transformation, we can write

$$R_{\dot{a}}{}^{\dot{b}}(1+\delta\omega) = \delta_{\dot{a}}{}^{\dot{b}} + \frac{i}{2}\delta\omega_{\mu\nu}(S_R^{\mu\nu})_{\dot{a}}{}^{\dot{b}}$$

We can prove that

$$(S_R^{\mu\nu})_{\dot{a}}{}^{\dot{b}} = -[(S_L^{\mu\nu})_a{}^b]^*$$

Define

$$\epsilon_{ab} \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Group theory can show that

$$L_a{}^c(\Lambda)L_b{}^d(\Lambda)\epsilon_{cd} = \epsilon_{ab}$$

which means that ϵ_{ab} is an invariant symbol of the Lorentz group: it does not change under a Lorentz transformation that acts on all of its indices. The inverse matrix of ϵ_{ab} is

$$\epsilon^{ab} \equiv \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

So,

$$\epsilon_{ab}\epsilon^{bc} = \delta_a{}^c \quad \epsilon^{ab}\epsilon_{bc} = \delta^a{}_c$$

We can use ϵ_{ab} and and its inverse ϵ^{ab} to raise and lower left-handed spinor indices,

$$\psi^a(x) \equiv \epsilon^{ab}\psi_b(x)$$

We also notice the minus sign when we contract indices,

$$\psi^a\chi_a = -\psi_a\chi^a$$

And we can verify the following equations,

$$-L^a{}_bL_a{}^c = \delta_b^c$$

$$\psi_a(x) = \epsilon_{ab}\psi^b(x)$$

$$L^a{}_c(\Lambda)L^b{}_d(\Lambda)\epsilon^{cd} = \epsilon^{ab}$$

$$U(\Lambda)^{-1}\psi^a(x)U(\Lambda) = -L^a{}_b(\Lambda)\psi^b(\Lambda^{-1}x)$$

In right handed representation; we can also deduce the existence of an invariant symbol $\epsilon_{\dot{a}\dot{b}}$. The value of $\epsilon_{\dot{a}\dot{b}}$ is the same as that of ϵ_{ab}.
Define

$$\sigma^{\mu}_{a\dot{a}} \equiv (I, \vec{\sigma})$$

It is an invariant symbol of the group $(2,1) \otimes (1,2) \otimes (2,2)$

The properties of invariance symbol can be used to derive the following equations. The detailed derivation can be found in chapter 35 from *Quantum Field Theory (Mark Srednicki)*.

Proposition 13.1

1.

$$\sigma^{\mu}_{a\dot{a}}\sigma_{\mu b\dot{b}} = -2\epsilon_{ab}\epsilon_{\dot{a}\dot{b}}$$
$$\epsilon^{ab}\epsilon^{\dot{a}\dot{b}}\sigma^{\mu}_{a\dot{a}}\sigma^{\nu}_{b\dot{b}} = -2\eta^{\mu\nu}$$

2.

$$(S_L^{\mu\nu})_{ab} = (S_L^{\mu\nu})_{ba}$$
$$(S_R^{\mu\nu})_{\dot{a}\dot{b}} = (S_R^{\mu\nu})_{\dot{b}\dot{a}}$$

3. Define

$$\overline{\sigma}^{\mu\dot{a}a} \equiv \epsilon^{ab}\epsilon^{\dot{a}\dot{b}}\sigma^{\mu}_{b\dot{b}}$$

Numerically,

$$\overline{\sigma}^{\mu\dot{a}a} = (I, -\vec{\sigma})$$

So,

$$\begin{aligned} (S_L^{\mu\nu})_a{}^b &= +\frac{i}{4}(\sigma^\mu\overline{\sigma}^\nu - \sigma^\nu\overline{\sigma}^\mu)_a{}^b \\ (S_R^{\mu\nu})^{\dot{a}}{}_{\dot{b}} &= -\frac{i}{4}(\overline{\sigma}^\mu\sigma^\nu - \overline{\sigma}^\nu\sigma^\mu)^{\dot{a}}{}_{\dot{b}} \end{aligned}$$

We adopt the following convention: a missing pair of contracted, undotted indices is understood to be written as ${}^c{}_c$, and a missing pair of contracted, dotted indices is understood to be written as ${}_{\dot{c}}{}^{\dot{c}}$. Thus, if χ and ψ are two left-handed Weyl fields, we have

$$\chi\psi = \chi^a\psi_a \quad \chi^\dagger\psi_\dagger = \chi^\dagger_{\dot{a}}\psi^{\dagger\dot{a}}$$

We expect Weyl fields to describe spin-one-half particles, and (by the spinstatistics theorem) these particles must be fermions. Therefore the corresponding fields must anticommute, rather than commute. That is, we should have

$$\chi_a(x)\psi_b(y) = -\psi_b(x)\chi_a(x)$$

Thus, we can get

$$\chi\psi = \psi\chi$$

Using the above convention, we can derive

Proposition 13.2

1.
$$(\chi\psi)^\dagger = \psi^\dagger\chi^\dagger$$

2.
$$[\psi^\dagger\overline{\sigma}^\mu\chi]^\dagger = \chi^\dagger\overline{\sigma}^\mu\psi$$

13.4 Lagrangians for spinor fields

Weyl field

$$\mathcal{L} = i\psi^\dagger\overline{\sigma}^\mu\partial_\mu\psi - \frac{1}{2}m\psi\psi - \frac{1}{2}m\psi^\dagger\psi^\dagger$$

Note:

$$(i\psi^\dagger\overline{\sigma}^\mu\partial_\mu\psi)^\dagger = i\psi^\dagger\overline{\sigma}^\mu\partial_\mu\psi - i\partial_\mu(i\psi^\dagger\overline{\sigma}^\mu\psi)$$

The second term is a total divergence, and vanishes (with suitable boundary conditions on the fields at infinity) when we integrate it over d^4x *to get the action* S. *Thus* $i\psi^\dagger\overline{\sigma}^\mu\partial_\mu\psi$ *has the hermiticity properties necessary for a term in* $\mathcal{L}$.

We can derive the equation of motion from Hamilton principle,

$$-i\overline{\sigma}^\mu\partial_\mu\psi + m\psi^\dagger = 0$$

$$-i\sigma^\mu\partial_\mu\psi^\dagger + m\psi = 0$$

Define gamma matrix as

$$\gamma^\mu \equiv \begin{bmatrix} 0 & \sigma^\mu_{a\dot{c}} \\ \overline{\sigma}^{\mu\dot{a}c} & 0 \end{bmatrix}$$

We can prove that

$$\{\gamma^\mu, \gamma^\nu\} = -2\eta^{\mu\nu}$$

Define a four-component Majorana field as

$$\Psi \equiv \begin{bmatrix} \psi_c \\ \psi^{\dagger\dot{c}} \end{bmatrix}$$

The equation of motion can be written as

$$(-i\gamma^\mu\partial_\mu + m)\Psi = 0$$

Dirac field

$$\mathcal{L} = i\chi^\dagger\overline{\sigma}^\mu\partial_\mu\chi + i\xi^\dagger\overline{\sigma}^\mu\partial_\mu\xi - \frac{1}{2}m\chi\xi - \frac{1}{2}m\xi^\dagger\chi^\dagger$$

It is invariant under

$$\chi \to e^{-i\alpha}\chi \quad \xi \to e^{i\alpha}\xi$$

Define a four-component Dirac field

$$\Psi \equiv \begin{bmatrix}\chi_a \\ \xi^{\dagger\dot{a}}\end{bmatrix}$$

So we take the hermitian conjugate of Ψ to get

$$\Psi^\dagger = (\chi^\dagger_{\dot{a}}, \xi^a)$$

Introduce the matrix

$$\beta \equiv \begin{bmatrix}0 & \delta^{\dot{a}}{}_{\dot{c}} \\ \delta_a{}^c & 0\end{bmatrix}$$

Given β, we define

$$\overline{\Psi} \equiv (\xi^a, \chi^\dagger_{\dot{a}})$$

Detailed calculation shows that

$$\overline{\Psi}\gamma^\mu\partial_\mu\Psi = \chi^\dagger\overline{\sigma}^\mu\partial_\mu\chi + \xi^\dagger\overline{\sigma}^\mu\partial_\mu\xi + \partial_\mu(\xi^\dagger\sigma^\mu\xi)$$

So, the Lagrangian can be written as

$$\mathcal{L} = i\overline{\Psi}\gamma^\mu\partial_\mu\Psi - m\overline{\Psi}\Psi$$

It is invariant under the $U(1)$ transformation

$$\Psi \to e^{-i\alpha}\Psi \quad \overline{\Psi} \to e^{i\alpha}\overline{\Psi}$$

The corresponding Noether current is

$$j^\mu = \overline{\Psi}\gamma^\mu\Psi = \chi^\dagger\overline{\sigma}^\mu\chi - \xi^\dagger\overline{\sigma}^\mu\xi$$

The equation of motion is

$$(-i\gamma^\mu\partial_\mu + m)\Psi = 0$$

Charge conjugation

Charge conjugation simply exchanges ξ and χ. We can define a unitary charge conjugation operator C that implements this,

$$C^{-1}\xi_a(x)C = \chi_a(x)$$

$$C^{-1}\chi_a(x)C = \xi_a(x)$$

We then have $C^{-1}\mathcal{L}(x)C = \mathcal{L}(x)$.
Introduce the charge conjugation matrix

$$\mathcal{C} \equiv \begin{bmatrix} \varepsilon_{ab} & \\ & \varepsilon^{\dot{a}\dot{b}} \end{bmatrix}$$

Take the transpose of Ψ

$$\overline{\Psi}^T = \begin{bmatrix} \xi^a \\ \chi^\dagger_{\dot{a}} \end{bmatrix}$$

Define the charge conjugate of Ψ,

$$\Psi^C \equiv \mathcal{C}\overline{\Psi}^T = \begin{bmatrix} \xi_a \\ \chi^{\dagger\dot{a}} \end{bmatrix}$$

We therefore have

$$C^{-1}\Psi(x)C = \Psi^C(x)$$

for a Dirac field.
The charge conjugation matrix has a number of useful properties.

$$\mathcal{C}^T = \mathcal{C}^\dagger = \mathcal{C}^{-1} = -\mathcal{C}$$

$$\mathcal{C}^{-1}\gamma^\mu\mathcal{C} = -(\gamma^\mu)^T$$

Now let us return to the Majorana field. It is obvious that a Majorana field is its own charge conjugate, that is, $\Psi^C = \Psi$. This condition is analogous to the condition $\phi^\dagger = \phi$ that is satisfied by a real scalar field. A Dirac field, with its $U(1)$ symmetry, is analogous to a complex scalar field, while a Majorana field, which has no $U(1)$ symmetry, is analogous to a real scalar field. For a Majorana field, we have $\overline{\Psi} = \Psi^T\mathcal{C}$. Then the Lagrangian can be written as

$$\mathcal{L} = \frac{i}{2}\Psi^T\mathcal{C}\gamma^\mu\partial_\mu\Psi - \frac{1}{2}m\Psi^T\mathcal{C}\Psi$$

Projection matrix

We can also recover the Weyl components of a Dirac or Majorana field by means of a suitable projection matrix. Define

$$\gamma^5 \equiv \begin{bmatrix} -\delta_a{}^c & 0 \\ 0 & \delta^{\dot{a}}{}_{\dot{c}} \end{bmatrix}$$

Then we can define left and right projection matrices

$$P_L \equiv \frac{1}{2}(1-\gamma^5) = \begin{bmatrix} \delta_a{}^c & 0 \\ 0 & 0 \end{bmatrix}$$

$$P_R \equiv \frac{1}{2}(1+\gamma^5) = \begin{bmatrix} 0 & 0 \\ 0 & \delta^{\dot{a}}{}_{\dot{c}} \end{bmatrix}$$

The matrix γ^5 can also be expressed as

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \frac{i}{24}\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma$$

The γ^5 has the following properties,

$$(\gamma^5)^\dagger = \gamma^5$$
$$(\gamma^5)^2 = 1$$
$$\{\gamma^5, \gamma^\mu\} = 0$$

The behavior of a Dirac field under a Lorentz transformation

Define

$$S^{\mu\nu} \equiv \begin{bmatrix} (S_L^{\mu\nu})_a{}^b & 0 \\ 0 & -(S_R^{\mu\nu})^{\dot{a}}{}_{\dot{b}} \end{bmatrix} = \frac{i}{4}[\gamma^\mu, \gamma^\nu]$$

Numerically, we have

$$S^{i0} = \frac{i}{2}\begin{bmatrix} \sigma^i & \\ & -\sigma^i \end{bmatrix} \quad S^{ij} = \frac{1}{2}\epsilon_{ijk}\begin{bmatrix} \sigma^k & \\ & \sigma^k \end{bmatrix}$$

Then, for either a Dirac or Majorana field Ψ, we can write

$$U(\Lambda)^{-1}\Psi(x)U(\Lambda) = D(\Lambda)\Psi(\Lambda^{-1}x)$$

where, for an infinitesimal transformation,

$$D(1+\delta\omega) = I + \frac{i}{2}\delta\omega_{\mu\nu}S^{\mu\nu}$$

So the $D(\Lambda)$ can be written as

$$D = \begin{bmatrix} L_a{}^b & 0 \\ 0 & -R^{\dot{a}}{}_{\dot{b}} \end{bmatrix}$$

We can verify that

$$U(\Lambda)^{-1}\overline{\Psi}(x)U(\Lambda) = \overline{\Psi}(\Lambda^{-1}x)[D(\Lambda)]^{-1}$$

From the identity

$$\sigma^\mu_{a\dot{a}} = L_a{}^b R_{\dot{a}}{}^{\dot{b}} \Lambda^\mu{}_\nu \sigma^\nu_{b\dot{b}} \quad \overline{\sigma}^{\mu\dot{a}a} = L^a{}_b R^{\dot{a}}{}_{\dot{b}} \Lambda^\mu{}_\nu \overline{\sigma}^{\nu\dot{b}b}$$

We have

$$L^a{}_b R^{\dot{a}}{}_{\dot{b}} \sigma^\mu_{a\dot{a}} = \Lambda^\mu{}_\nu \sigma^\nu_{b\dot{b}} \quad L_a{}^b R_{\dot{a}}{}^{\dot{b}} \overline{\sigma}^{\mu\dot{a}a} = \Lambda^\mu{}_\nu \overline{\sigma}^{\nu\dot{b}b}$$

So we can verify that

$$D^{-1}\gamma^\mu D = \Lambda^\mu{}_\nu \gamma^\nu$$

Recall the commutation relation of $M^{\mu\nu}$ and P^μ, we have

$$[\gamma^\mu, S^{\rho\sigma}] = i(\eta^{\mu\sigma}\gamma^\mu - \eta^{\mu\rho}\gamma^\sigma)$$

We also have

$$[\gamma^5, S^{\mu\nu}] = 0$$

which means that

$$D^{-1}\gamma^5 D = \gamma^5$$

13.5 Canonical quantization of Dirac field

Canonical momentum and Hamiltonian

$$\Pi \equiv \frac{\partial \mathcal{L}}{\partial(\partial_0 \Psi)} = i\overline{\Psi}\gamma^0 = i\Psi^\dagger \quad (\overline{\Psi} = -i\Pi\gamma^0 \quad \Psi^\dagger = -i\Pi)$$
$$\mathcal{H} = -\Pi(\vec{\alpha}\cdot\vec{\nabla} + i\beta m)\Psi \quad (\alpha_i = \gamma^0\gamma^i \quad \beta = \gamma_0)$$
$$H = \int \mathcal{H} d^3x$$

Momentum and angular momentum

$$T^{\mu\nu} = i\overline{\Psi}\left[\eta^{\mu\nu}\gamma^\rho\partial_\rho - \gamma^\mu\partial^\nu\right]\Psi - m\eta^{\mu\nu}\overline{\Psi}\Psi$$
$$P^0 = H \quad P^i = -\int \Pi\nabla^i\Phi\, d^3x$$
$$J_i = -\epsilon_{ijk}\int \Pi(x^j\nabla^k + \frac{i}{2}S^{jk})\Psi\, d^3x$$

Define $\Sigma_i \equiv \frac{1}{2}\epsilon_{ijk}S^{jk}$, so

$$\Sigma_i = \frac{1}{2}\begin{bmatrix}\sigma^i & \\ & \sigma^i\end{bmatrix}$$

The above equation can be written as

$$\vec{J} = -\int \Pi(\vec{x}\times\vec{\nabla} + i\vec{\Sigma})\Psi\, d^3x$$

Canonical quantization

$$\{\Psi_a(\boldsymbol{x},t), \Psi_b(\boldsymbol{x},t)\} = 0$$
$$\{\Psi_a(\boldsymbol{x},t), \Pi^b(\boldsymbol{y},t)\} = i\delta_a^b\delta(\boldsymbol{x}-\boldsymbol{y})$$
$$\{\Psi_a(\boldsymbol{x},t), \Psi^{\dagger b}(\boldsymbol{y},t)\} = \delta_a^b\delta(\boldsymbol{x}-\boldsymbol{y})$$

Solution of Dirac equation

$$\Psi(x) = \sum_{s=\pm}\int \widetilde{dp}\left[b_s(\boldsymbol{p})u_s(\boldsymbol{p})e^{ipx} + d_s^\dagger(\boldsymbol{p})v_s(\boldsymbol{p})e^{-ipx}\right]$$

Here, we introduce the Feynman slash: given any four-vector a^μ, we define

$$\not{a} \equiv a_\mu\gamma^\mu$$

The Dirac equation implies that

$$(\not{p} + m)u(\boldsymbol{p}) = 0$$

$$(-\not{p} + m)v(\boldsymbol{p}) = 0$$

Each of these equations has two solutions, which we label via $s = +$ and $s = -$. For $m \neq 0$, we can go to the rest frame, $\boldsymbol{p} = 0$. We will then distinguish the two solutions by the eigenvalue of the spin matrix Σ_3. Specifically, we will require

$$\Sigma_3 u_\pm(\mathbf{0}) = \pm\frac{1}{2}u_\pm(\mathbf{0})$$

$$\Sigma_3 v_\pm(\mathbf{0}) = \mp\frac{1}{2}v_\pm(\mathbf{0})$$

The solutions are

$$u_+(\mathbf{0}) = \sqrt{m}\begin{pmatrix}1\\0\\1\\0\end{pmatrix} \quad u_-(\mathbf{0}) = \sqrt{m}\begin{pmatrix}0\\1\\0\\1\end{pmatrix}$$

$$v_+(\mathbf{0}) = \sqrt{m}\begin{pmatrix}0\\1\\0\\-1\end{pmatrix} \quad u_-(\mathbf{0}) = \sqrt{m}\begin{pmatrix}-1\\0\\1\\0\end{pmatrix}$$

For later use we also compute the barred spinors

$$\overline{u}_s(\boldsymbol{p}) \equiv u_s^\dagger(\boldsymbol{p})\beta \quad \overline{v}_s(\boldsymbol{p}) \equiv v_s^\dagger(\boldsymbol{p})\beta \quad (\beta = \gamma^0)$$

We get

$$\overline{u}_+(\mathbf{0}) = \sqrt{m}(1,0,1,0)$$

$$\overline{u}_-(\mathbf{0}) = \sqrt{m}(0,1,0,1)$$

$$\overline{v}_+(\mathbf{0}) = \sqrt{m}(0,-1,0,1)$$

$$\overline{v}_-(\mathbf{0}) = \sqrt{m}(1,0,-1,0)$$

We can now find the spinors corresponding to an arbitrary three-momentum $\boldsymbol{p}$ by applying to $u_s(\mathbf{0})$ and $v_s(\mathbf{0})$ the matrix $D(\Lambda)$ that corresponds to an appropriate boost. This is given by

$$D(\Lambda) = \exp(i\eta\hat{\boldsymbol{p}}\cdot\boldsymbol{D})$$

where $\hat{\boldsymbol{p}}$ is a unit vector in the $\boldsymbol{p}$ direction, $D^i = \frac{i}{4}[\gamma^i,\gamma^0] = \frac{i}{2}\gamma^i\gamma^0$ is the boost matrix, and η is the rapidity. Thus we have

$$u_s(\boldsymbol{p}) = \exp(i\eta\hat{\boldsymbol{p}}\cdot\boldsymbol{D})u_s(\mathbf{0})$$

$$v_s(\boldsymbol{p}) = \exp(i\eta\hat{\boldsymbol{p}}\cdot\boldsymbol{D})v_s(\mathbf{0})$$

We also have

$$\overline{u}_s(\boldsymbol{p}) = \overline{u}_s(\mathbf{0})\exp(-i\eta\hat{\boldsymbol{p}}\cdot\boldsymbol{D})$$

$$\overline{v}_s(\boldsymbol{p}) = \overline{v}_s(\mathbf{0})\exp(-i\eta\hat{\boldsymbol{p}}\cdot\boldsymbol{D})$$

This follows from $\overline{K^j} = K^j$, where, for any general combination of gamma matrices,

$$\overline{A} \equiv \beta A^\dagger \beta$$

In particular, it turns out that

$$\gamma^\mu \quad S^{\mu\nu} \quad i\gamma^5 \quad \gamma^\mu\gamma^5 \quad i\gamma^5 S^{\mu\nu}$$

all satisfy $\overline{A} = A$ The barred spinors satisfy the equations

$$\overline{u}_s(\boldsymbol{p})(\not{p} + m) = 0 \quad \overline{v}_s(\boldsymbol{p})(-\not{p} + m) = 0$$

Proposition 13.3

1.

$$\overline{u}_{s'}(\boldsymbol{p})u_s(\boldsymbol{p}) = 2m\delta_{ss'}$$
$$\overline{v}_{s'}(\boldsymbol{p})v_s(\boldsymbol{p}) = -2m\delta_{ss'}$$
$$\overline{u}_{s'}(\boldsymbol{p})v_s(\boldsymbol{p}) = 0$$
$$\overline{v}_{s'}(\boldsymbol{p})u_s(\boldsymbol{p}) = 0$$

2.

$$2m\overline{u}_{s'}(\boldsymbol{p}')\gamma^\mu u_s(\boldsymbol{p}) = \overline{u}_{s'}[(p' + p)^\mu - 2iS^{\mu\nu}(p' - p)_\nu]u_s(\boldsymbol{p})$$
$$-2m\overline{v}_{s'}(\boldsymbol{p}')\gamma^\mu v_s(\boldsymbol{p}) = \overline{v}_{s'}[(p' + p)^\mu - 2iS^{\mu\nu}(p' - p)_\nu]v_s(\boldsymbol{p})$$

3.

$$\overline{u}_{s'}(\boldsymbol{p})\gamma^\mu u_s(\boldsymbol{p}) = 2p^\mu\delta_{ss'}$$
$$\overline{v}_{s'}(\boldsymbol{p})\gamma^\mu v_s(\boldsymbol{p}) = 2p^\mu\delta_{ss'}$$
$$\overline{u}_{s'}(\boldsymbol{p})\gamma^0 v_s(-\boldsymbol{p}) = 0$$
$$\overline{v}_{s'}(\boldsymbol{p})\gamma^0 u_s(-\boldsymbol{p}) = 0$$

4.

$$\sum_{s=\pm} u_s(\boldsymbol{p})\overline{u}_s(\boldsymbol{p}) = -\not{p} + m$$
$$\sum_{s=\pm} v_s(\boldsymbol{p})\overline{v}_s(\boldsymbol{p}) = -\not{p} - m$$

5.

$$u_s(\boldsymbol{p})\overline{u}_s(\boldsymbol{p}) = \frac{1}{2}(1 - s\gamma^5\not{z})(-\not{p} + m)$$
$$v_s(\boldsymbol{p})\overline{v}_s(\boldsymbol{p}) = \frac{1}{2}(1 - s\gamma^5\not{z})(-\not{p} - m)$$

Here, z^μ is the boost of the vector $(0, 0, 0, 1)$ from the frame with $\boldsymbol{p}' = 0$ to the frame with $\boldsymbol{p}' = \boldsymbol{p}$.

The proof can be found in chapter 38 from *Quantum field theory (Mark Srednicki)*

It is interesting to consider the extreme relativistic limit of this formula. Let us take the three-momentum to be in the z direction, so that it is parallel to the spin-quantization axis. The component of the spin in the direction of the three-momentum is called the helicity. A fermion with helicity $+1/2$ is said to be right-handed, and a fermion with helicity $-1/2$ is said to be left-handed. For rapidity η, we have

$$\frac{p^\mu}{m} = (\cosh\eta, 0, 0, \sinh\eta) \quad z^\mu = (\sinh\eta, 0, 0, \cosh\eta)$$

In the limit of large η,

$$z^\mu = \frac{p^\mu}{m} + \mathcal{O}(e^{-\eta})$$

Then the equations 3.3.5 can be written as

$$u_s(\boldsymbol{p})\overline{u}_s(\boldsymbol{p}) \to \frac{1}{2}(1 + s\gamma^5)(-\not{p})$$

$$v_s(\boldsymbol{p})\overline{v}_s(\boldsymbol{p}) \to \frac{1}{2}(1 - s\gamma^5)(-\not{p})$$

So, we can see the spinor corresponding to a right-handed fermion is $u_+(\boldsymbol{p})$ for a b-type particle and $v_-\boldsymbol{p}$ for a d-type particle. The spinor corresponding to a left-handed fermion is $u_-(\boldsymbol{p})$ for a b-type particle and $v_+\boldsymbol{p}$ for a d-type particle.

Note that $\beta u_s(\mathbf{0}) = +u_s(\mathbf{0})$ and $\beta v_s(\mathbf{0}) = -v_s(\mathbf{0})$. Also, $\beta D^j = -D^j\beta$. We then have

$$u_s(-\boldsymbol{p}) = \beta u_s(\boldsymbol{p}) \quad v_s(-\boldsymbol{p}) = -\beta v_s(\boldsymbol{p})$$

For charge conjugation matrix, note that that $\mathcal{C}\overline{u}_s^T(\mathbf{0}) = v_s(\mathbf{0})$, $\mathcal{C}\overline{v}_s^T(\mathbf{0}) = v_s(\mathbf{0})$, and $\mathcal{C}^{-1}D^j\mathcal{C} = -(D^j)^T$, we have

$$\mathcal{C}\overline{u}_s^T(\boldsymbol{p}) = v_s(\boldsymbol{p}) \quad \mathcal{C}\overline{v}_s^T(\boldsymbol{p}) = u_s(\boldsymbol{p})$$

Taking the complex conjugate of the above equation, we have

$$u_s^*(\boldsymbol{p}) = \mathcal{C}\beta v_s(\boldsymbol{p}) \quad v_s^*(\boldsymbol{p}) = \mathcal{C}\beta u_s(\boldsymbol{p})$$

Note that $\gamma^5 u_s(\mathbf{0}) = +sv_{-s}(\mathbf{0})$ and $\gamma^5 v_s(\mathbf{0}) = -su_{-s}(\mathbf{0})$, and that $\gamma^5 D^j = D^j\gamma^5$, we have

$$\gamma^5 u_s(\boldsymbol{p}) = +sv_{-s}(\boldsymbol{p}) \quad \gamma^5 v_s(\boldsymbol{p}) = -su_{-s}(\boldsymbol{p})$$

Combine the above equations, we can derive

$$u_{-s}^*(\boldsymbol{p}) = -s\mathcal{C}\gamma^5 u_s(\boldsymbol{p}) \quad v_{-s}^*(\boldsymbol{p}) = -s\mathcal{C}\gamma^5 v_s(\boldsymbol{p})$$

Fourier expansion

$$\Psi(x) = \sum_{s=\pm}\int \widetilde{dp}\left[b_s(\boldsymbol{p})u_s(\boldsymbol{p})e^{ipx} + d_s^\dagger(\boldsymbol{p})v_s(\boldsymbol{p})e^{-ipx}\right]$$

$$\Pi(x) = i\sum_{s=\pm}\int \widetilde{dp}\left[b_s^\dagger(\boldsymbol{p})u_s^\dagger(\boldsymbol{p})e^{-ipx} + d_s(\boldsymbol{p})v_s^\dagger(\boldsymbol{p})e^{+ipx}\right]$$

$$
\begin{aligned}
b_s(\boldsymbol{p}) &= \int d^3x e^{-ipx}\overline{u}_s(\boldsymbol{p})\gamma^0\Psi(x) \\
b_s^\dagger(\boldsymbol{p}) &= \int d^3x e^{ipx}\overline{\Psi}(x)\gamma^0 u_s(\boldsymbol{p}) \\
d_s(\boldsymbol{p}) &= \int d^3x e^{-ipx}\overline{\Psi}(x)\gamma^0 v_s(\boldsymbol{p}) \\
d_s^\dagger(\boldsymbol{p}) &= \int d^3x e^{ipx}\overline{v}_s(\boldsymbol{p})\gamma^0\Psi(x)
\end{aligned}
$$

We can get the anticommutation relation in terms of $b, b^\dagger, d, d^\dagger$. The only non-vanishing terms are

$$
\begin{aligned}
\{b_s(\boldsymbol{p}), b_{s'}^\dagger(\boldsymbol{p}')\} &= (2\pi)^3\delta^3(\boldsymbol{p}-\boldsymbol{p}')2\omega\delta_{ss'} \\
\{d_s(\boldsymbol{p}), d_{s'}^\dagger(\boldsymbol{p}')\} &= (2\pi)^3\delta^3(\boldsymbol{p}-\boldsymbol{p}')2\omega\delta_{ss'}
\end{aligned}
$$

Operator represented by $b, b^\dagger, d, d^\dagger$

Define

$$
N^+(\boldsymbol{p}, s) = b_s^\dagger(\boldsymbol{p})b_s(\boldsymbol{p}) \quad N^-(\boldsymbol{p}, s) = d_s^\dagger(\boldsymbol{p})d_s(\boldsymbol{p})
$$

So we can derive

$$
\begin{aligned}
H &= \sum_{s=\pm}\int \widetilde{dp}\,\omega\left[N^+(\boldsymbol{p}, s) + N^-(\boldsymbol{p}, s)\right] - 4\mathcal{E}_0 V \\
P^i &= \sum_{s=\pm}\int \widetilde{dp}\,p^i\left[N^+(\boldsymbol{p}, s) + N^-(\boldsymbol{p}, s)\right] \\
S_3 &= \sum_{s=\pm}\int \widetilde{dp}\,\frac{s}{2}\left[N^+(\boldsymbol{p}, s) + N^-(\boldsymbol{p}, s)\right] \\
Q &= \sum_{s=\pm}\int \widetilde{dp}\,\left[N^+(\boldsymbol{p}, s) - N^-(\boldsymbol{p}, s)\right]
\end{aligned}
$$

Causality

Firstly, we derive the anticommutation relation for field operators at any space-time.

$$
\{\overline{\Psi}_a(x), \Psi_b(y)\} = (i\partial\!\!\!/_x + m)_{ab} i\Delta(x-y)
$$

Here,

$$
i\Delta(x-y) = \int \widetilde{dp}\,[e^{ip(x-y)} - e^{-ip(x-y)}]
$$

For $(x-y)^2 > 0$ the anti-commutators vanish, because $\Delta(x-y)$ also vanishes. We then can verify that

$$
[\overline{\Psi}_a(x)\Psi_b(x), \overline{\Psi}_c(y)\Psi_d(y)] = 0
$$

for $(x-y)^2 > 0$.

In this way the microscopic causality is satisfied for the physical observables, such as the charge density or the momentum density.

The Dirac propagator

$$\langle 0|\Psi_a(x)\overline{\Psi}_b(y)|0\rangle = (i\partial\!\!\!/_x + m)_{ab}\int \widetilde{dp}\, e^{ip(x-y)}$$
$$\langle 0|\overline{\Psi}_b(y)\Psi_a(x)|0\rangle = -(i\partial\!\!\!/_x + m)_{ab}\int \widetilde{dp}\, e^{ip(y-x)}$$

Define retarded green function as

$$S_R(x-y)_{ab} \equiv \theta(x^0-y^0)\langle 0|\{\Psi_a(x)\overline{\Psi}_b(y)\}|0\rangle$$

It is easy to verify that

$$S_R(x-y) = (i\partial\!\!\!/_x + m)D_R(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{i(\not p - m)}{p^2+m^2}e^{ip(x-y)}$$

and

$$(i\partial\!\!\!/_x - m)S_R(x-y) = i\delta(x-y)\cdot \mathbf{1}_{4\times 4}$$

Now, we define the time ordered product for fermion fields

$$T\eta(x)\eta(y) \equiv \theta(x^0-y^0)\eta(x)\eta(y) - \theta(y^0-x^0)\eta(y)\eta(x)$$

So,

$$S_F(x-y) \equiv \langle 0|T\Psi(x)\overline{\Psi}(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{i(\not p - m)}{p^2+m^2-i\epsilon}e^{ip(x-y)}$$

It is easy to verify that

$$\langle 0|T\overline{\Psi}_a(x)\Psi_b(y)|0\rangle = -\langle 0|T\Psi_b(y)\overline{\Psi}_a(x)|0\rangle = -S_F(y-x)_{ba}$$

13.6 Parity,time reversal and charge conjugation

Parity

We assume that

$$P^{-1}b_s(\boldsymbol{p})P = \eta_b b_s(-\boldsymbol{p})$$
$$P^{-1}d_s(\boldsymbol{p})P = \eta_d d_s(-\boldsymbol{p})$$

Here, η_b and η_d is the phase factor. Then, we can verify that

$$P^{-1}\boldsymbol{P}P = -\boldsymbol{P} \quad P^{-1}\boldsymbol{S}P = \boldsymbol{S}$$

So a parity transformation reverse the three momentum while leaving the spin direction unchanged. Further more, we have

$$P^{-1}\Psi(x)P = \sum_{s=\pm}\int \widetilde{dp}\,\left[\eta_b b_s(\boldsymbol{p})\beta u_s(\boldsymbol{p})e^{ip\mathcal{P}x} - \eta_d^* d_s^\dagger(\boldsymbol{p})\beta v_s(\boldsymbol{p})e^{-ip\mathcal{P}x}\right]$$

where,$\mathcal{P}^\mu{}_\nu = \mathrm{diag}(1,-1,-1,-1)$. So, we should demand that $\eta_b = -\eta_d^*$. And we can get

$$P^{-1}\Psi(x)P = \eta_b\beta\Psi(\mathcal{P}x) \quad P^{-1}\overline{\Psi}(x)P = \eta_b^*\overline{\Psi}(\mathcal{P}x)\beta$$

Generally, we have

$$P^{-1}(\overline{\Psi}A\Psi)P = \overline{\Psi}(\beta A\beta)\Psi$$

Time reversal

Note in quantum mechanics, we have shown that time reversal operator is antiunitary. Firstly, we assume that

$$T^{-1}b_s(\boldsymbol{p})T = \zeta_{b,s}b_{-s}(-\boldsymbol{p})$$

$$T^{-1}d_s(\boldsymbol{p})T = \zeta_{d,s}d_{-s}(-\boldsymbol{p})$$

Then we can verify that

$$T^{-1}\boldsymbol{P}T = -\boldsymbol{P} \quad T^{-1}\boldsymbol{S}T = -\boldsymbol{S}$$

So a parity transformation reverse the three momentum and the spin direction. Further more, we have

$$T^{-1}\Psi(x)T = \sum_{s=\pm}\int \widetilde{dp} \; - s\mathcal{C}\gamma^5\left[\zeta_{b,-s}b_s(\boldsymbol{p})u_s(\boldsymbol{p})e^{ip\mathcal{T}x} + \zeta^*_{d,-s}d^\dagger_s(\boldsymbol{p})\beta v_s(\boldsymbol{p})e^{-ip\mathcal{T}x}\right]$$

where,$\mathcal{T}^\mu_{\ \nu} = \mathrm{diag}(-1,1,1,1)$. So, we should demand that $\zeta_{s,b} = s\zeta$ and $\zeta_{s,d} = s\zeta^*$. And we can get

$$T^{-1}\Psi(x)T = \zeta\mathcal{C}\gamma^5\Psi(\mathcal{T}x) \quad T^{-1}\overline{\Psi}(x)T = \zeta^*\overline{\Psi}(\mathcal{T}x)\gamma^5\mathcal{C}^{-1}$$

Generally, we have

$$T^{-1}(\overline{\Psi}A\Psi)T = \overline{\Psi}(\gamma^5\mathcal{C}^{-1}A\mathcal{C}\gamma^5)\Psi$$

Charge conjugation

We have already shown that

$$C^{-1}\Psi(x)C = \mathcal{C}\overline{\Psi}^T(x) \quad C^{-1}\overline{\Psi}(x)C = \Psi^T(x)\mathcal{C}$$

Generally, we have

$$C^{-1}(\overline{\Psi}A\Psi)C = \overline{\Psi}(\mathcal{C}^{-1}A^T\mathcal{C})\Psi$$

Summary

The transformation properties of the various fermion bilinears under C,P and T are summarized in the table below. Here we use the shorthand $(-1)^\mu \equiv= 1$ for $\mu = 0$ and $(-1)^\mu \equiv= -1$ for $\mu = 1,2,3$.

	$\bar{\psi}\psi$	$i\bar{\psi}\gamma^5\psi$	$\bar{\psi}\gamma^\mu\psi$	$\bar{\psi}\gamma^\mu\gamma^5\psi$	$\bar{\psi}\sigma^{\mu\nu}\psi$	∂_μ
P	$+1$	-1	$(-1)^\mu$	$-(-1)^\mu$	$(-1)^\mu(-1)^\nu$	$(-1)^\mu$
T	$+1$	-1	$(-1)^\mu$	$(-1)^\mu$	$-(-1)^\mu(-1)^\nu$	$-(-1)^\mu$
C	$+1$	$+1$	-1	$+1$	-1	$+1$
CPT	$+1$	$+1$	-1	-1	$+1$	-1

Figure 13.1: Transformation of fermion bilinears under CPT

Theorem 13.2 CPT theorem

Any hermitian combination of any set of fields (scalar, vector, Dirac, Majorana) and their derivatives that is a Lorentz scalar (and so carries no indices) is even under CPT. Since the lagrangian must be formed out of such combinations, we have $\mathcal{L}(x) \to \mathcal{L}(-x)$ under CPT, and so the action $S = \int d^4x\mathcal{L}$ is invariant.

13.7 Perturbation theory for canonical quantization

We use Yukawa theory as an example.

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}M_0^2\phi^2 + i\overline{\Psi}\gamma^\mu\partial_\mu\Psi - m_0\overline{\Psi}\Psi - g_0\overline{\Psi}\Psi\phi$$

$$H = H_0 + H_{int} \quad H_{int} = \int d^3x\, g_0\overline{\Psi}\Psi\phi$$

Similar to ϕ^4 theory, the perturbation expansion of correlation functions is

$$\langle\Omega|T\{\Psi(x)\overline{\Psi}(y)\phi(z)\}|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\langle 0|T\left\{\Psi_I(x)\overline{\Psi}_I(y)\phi_I(z)\exp\left[-i\int_{-T}^{T} dtH_I\right]\right\}|0\rangle}{\langle 0|T\left\{\exp\left[-i\int_{-T}^{T} dtH_I\right]\right\}|0\rangle}$$

Before we state the Wick's theorem, we must note the following conventions:

1. The time-ordered product picks up one minus sign for each interchange of operators that is necessary to put the fields in time order.
2. The normal-ordered product picks up one minus sign for each interchange of operators that is necessary to put the fields in normal order.
3. Define contractions under the normal-ordering symbol to include minus signs for operator interchanges.

With these conventions, Wick's theorem takes the same form as before:

$$T\left\{\Psi_I(x_1)\overline{\Psi}_I(x_2)\Psi_I(x_3)\cdots\right\} = N\left\{\Psi_I(x_1)\overline{\Psi}_I(x_2)\Psi_I(x_3)\cdots + \text{ all possible contractions }\right\}$$

Example:

$$\begin{aligned}\langle 0|T\left\{\Psi_{Ia}(x_1)\overline{\Psi}_{Ib}(x_2)\Psi_{Ic}(x_3)\overline{\Psi}_{Id}(x_4)\right\}|0\rangle &= S_F(x_1-x_2)_{ab}S_F(x_3-x_4)_{cd} \\ &- S_F(x_1-x_4)_{ad}S_F(x_3-x_2)_{cb}\end{aligned}$$

Then we can derive the Feynman rule for Yukawa theory. Expand

$$\langle 0|T\left\{\Psi_{Ia}(x)\overline{\Psi}_{Ib}(y)\phi_I(z)\exp\left[-i\int_{-T}^{T} dtH_I\right]\right\}|0\rangle$$

to the first order of g_0

$$\begin{aligned}
&\langle 0|T\left\{\Psi_{Ia}(x)\overline{\Psi}_{Ib}(y)\phi_I(z)(-ig_0)\int dw^4\overline{\Psi}_I(w)\Psi_I(w)\phi_I(w)\right\}|0\rangle \\
= \quad & -(-ig_0)S_F(x-y)_{ab}\int d^4wD_F(z-w)\mathrm{Tr}[S_F(w-w)] \\
+ \quad & (-ig_0)\int d^4w[S_F(x-w)S_F(w-y)]_{ab}D_F(w-z)
\end{aligned}$$

It can be represented by the so called Feynman diagram.

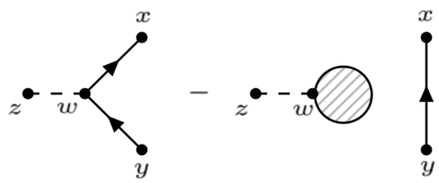


Figure 13.2: Feynman diagram representation of perturbation expansion

The Feynman rules for Yukama theory are:

1. For each Fermion propagator from y to x, $P = S_F(x-y)$
2. For each scalar propagator, $P = D_F(x-y)$
3. For each vertex, $V = (-ig_0)\int d^4w$
4. For each external point, $E = 1$
5. Divided by the symmetry factor

13.8 Path integral quantization

13.8.1 Grassmann numbers

Formal definition

Grassmann numbers are individual elements or points of the exterior algebra generated by a set of n Grassmann variables or Grassmann directions or supercharges $\{\theta_i\}$ with n possibly being infinite.

The Grassman variables are the basis vectors of a vector space (of dimension n) They form an algebra over a field, with the field usually being taken to be the complex numbers, although one could contemplate other fields, such as the reals. The algebra is a unital algebra, and the generators are anti-commuting:

$$\theta_i\theta_j = -\theta_j\theta_i$$

Since the θ_i form a vector space over the complex numbers, it is trivial that they commute with the complex numbers; this is by definition. That is, for complex x, one has

$$\theta_i x = x\theta_i$$

The squares of the generators vanish:

$$\theta_i\theta_i = -\theta_i\theta_i$$

In other words, a Grassmann variable is a non-zero square-root of zero.

Let V denote this n-dimensional vector space of Grassman variables. Note that it is independent of the choice of basis. The corresponding exterior algebra is defined as

$$\Lambda = \mathbb{C} \oplus V \oplus (V \wedge V) \oplus (V \wedge V \wedge V) \oplus \cdots$$

where $\wedge$ is the exterior product and $\oplus$ is the direct sum. The individual elements of this algebra are then called Grassman numbers. It is standard to completely omit the wedge symbol $\wedge$ when writing a Grassman number; it is used here only to clearly illustrate how the exterior algebra is built up out of the Grassman variables. Thus, a completely general Grassman number can be written as

$$z = \sum_{k=0}^{\infty} \sum_{i_1, i_2, \cdots, i_k} c_{i_1 i_2 \cdots i_k} \theta_{i_1}\theta_{i_2} \cdots \theta_{i_k}$$

where the c's are complex numbers, or, equivalently, $c_{i_1 i_2 \cdots i_k}$ is a complex-valued, completely antisymmetric tensor of rank k. Again, the θ_i can be clearly seen here to be playing the role of a basis vector of a vector space.

Observe that the Grassmann algebra generated by n linearly independent Grassmann variables has dimension $2n$; this follows from the binomial theorem applied to the above sum, and the fact that the $n + 1$-fold product of variables must vanish, by the anti-commutation relations, above. In other words, for n variables, the sum terminates

$$\Lambda = \mathbb{C} \oplus \Lambda^1 V \oplus \Lambda^2 V \oplus \cdots \oplus \Lambda^n V$$

where $\Lambda^k V$ is the k-fold alternating product. The dimension of $\Lambda^k V$ is given by n choose k, the binomial coefficient. The special case of $n = 1$ is called a dual number, and was introduced by William Clifford in 1873.

Integral over Grassmann number

Single Variable:

$$\int d\theta (A + B\theta) \equiv B$$

Multi-variable:

$$\int d\theta d\eta \; \eta\theta \equiv 1$$

Complex Grassmann number:

$$(\theta\eta)^* \equiv \eta^*\theta^* = -\theta^*\eta^*$$

$$\int d\theta^* d\theta \; \theta\theta^* \equiv 1$$

Gaussian integral over a complex Grassmann number:

$$\int d\theta^* d\theta e^{-\theta^* b\theta} = b$$

$$\int d\theta^* d\theta \; \theta\theta^* e^{-\theta^* b\theta} = 1$$

Unitary transformation:
If $\theta'_i = U_{ij}\theta_j$ and U is unitary matrix, then we can derive

$$\prod_i \theta'_i = (\det U)(\prod_i \theta_i)$$

In a general integral

$$(\prod_i \int d\theta_i^* d\theta) f(\theta)$$

the only term of $f(\theta)$ that survives has exactly one factor of each θ_i and θ_i^*; it is proportional to $(\prod_i \theta_i)(\prod_i \theta_i^*)$. If we replace θ by $U\theta$, this term acquires a factor of $\det U \det U^* = 1$, so the integral is unchanged under the unitary transformation.

Gaussian integral over multiple complex Grassmann numbers:

$$(\prod_i \int d\theta_i^* d\theta) e^{-\theta^* B_{ij}\theta_j} = \det B$$

$$(\prod_i \int d\theta_i^* d\theta)\theta_k\theta_l^* e^{-\theta^* B_{ij}\theta_j} = B_{kl}^{-1} \det B$$

13.8.2 Path integral formulation for free Dirac field

A Grassmann field is a function of space-time whose values are Grassmann numbers. The classical Dirac field being used to evaluate path integral is a Grassmann field. The correlation function is given by

$$\langle\Omega|T\Psi_H(x_1)\overline{\Psi}_H(x_2)|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\overline{\Psi}\mathcal{D}\Psi \; \exp\left[i\int_{-T}^{T} d^4x\overline{\Psi}(i\partial\!\!\!/ - m)\Psi\right] \Psi(x_1)\overline{\Psi}(x_2)}{\int \mathcal{D}\overline{\Psi}\mathcal{D}\Psi \; \exp\left[i\int_{-T}^{T} d^4x\overline{\Psi}(i\partial\!\!\!/ - m)\Psi\right]}$$

The generating function is

$$Z[\overline{\eta},\eta] = \int \mathcal{D}\overline{\Psi}\mathcal{D}\Psi \; \exp\left[i\int d^4x\overline{\Psi}(i\partial\!\!\!/ - m)\Psi + \overline{\eta}\Psi + \overline{\Psi}\eta\right]$$

Here, $\eta(x)$ is a Grassmann-valued source field. Define

$$\Psi'(x) \equiv \Psi(x) - i\int d^4y S_F(x-y)\eta(y)$$

Then we can derive that

$$\overline{\Psi}'(x) \equiv \overline{\Psi}(x) - i\int d^4y\overline{\eta}(y)S_F(y-x)$$

Recall that

$$(i\partial\!\!\!/_x - m)S_F(x-y) = i\delta(x-y)$$

and

$$S_F(y-x)(i\overleftarrow{\partial\!\!\!/}_x+m)=-i\delta(x-y)$$

we can derive that

$$\int d^4x\overline{\Psi}(i\partial\!\!\!/-m)\Psi+\overline{\eta}\Psi+\overline{\Psi}\eta=\int d^4x\overline{\Psi}'(i\partial\!\!\!/-m)\Psi'+i\int d^4xd^4y\overline{\eta}(x)S_F(x-y)\eta(y)$$

After integration, we have

$$Z[\overline{\eta},\eta]=Z_0\mathrm{exp}\left[-\int d^4xd^4y\ \overline{\eta}(x)S_F(x-y)\eta(y)\right]$$

If we adopt the convention that

$$\frac{d}{d\eta}\theta\eta=-\frac{d}{d\eta}\eta\theta=-\theta$$

the two point correlation functions are

$$\langle 0|T\Psi_H(x_1)\overline{\Psi}_H(x_2)|0\rangle=Z_0^{-1}\left(-i\frac{\delta}{\delta\overline{\eta}(x_1)}\right)\left(i\frac{\delta}{\delta\eta(x_2)}\right)Z[\overline{\eta},\eta]|_{\overline{\eta},\eta=0}=S_F(x_1-x_2)$$

13.8.3 Perturbation theory for path integral quantization

We use Yukawa theory as an example.

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi-\frac{1}{2}M_0^2\phi^2+i\overline{\Psi}\gamma^\mu\partial_\mu\Psi-m_0\overline{\Psi}\Psi-g_0\overline{\Psi}\Psi\phi \\
\mathcal{L} &= \mathcal{L}_0+\mathcal{L}_1 \quad \mathcal{L}_1=-g_0\overline{\Psi}\Psi\phi \\
Z[J] &= \int\mathcal{D}\phi\mathcal{D}\overline{\Psi}\mathcal{D}\Psi e^{i\int d^4x[\mathcal{L}_0+\mathcal{L}_1+J\phi+\overline{\eta}\Psi+\overline{\Psi}\eta]} \\
&= e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)},\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)},-\frac{1}{i}\frac{\delta}{\delta\eta(x)})}\int\mathcal{D}\phi\mathcal{D}\overline{\Psi}\mathcal{D}\Psi e^{i\int d^4y[\mathcal{L}_0+J\phi+\overline{\eta}\Psi+\overline{\Psi}\eta]} \\
&\propto e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)},\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)},-\frac{1}{i}\frac{\delta}{\delta\eta(x)})}\mathrm{exp}[-\int d^4yd^4z\frac{1}{2}J(y)D_F(y-z)J(z)+\overline{\eta}(y)S_F(y-z)\eta(z)] \\
&= \sum_{V=0}^{\infty}\frac{1}{V!}[-ig_0\int d^4x(\frac{1}{i}\frac{\delta}{\delta J(x)}\cdot-\frac{1}{i}\frac{\delta}{\delta\eta(x)}\cdot\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)})]^V \\
&\times\sum_{P_1=0}^{\infty}\frac{1}{P_1!}[-\frac{1}{2}\int d^4y_1d^4z_1J(y_1)D_F(y_1-z_1)J(z_1)]^{P_1} \\
&\times\sum_{P_2=0}^{\infty}\frac{1}{P_2!}[-\int d^4y_2d^4z_2\overline{\eta}(y_2)S_F(y_2-z_2)\eta(z_2)]^{P_2}
\end{aligned}$$

If we focus on a term with particular values of V, P_1 and P_2, then the number of surviving scalar sources is $E_1=2P_1-V$, the number of surviving fermion sources is $E_2=2P_2-2V$. We can introduce Feynman diagrams as in the ϕ^4 theory. In these diagrams, a dashed line segment stands for a scalar propagator $D_F(x-y)$, a line with an arrow pointing from y to x for a fermion propagator $S_F(x-y)$, a filled circle at one end of a dashed line segment for a scalar source $i\int d^4xJ(x)$, a filled circle at the start of a line with an arrow for a fermion source $i\int d^4x\eta(x)$, a filled circle at the end of a line with an arrow for a anti-fermion source $i\int d^4x\overline{\eta}(x)$, a vertex joining three line segments for $-ig_0\int d^4x$.

13.9 LSZ reduction formula

Similarly to the scalar field theory, we can derive the structure of the exact propagator of Dirac fermions in interaction theory.

$$\int d^4x\, e^{-ipx}\langle\Omega|T\Psi(x)\overline{\Psi}(0)|\Omega\rangle_C = \frac{iZ_2(\not{p}-m)}{p^2+m^2-i\epsilon}+\cdots$$

We eliminate the terms contributing none isolate poles for p^2. Here, m is the exact mass of the fermions. The constant Z_2 is the probability for the quantum field to create or annihilate an exact one-particle eigenstate of H:

$$\langle\Omega|\Psi(0)|p,s\rangle = \sqrt{Z_2}u^s(p)$$

The LSZ reduction formula for fermions would take the form as

Theorem 13.3 LSZ reduction formula

$$\begin{aligned}
&\langle \boldsymbol{p}_1\cdots\boldsymbol{p}_n\,\overline{\boldsymbol{p}}_1\cdots\overline{\boldsymbol{p}}_{\overline{n}}\,|S|\,\boldsymbol{k}_1\cdots\boldsymbol{k}_m\,\overline{\boldsymbol{k}}_1\cdots\overline{\boldsymbol{k}}_{\overline{n}}\rangle\\
=\;&\prod_1^n\int d^4x_ie^{-ip_ix_i}\prod_1^{\overline{n}}\int d^4\overline{x}_ie^{-i\overline{p}_i\overline{x}_i}\prod_1^m\int d^4y_je^{ik_jy_j}\prod_1^{\overline{m}}d^4\overline{y}_je^{i\overline{k}_j\overline{y}_j}\\
\times\;&[\overline{u}_{s_1}(\boldsymbol{p}_1)(\not{p}_1+m)]\cdots[\overline{u}_{s_n}(\boldsymbol{p}_n)(\not{p}_n+m)]\\
\times\;&[\overline{v}_{\overline{r}_1}(\overline{\boldsymbol{k}}_1)(\overline{\not{k}}_1-m)]\cdots[\overline{v}_{\overline{r}_{\overline{m}}}(\overline{\boldsymbol{k}}_{\overline{m}})(\overline{\not{k}}_{\overline{m}}-m)]\\
\times\;&\langle\Omega|T\{\Psi(x_1)\cdots\Psi(x_n)\overline{\Psi}(\overline{x}_1)\cdots\overline{\Psi}(\overline{x}_{\overline{n}})\overline{\Psi}(y_1)\cdots\overline{\Psi}(y_m)\Psi(\overline{y}_1)\cdots\Psi(\overline{y}_{\overline{m}})\}|\Omega\rangle\\
\times\;&[(\not{k}_1+m)u_{r_1}(\boldsymbol{k}_1)]\cdots[(\not{k}_m+m)u_{r_m}(\boldsymbol{k}_m)]\\
\times\;&[(\overline{\not{p}}_1-m)v_{\overline{s}_1}(\overline{\boldsymbol{p}}_1)]\cdots[(\overline{\not{p}}_{\overline{n}}-m)v_{\overline{s}_{\overline{n}}}(\overline{\boldsymbol{p}}_{\overline{n}})]\\
\times\;&\left(\frac{i}{\sqrt{Z_2}}\right)^{m+\overline{m}+n+\overline{n}}
\end{aligned}$$

From this equation, we can see that the scattering amplitude would vanish unless $n+\overline{m}=\overline{n}+m$, which implies the conservation of Q. The term e^{ipx} will impose the condition of momentum conservation, and the term $\not{p}\pm m$ will remove the external legs.
Finally, we list the Feymann rules of Yukawa theory in momentum space as follows:

1. For each incoming electron, draw a solid line with an arrow pointed towards the vertex, and label it with the electron's four-momentum, k_i.
2. For each outgoing electron, draw a solid line with an arrow pointed away from the vertex, and label it with the electron's four-momentum, p_i.
3. For each incoming positron, draw a solid line with an arrow pointed away from the vertex, and label it with minus the positron's four-momentum, $-\bar{k}_i$.
4. For each outgoing positron, draw a solid line with an arrow pointed towards the vertex, and label it with minus the positron's four-momentum, $-\bar{p}_i$.

5. For each incoming scalar, draw a dashed line with an arrow pointed towards the vertex, and label it with the scalar's four-momentum, q_i.

6. For each outgoing scalar, draw a dashed line with an arrow pointed away from the vertex, and label it with the scalar's four-momentum, q_i'.

7. The only allowed vertex joins two solid lines, one with an arrow pointing towards it and one with an arrow pointing away from it, and one dashed line (whose arrow can point in either direction). Using this vertex, join up all the external lines, including extra internal lines as needed. In this way, draw all possible diagrams that are topologically inequivalent.

8. Assign each internal line its own four-momentum. Think of the four-momenta as flowing along the arrows, and conserve four-momentum at each vertex. For a tree diagram, this fixes the momenta on all the internal lines.

9. The value of a diagram consists of the following factors:

   - for each incoming or outgoing scalar, 1;
   - for each incoming electron, $u_r(\boldsymbol{k})$;
   - for each outgoing electron, $\overline{u}_s(\boldsymbol{p})$;
   - for each incoming positron, $\overline{v}_{\overline{r}}(\overline{\boldsymbol{k}})$;
   - for each outgoing positron, $v_{\overline{s}}(\overline{\boldsymbol{p}})$;
   - for each vertex, $-ig_0$;
   - for each internal scalar, $\frac{-i}{p^2+M^2-i\epsilon}$;
   - for each internal fermion, $\frac{i(\not{p}-m)}{p^2+m^2-i\epsilon}$

10. Spinor indices are contracted by starting at one end of a fermion line: specifically, the end that has the arrow pointing away from the vertex. The factor associated with the external line is either $\bar{u}$ or $\bar{v}$. Go along the complete fermion line, following the arrows backwards, and write down (in order from left to right) the factors associated with the vertices and propagators that you encounter. The last factor is either a u or v. Repeat this procedure for the other fermion lines, if any.

11. The overall sign of a tree diagram is determined by drawing all contributing diagrams in a standard form: all fermion lines horizontal, with their arrows pointing from left to right, and with the left endpoints labeled in the same fixed order (from top to bottom); if the ordering of the labels on the right endpoints of the fermion lines in a given diagram is an even (odd) permutation of an arbitrarily chosen fixed ordering, then the sign of that diagram is positive (negative).

12. Each closed fermion loop contributes an extra minus sign.

13. Value of $i\mathcal{M}$ is given by a sum over the values of the contributing diagrams.

14. $\langle f|S|i\rangle = (Z_1)^{\frac{n_s}{2}}(Z_2)^{\frac{n_f}{2}} i\mathcal{M}\delta(\sum p_f - \sum p_i)$

Chapter 14
Vector Field

14.1 Vector field

Consider a vector field $A^\mu(x)$. Here the index μ is a vector index that takes on four possible values. Under a Lorentz transformation, we have

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

For an infinitesimal transformation, we can write

$$\delta^\mu{}_\nu + \delta\omega^\mu{}_\nu = \delta^\mu{}_\nu + \frac{i}{2}\delta\omega_{\rho\sigma}(S_V^{\rho\sigma})^\mu{}_\nu$$

Here

$$(S_V^{\rho\sigma})^\mu{}_\nu = -i(\eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho{}_\nu)$$

It is obvious that $A^{\dagger\mu}$ is also a vector field. We know that $\eta^{\mu\nu}$ is invariant under Lorentz transformation, i.e.

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

We can use $\eta^{\mu\nu}$ and and its inverse $\eta_{\mu\nu}$ to raise and lower vector indices of the vector field,

$$A_\mu \equiv \eta_{\mu\nu}A^\nu$$

And we can verify the following equations

$$\Lambda^\mu{}_\nu\Lambda_\mu{}^\rho = \delta^\rho_\nu$$

$$A^\mu(x) = \eta^{\mu\nu}A_\nu(x)$$

$$\Lambda_\mu{}^\rho\Lambda_\nu{}^\sigma\eta_{\rho\sigma} = \eta_{\mu\nu}$$

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

Define $C_i \equiv \frac{1}{2}\epsilon_{ijk}S_V^{jk}$,$D_i \equiv S_V^{i0}$. For example, we have

$$(C_3)_\mu{}^\nu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The eigenvectors of C_3 are

$$[(-1, [(0,\ 1,\ -i,\ 0)], 1), (1, [(0,\ 1,\ i,\ 0)], 1), (0, [(1,\ 0,\ 0,\ 0), (0,\ 0,\ 0,\ 1)], 2)]$$

We further define $N_i \equiv \frac{1}{2}(C_i - iD_i)$ and $N_i^\dagger \equiv \frac{1}{2}(C_i + iD_i)$. For example, we have

$$(N_1)_\mu{}^\nu = \begin{pmatrix} 0 & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2}i \\ 0 & 0 & \frac{1}{2}i & 0 \end{pmatrix}$$

The eigenvectors of N_1 are

$$\left[\left(-\frac{1}{2}, [(1,\ 1,\ 0,\ 0)\,, (0,\ 0,\ 1,\ -i)]\,, 2\right), \left(\frac{1}{2}, [(1,\ -1,\ 0,\ 0)\,, (0,\ 0,\ 1,\ i)]\,, 2\right)\right]$$

And we can conclude that vector is in the $(2,2)$ representation of the Lie algebra of the Lorentz group.

14.2 Electromagnetic field and gauge invariance

The Lagrangian of EM field is

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

Here,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad \text{and} \quad A^\mu = (\phi, \boldsymbol{A})$$

So,

$$F_{0i} = \dot{A}^i + \nabla_i \phi \equiv -E^i \quad \text{and} \quad F_{ij} = \nabla_i A^j - \nabla_j A^i \equiv \epsilon_{ijk} B^k$$

We can derive the equation of motion of the EM field by variation method,

$$\partial_\mu F^{\mu\nu} = 0$$

It can be rewritten in terms of $\boldsymbol{E}$ and $\boldsymbol{B}$, i.e. Maxwell equations:

$$\begin{aligned} \boldsymbol{\nabla} \cdot \boldsymbol{E} = 0 \quad & \frac{\partial \boldsymbol{E}}{\partial t} = \boldsymbol{\nabla} \times \boldsymbol{B} \\ \boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 \quad & \frac{\partial \boldsymbol{B}}{\partial t} = -\boldsymbol{\nabla} \times \boldsymbol{E} \end{aligned}$$

The massless vector field A_μ has 4 components, which would naively seem to tell us that the gauge field has 4 degrees of freedom.But there are two related comments which will ensure that quantizing the gauge field A_μ gives rise to 2 degrees of freedom, rather than 4.

- The field A_0 has no kinetic term $\dot{A}_0$ in the Lagrangian: it is not dynamical. This means that if we are given some initial data A_i and $\dot{A}_i$ at a time t_0, then the field A_0 is fully determined by the equation of motion $\boldsymbol{\nabla} \cdot \boldsymbol{E} = 0$,which, expanding out, reads

  $$\nabla^2 A_0 = \boldsymbol{\nabla} \cdot \frac{\partial \boldsymbol{A}}{\partial t}$$

  So A_0 is not independent: we don't get to specify A_0 on the initial time slice.

- If we transform the EM field as
$$A_\mu \to A_\mu + \partial_\mu \lambda(x)$$
we can derive that
$$F_{\mu\nu} \to F_{\mu\nu} \quad \mathcal{L} \to \mathcal{L}$$
The seemed infinite number of symmetries, one for each function $\lambda(x)$, is to be viewed as a redundancy in our description. That is, two states related by a gauge symmetry are to be identified: they are the same physical state. One way to see that this interpretation is necessary is to notice that Maxwell's equations are not sufficient to specify the evolution of A_μ.The equations read,
$$(\eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu)A^\nu = 0$$
But the operator $(\eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu)$ is not invertible: it annihilates any function of the form $\partial_\mu\lambda$. This means that given any initial data, we have no way to uniquely determine A_μ at a later time since we can't distinguish between A_μ and $A_\mu + \partial_\mu\lambda$. This would be problematic if we thought that A_μ is a physical object. However, if we're happy to identify A_μ and $A_\mu + \partial_\mu\lambda$ as corresponding to the same physical state, then our problems disappear.

The picture that emerges for the theory of electromagnetism is of an enlarged phase space, foliated by gauge orbits. All states that lie along a given gauge orbit can be reached by a gauge transformation and are identified. To make progress, we pick a representative from each gauge orbit. It doesn't matter which representative we pick after all, they're all physically equivalent. But we should make sure that we pick a "good" gauge, in which we cut the orbits. Here we'll look at two different gauges:

- Coulomb Gauge: $\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$
We can make use of the residual gauge transformations in Lorentz gauge to pick $\boldsymbol{\nabla} \cdot \dot{\boldsymbol{A}} = 0$. We have as a consequence $A_0 = 0$. Coulomb gauge is sometimes called radiation gauge.

- Lorentz Gauge: $\partial^\mu A_\mu = 0$
In fact this condition doesn't pick a unique representative from the gauge orbit. We're always free to make further gauge transformations with $\partial^\mu\partial_\mu\lambda = 0$, which also has non-trivial solutions. As the name suggests, the Lorentz gauge has the advantage that it is Lorentz invariant.

14.3 Canonical quantization of EM field

14.3.1 Canonical quantization in Coulomb gauge

Canonical momentum and Hamiltonian

$$\pi^0 = \frac{\partial\mathcal{L}}{\partial\dot{A}_0} = 0 \quad \pi^i = \frac{\partial\mathcal{L}}{\partial(\partial_0 A_i)} = \dot{A}^i + \nabla_i\phi = -E^i$$

$$\mathcal{H} = \frac{1}{2}(\boldsymbol{\pi}^2 + \boldsymbol{B}^2) + (\boldsymbol{\pi} \cdot \boldsymbol{\nabla})A_0$$

Integration by parts can give

$$H = \int d^3x \frac{1}{2}(\boldsymbol{\pi}^2 + \boldsymbol{B}^2)$$

Momentum and angular momentum

$$P^0 = H \quad \vec{P} = \int -\boldsymbol{\pi}\vec{\nabla}\boldsymbol{A} d^3x$$

$$\vec{J} = -\int \boldsymbol{\pi}(\vec{x} \times \vec{\nabla} + i\vec{C})\boldsymbol{A}\ d^3x \quad \vec{S} = -i\int \boldsymbol{\pi}\vec{C}\boldsymbol{A}\ d^3x$$

Canonical quantization

In Coulomb gauge, we have

$$A_0 = \pi^0 = 0 \quad \pi^i = \dot{A}^i$$

Three pairs of A_i and π^i are not independent from each other. They must satisfy the constraint equations

$$\nabla \cdot \boldsymbol{A} = 0 \quad \nabla \cdot \boldsymbol{\pi} = 0$$

A reasonable quantization condition can be written as

$$[A_i(\boldsymbol{x}, t), A_j(\boldsymbol{x}', t)] = 0 \quad [\pi^i(\boldsymbol{x}, t), \pi^j(\boldsymbol{x}', t)] = 0$$

$$[A_i(\boldsymbol{x}, t), \pi^j(\boldsymbol{x}', t)] = i\left(\delta_i^j - \frac{\partial_i\partial^j}{\nabla^2}\right)\delta(\boldsymbol{x} - \boldsymbol{x}') \equiv i\int \frac{d^3k}{(2\pi)^3}(\delta_i^j - \frac{k_ik^j}{\boldsymbol{k}^2})e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')}$$

In this case, we can verify that

$$\dot{A}_i = -i[A_i(\boldsymbol{x}, t), H] = \pi_i(\boldsymbol{x}, t)$$

$$\dot{\pi}^i = -i[\pi^i(\boldsymbol{x}, t), H] = \nabla^2 A^i(\boldsymbol{x}, t)$$

It is constant with the field equation we derive from Euler-Lagrange equation.

Fourier expansion

$$\boldsymbol{A}(x) = \sum_{r=\pm}\int \widetilde{dp}[a_r(\boldsymbol{p})\boldsymbol{\epsilon}_r(\boldsymbol{p})e^{ipx} + a_r^\dagger(\boldsymbol{p})\boldsymbol{\epsilon}_r^*(\boldsymbol{p})e^{-ipx}]$$

And we can derive from constraint condition that

$$\boldsymbol{\epsilon} \cdot \boldsymbol{p} = 0$$

We will choose $\boldsymbol{\epsilon}$ to satisfy that

$$\boldsymbol{\epsilon}_r \cdot \boldsymbol{\epsilon}_s^* = \delta_{rs}$$

So, the completeness relation for the polarization vectors is

$$\sum_{r=\pm} \epsilon_r^i(\boldsymbol{p})\epsilon_r^{*j}(\boldsymbol{p}) = \delta^{ij} - \frac{p^i p^j}{|\boldsymbol{p}|^2}$$

Example: If $\boldsymbol{p} = (0, 0, p)$, we usually choose

$$\boldsymbol{\epsilon}_+ = \frac{1}{\sqrt{2}}(1, i, 0) \quad \boldsymbol{\epsilon}_- = \frac{1}{\sqrt{2}}(1, -i, 0)$$

$\boldsymbol{\epsilon}_+$ corresponds to left-handed rotation and it is the eigenvectors of the space-part of C_3 with eigenvalue $+1$. $\boldsymbol{\epsilon}_-$ corresponds to right-handed rotation and it is eigenvector of the space-part of C_3 with eigenvalue -1.

We can further derive from above discussion that

$$\boldsymbol{\pi}(x) = -i \sum_{r=\pm} \int \widetilde{dp}\omega[a_r(\boldsymbol{p})\boldsymbol{\epsilon}_r(\boldsymbol{p})e^{ipx} - a_r^\dagger(\boldsymbol{p})\boldsymbol{\epsilon}_r^*(\boldsymbol{p})e^{-ipx}]$$

$$a_r(\boldsymbol{p}) = \boldsymbol{\epsilon}_r^* \int d^3x e^{-ikx}(i\boldsymbol{\pi} + \omega \boldsymbol{A})$$

$$a_r^\dagger(\boldsymbol{p}) = \boldsymbol{\epsilon}_r \int d^3x e^{ikx}(-i\boldsymbol{\pi} + \omega \boldsymbol{A})$$

$$[a_r(\boldsymbol{p}), a_{r'}(\boldsymbol{p'})] = 0 \quad [a_r^\dagger(\boldsymbol{p}), a_{r'}^\dagger(\boldsymbol{p'})] = 0 \quad [a_r(\boldsymbol{p}), a_{r'}^\dagger(\boldsymbol{p'})] = (2\pi)^3 2\omega\delta_{rr'}\delta(\boldsymbol{p} - \boldsymbol{p}')$$

Operator represented by a and $a^\dagger$

Define that

$$N(\boldsymbol{p}, r) \equiv a_r^\dagger(\boldsymbol{p})a_r(\boldsymbol{p})$$

So, we can derive

$$H = \sum_{r=\pm} \int \widetilde{dp}\, \omega N(\boldsymbol{p}, r) + 2\mathcal{E}_0 V$$

$$\vec{P} = \sum_{r=\pm} \int \widetilde{dp}\, \vec{p} N(\boldsymbol{p}, r)$$

$$\vec{S} = \sum_{r,s=\pm} \int \widetilde{dp}\, \frac{1}{2}(\boldsymbol{\epsilon}_s^* \vec{C} \boldsymbol{\epsilon}_r - \boldsymbol{\epsilon}_r \vec{C} \boldsymbol{\epsilon}_s^*) a_s^\dagger(\boldsymbol{p}) a_r(\boldsymbol{p})$$

From above equation, we can say that $a_r^\dagger(\boldsymbol{p})$ create an photon with energy ω, momentum $\boldsymbol{p}$ and spin angular momentum along the direction of momentum r.

Propagator

$$G_F(x-y)_{ij} \equiv \langle 0|TA_i(x)A_j(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4} \frac{-i}{p^2 - i\epsilon} \left(\delta_{ij} - \frac{p_i p_j}{|\boldsymbol{p}|^2}\right) e^{ip(x-y)}$$

14.3.2 Canonical quantization in Lorentz gauge

Undefined metric formalism

Modify the Maxwell Lagrangian introducing a new term

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial_\mu A^\mu)^2$$

The equations of motion are now

$$\partial^2 A_\mu - (1 - \frac{1}{\xi})\partial^\mu(\partial \cdot A) = 0$$

Canonical momentums are

$$\pi^0 = \frac{1}{\xi}\partial \cdot A = \frac{1}{\xi}(-\dot{A}_0 + \partial_i A^i) \quad \pi^i = \dot{A}^i + \nabla^i A^0 = -E^i$$

Hamiltonian is

$$\mathcal{H} = \frac{1}{2}(\boldsymbol{\pi}^2 + \boldsymbol{B}^2 - \xi\pi^0\pi^0) + (\boldsymbol{\pi} \cdot \boldsymbol{\nabla})A_0 + \pi^0(\boldsymbol{\nabla} \cdot \boldsymbol{A})$$

$$H = \int \left[\frac{1}{2}(\boldsymbol{\pi}^2 + \boldsymbol{B}^2 - \xi\pi^0\pi^0) - A_0(\boldsymbol{\nabla} \cdot \boldsymbol{\pi}) + \pi^0(\boldsymbol{\nabla} \cdot \boldsymbol{A})\right] d^3x$$

We remark that the above Lagrangian and the equations of motion, reduce to Maxwell theory in the gauge $\partial \cdot A = 0$. This why we say that our choice corresponds to a class of Lorenz gauges with parameter ξ. With this abuse of language (in fact we are not setting $\partial \cdot A = 0$, otherwise the problems would come back) the value of $\xi = 1$ is known as the Feynman gauge and $\xi = 0$ as the Landau gauge. From now on we will take the case of the so-called Feynman gauge, where $\xi = 1$. Then the equation of motion coincide with the Maxwell theory in the Lorenz gauge. In Feymann gauge,the canonical quantization conditions can be written as

$$[A_\mu(\boldsymbol{x},t), A_\nu(\boldsymbol{x}',t)] = 0 \quad [\pi^\mu(\boldsymbol{x},t), \pi^\nu(\boldsymbol{x}',t)] = 0 \quad [A_\mu(\boldsymbol{x},t), \pi^\nu(\boldsymbol{x}',t)] = i\delta^\nu_\mu\delta(\boldsymbol{x} - \boldsymbol{x}')$$

we can also derive that

$$[\dot{A}_\mu(\boldsymbol{x},t), \dot{A}_\nu(\boldsymbol{x}',t)] = 0 \quad [A_\mu(\boldsymbol{x},t), \dot{A}_\nu(\boldsymbol{x}',t)] = i\eta_{\mu\nu}\delta(\boldsymbol{x} - \boldsymbol{x}')$$

Fourier expansion

$$A(x) = \sum_{\lambda=0}^{3}\int \widetilde{dp}[a_\lambda(\boldsymbol{p})\epsilon_\lambda(\boldsymbol{p})e^{ipx} + a^\dagger_\lambda(\boldsymbol{p})\epsilon^*_\lambda(\boldsymbol{p})e^{-ipx}]$$

where $\epsilon_{\lambda\mu}$ are a set of four independent 4-vectors. We will now make a choice for these 4-vectors. We choose $\epsilon_{1\mu}$ and $\epsilon_{2\mu}$ orthogonal to k^μ and n^μ, such that

$$\epsilon_{\lambda\mu}\epsilon^{*\mu}_{\lambda} = \delta_{\lambda\lambda'} \quad \lambda, \lambda' = 1, 2$$

After, we choose $\epsilon_{3\mu}$ in the plane (k^μ, n^μ) and perpendicular to n^μ such that

$$\epsilon_{3\mu}n^\mu = 0 \quad \epsilon_{3\mu}\epsilon_3^{*\mu} = 1$$

Finally we choose $\epsilon_{0\mu} = n_\mu$. The vectors $\epsilon_{1\mu}$ and $\epsilon_{2\mu}$ are called transverse polarizations, while $\epsilon_{3\mu}$ and $\epsilon_{0\mu}$ longitudinal and scalar polarizations, respectively.
In general we can show that

$$\epsilon_\lambda \cdot \epsilon^*_{\lambda'} = \eta_{\lambda\lambda'} \quad \eta^{\lambda\lambda'}\epsilon_{\lambda\mu}\epsilon^*_{\lambda'\nu} = \eta_{\mu\nu}$$

We can further derive from above discussion that

$$\dot{A}(x) = -i\sum_{\lambda=0}^{3}\int \widetilde{dp}\omega[a_\lambda(\boldsymbol{p})\epsilon_\lambda(\boldsymbol{p})e^{ipx} - a^\dagger_\lambda(\boldsymbol{p})\epsilon^*_\lambda(\boldsymbol{p})e^{-ipx}]$$

$$a_\lambda(\boldsymbol{p}) = \eta_{\lambda\lambda'}\epsilon^*_{\lambda'} \cdot \int d^3x e^{-ipx}(i\dot{A} + \omega A)$$

$$a^\dagger_\lambda(\boldsymbol{p}) = \eta_{\lambda\lambda'}\epsilon_{\lambda'} \cdot \int d^3x e^{ipx}(-i\dot{A} + \omega A)$$

$$[a_\lambda(\boldsymbol{p}), a_{\lambda'}(\boldsymbol{p'})] = 0 \quad [a^\dagger_\lambda(\boldsymbol{p}), a^\dagger_{\lambda'}(\boldsymbol{p'})] = 0 \quad [a_\lambda(\boldsymbol{p}), a^\dagger_{\lambda'}(\boldsymbol{p'})] = (2\pi)^3 2\omega\eta_{\lambda\lambda'}\delta(\boldsymbol{p} - \boldsymbol{p'})$$

Indefinite metric problem

We Introduce the vacuum state defined by

$$a_\lambda(\boldsymbol{p})|0\rangle = 0$$

To see the problem with the sign we construct the one-particle state with scalar polarization, that is

$$|1\rangle = \int \widetilde{dp} a^\dagger_0(\boldsymbol{p})|0\rangle$$

and calculate its norm

$$\langle 1|1\rangle = -\langle 0|0\rangle \int \widetilde{dp}|f(p)|^2$$

The state $|1\rangle$ has a negative norm.

To solve this problem we note that we are not working anymore with the classical Maxwell theory because we modified the Lagrangian. What we would like to do is to impose the condition $\partial \cdot A = 0$, but that is impossible as an equation for operators. We can, however, require that condition on a weaker form, as a condition only to be verified by the physical states.

More specifically, we require that the part of $\partial \cdot A$ that contains the annihilation operator (positive frequencies) annihilates the physical states,

$$\partial^\mu A^+_\mu|\psi\rangle = 0$$

The states $|\psi\rangle$ can be written in the form

$$|\psi\rangle = |\psi_T\rangle|\phi\rangle$$

where $|\psi_T\rangle$ is obtained from the vacuum with creation operators with transverse polarization and $|\phi\rangle$ with scalar and longitudinal polarization.

$\partial^\mu A_\mu^+$ contains only scalar and longitudinal polarizations

$$\partial^\mu A_\mu^+ = i \sum_{\lambda=0,3} \int \widetilde{dp} a_\lambda(\boldsymbol{p})(p \cdot \epsilon_\lambda(\boldsymbol{p}))e^{ipx}$$

Therefore the previous condition becomes

$$i \sum_{\lambda=0,3} (p \cdot \epsilon_\lambda(\boldsymbol{p}))a_\lambda(\boldsymbol{p})|\phi\rangle = 0$$

The condition is equivalent to,

$$(a_0(\boldsymbol{p}) - a_3(\boldsymbol{p}))|\phi\rangle = 0$$

We can construct $|\phi\rangle$ as a linear combination of states $|\phi\rangle$ with n scalar or longitudinal photons:

$$|\phi\rangle = C_0|\phi_0\rangle + C_1|\phi\rangle + \cdots \quad \text{Here,} |\phi_0\rangle \equiv |0\rangle$$

The states $|\phi_n\rangle$ are eigenstates of the operator number for scalar or longitudinal photons

$$N'|\phi_n\rangle = n|\phi_n\rangle$$

where,

$$N' = \int \widetilde{dp}[a_3^\dagger(\boldsymbol{p})a_3(\boldsymbol{p}) - a_0^\dagger(\boldsymbol{p})a_0(\boldsymbol{p})]$$

Then

$$n\langle\phi_n|\phi_n\rangle = \langle\phi_n|N'|\phi_n\rangle = 0$$

This means that

$$\langle\phi_n|\phi_n\rangle = \delta_{n0}$$

that is, for $n \neq 0$, the state $|\phi_n\rangle$ has zero norm. We have then for the general state $|\phi\rangle$,

$$\langle\phi|\phi\rangle = |C_0|^2 \geq 0$$

and the coefficients $C_i(i = 1, 2, \cdots)$ are arbitrary.

Operator represented by a and $a^\dagger$

Define that

$$N'(\boldsymbol{p}) \equiv a_3^\dagger(\boldsymbol{p})a_3(\boldsymbol{p}) - a_0^\dagger(\boldsymbol{p})a_0(\boldsymbol{p})$$

$$N(\boldsymbol{p}, 1) \equiv a_1^\dagger(\boldsymbol{p})a_1(\boldsymbol{p}) \quad N(\boldsymbol{p}, 2) \equiv a_2^\dagger(\boldsymbol{p})a_2(\boldsymbol{p}) \quad N_T(\boldsymbol{p}) \equiv N(\boldsymbol{p}, 1) + N(\boldsymbol{p}, 2)$$

We have that

$$\langle\psi|N'(\boldsymbol{p})|\psi\rangle = 0 \quad \langle\psi|N_T(\boldsymbol{p})|\psi\rangle = \langle\psi_T|N_T(\boldsymbol{p})|\psi_T\rangle$$

We can derive

$$H = \int \widetilde{dp}\, \omega[N'(\boldsymbol{p}) + N_T(\boldsymbol{p})] + 2\mathcal{E}_0 V$$

$$\vec{P} = \int \widetilde{dp}\, \vec{p}[N'(\boldsymbol{p}) + N_T(\boldsymbol{p})]$$

So, the arbitrariness of $C_i(i = 1, 2, \cdots)$ does not affect the physical observables. Only the physical transverse polarizations contribute to the result. Two states that differ only in their timelike and longitudinal photon content, $|\phi_n\rangle$ with $n \geq 1$ are said to be physically equivalent. We can think of the gauge symmetry of the classical theory as descending to the Hilbert space of the quantum theory.

It is important to note that although for the average values of the physical observables only the transverse polarizations contribute, the scalar and longitudinal polarizations are necessary for the consistency of the theory. In particular they show up when we consider complete sums over the intermediate states.

Propagator

$$G_F(x-y)_{\mu\nu} \equiv \langle 0|TA_\mu(x)A_\nu(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{-i\eta_{\mu\nu}}{p^2 - i\epsilon}e^{ip(x-y)}$$

It is easy to verify that $G_F(x-y)_{\mu\nu}$ is the Green's function of the equation of motion, that for $\xi = 1$ is the wave equation, that is

$$\partial^2 G_F(x-y)_{\mu\nu} = i\eta_{\mu\nu}\delta(x-y)$$

For the general case, $\xi \neq 0$, the equal times commutation relations are more complicated. And the propagator will be

$$G_F(x-y)_{\mu\nu} = \int \frac{d^4p}{(2\pi)^4}\left[\frac{-i\eta_{\mu\nu}}{p^2 - i\epsilon} + i(1-\xi)\frac{p_\mu p_\nu}{(p^2 - i\epsilon)^2}\right]e^{ip(x-y)}$$

14.4 Perturbation theory for canonical quantization

14.4.1 Lagrangian of QED

The Lagrangian of QED is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\Psi}(i\partial\!\!\!/ - m_0)\Psi + e_0 j^\mu A_\mu,$$

where $j^\mu \equiv \overline{\Psi}\gamma^\mu\Psi$. Usually, we also define a covariant derivative,

$$D_\mu\Psi \equiv \partial_\mu\Psi - ie_0A_\mu\Psi$$

So, the Lagrangian can also be written as

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\Psi}(i D\!\!\!/ - m_0)\Psi$$

The Lagrangian is invariant under the gauge transformation

$$A_\mu(x) \to A_\mu(x) + \frac{1}{e_0}\partial_\mu\alpha(x) \quad \Psi(x) \to e^{i\alpha(x)}\Psi(x)$$

14.4.2 Coulomb gauge

The constraint equations are

$$\boldsymbol{\nabla}\cdot\boldsymbol{A}=0\quad \boldsymbol{\nabla}^2A_0=e_0j^0$$

The solution for A_0 is

$$A_0(\boldsymbol{x},t)=-e_0\int d^3x'\frac{j^0(\boldsymbol{x}',t)}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}$$

Finally, we can derive that

$$H=H_D+H_M+H_{\text{int}}$$

where,

$$H_D=\int d^3x\ -\Pi(\vec{\alpha}\cdot\vec{\nabla}+i\beta m)\Psi\quad H_M=\int d^3x\ \frac{1}{2}(\boldsymbol{\pi}^2+\boldsymbol{B}^2)$$

$$H_{\text{int}}=\int d^3x\ \left[-e_0\boldsymbol{j}\cdot\boldsymbol{A}+\frac{e_0^2}{2}\int d^3x'\frac{j^0(\boldsymbol{x})j^0(\boldsymbol{x}')}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}\right]$$

The perturbation expansion of correlation functions is

$$\langle\Omega|T\{\Psi(x)\overline{\Psi}(y)A(z)\}|\Omega\rangle=\lim_{T\to\infty(1-i\epsilon)}\frac{\langle 0|T\left\{\Psi_I(x)\overline{\Psi}_I(y)A_I(z)\text{exp}\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle}{\langle 0|T\left\{\text{exp}\left[-i\int_{-T}^{T}dtH_I\right]\right\}|0\rangle}$$

Here, $\int_T^T dtH_I$ can be written as

$$\left[-\int d^4xe_0\overline{\Psi}_I\boldsymbol{\gamma}\Psi_I\cdot\boldsymbol{A}_I\right]+\left[\int d^4x\int d^4x'\frac{e_0^2\delta(t-t')}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}\frac{1}{2}\overline{\Psi}_I(\boldsymbol{x},t)\gamma^0\Psi_I(\boldsymbol{x},t)\overline{\Psi}_I(\boldsymbol{x}',t')\gamma^0\Psi_I(\boldsymbol{x}',t')\right]$$

Wick’s theorem for photons takes is similar to that in ϕ^4 theory:

$$T\left\{A_I(x_1)A_I(x_2)A_I(x_3)\cdots\right\}=N\left\{A_I(x_1)A_I(x_2)A_I(x_3)\cdots+\text{ all possible contractions }\right\}$$

Example:

$$\begin{aligned}&\langle 0|T\left\{A_{Ii}(x_1)A_{Ij}(x_2)A_{Ik}(x_3)A_{Il}(x_4)\right\}|0\rangle\\=\ &G_F(x_1-x_2)_{ij}G_F(x_3-x_4)_{kl}+G_F(x_1-x_3)_{ik}G_F(x_2-x_4)_{jl}+G_F(x_1-x_4)_{il}G_F(x_2-x_3)_{jk}\end{aligned}$$

Now we can derive the Feynman rule for QED theory. Firstly, we evaluate this term,

$$\langle 0|T\left\{\Psi_{Ia}(x)\overline{\Psi}_{Ib}(y)A_{Ii}(z)(ie_0)\int dw^4\overline{\Psi}_I(w)\boldsymbol{\gamma}\Psi_I(w)\cdot A_I(w)\right\}|0\rangle$$

After contraction, it can be written as

$$\begin{aligned}&-\ (ie_0)S_F(x-y)_{ab}\int d^4wG_F(z-w)_{ik}\text{Tr}[\gamma^kS_F(w-w)]\\&+\ (ie_0)\int d^4wG_F(w-z)_{ik}[S_F(x-w)\gamma^kS_F(w-y)]_{ab}\end{aligned}$$

It can be represented by the following Feynman diagram.

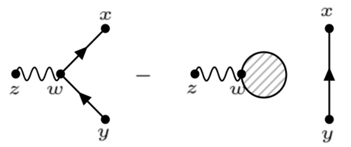


Figure 14.1: Feynman diagram representation of perturbation expansion

Secondly, we evaluate this term,

$$\langle 0|T\left\{\Psi_{Ia}(x)\overline{\Psi}_{Ib}(y)\int d^4w\int d^4w'\frac{-ie_0^2\delta(w^0-w'^0)}{4\pi|\boldsymbol{w}-\boldsymbol{w}'|}\frac{1}{2}\overline{\Psi}_I(w)\gamma^0\Psi_I(w)\overline{\Psi}_I(w')\gamma^0\Psi_I(w')\right\}|0\rangle$$

After contraction, it can be written as

$$\begin{aligned}
&-\quad (ie_0)^2\int d^4wd^4w'\frac{i\delta(w^0-w'^0)}{4\pi|\boldsymbol{w}-\boldsymbol{w}'|}[S_F(x-w)\gamma^0S_F(w-y)]_{ab}\mathrm{Tr}[\gamma^0S_F(w'-w')]\\
&+\quad (ie_0)^2\int d^4wd^4w'\frac{i\delta(w^0-w'^0)}{4\pi|\boldsymbol{w}-\boldsymbol{w}'|}[S_F(x-w)\gamma^0S_F(w-w')\gamma^0S_F(w'-y)]_{ab}\\
&+\quad \frac{1}{2}(ie_0)^2S_F(x-y)_{ab}\int d^4wd^4w'\frac{i\delta(w^0-w'^0)}{4\pi|\boldsymbol{w}-\boldsymbol{w}'|}\mathrm{Tr}[\gamma^0S_F(w-w)]\mathrm{Tr}[\gamma^0S_F(w'-w')]\\
&-\quad \frac{1}{2}(ie_0)^2S_F(x-y)_{ab}\int d^4wd^4w'\frac{i\delta(w^0-w'^0)}{4\pi|\boldsymbol{w}-\boldsymbol{w}'|}\mathrm{Tr}[\gamma^0S_F(w-w')\gamma^0S_F(w'-w)]
\end{aligned}$$

It can be represented by the following Feynman diagram.

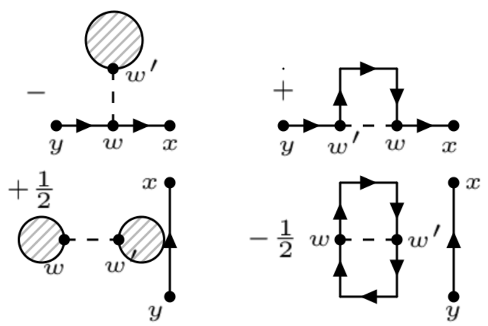


Figure 14.2: Feynman diagram representation of perturbation expansion

Before we write down Feymann rules, we notice that the offending non-local interaction comes from the A_0 component of the gauge field, we could try to redefine the propagator to include a $G_F(x-y)_{00}$ piece which will capture this term. We can verify that

$$\frac{i\delta(w^0-w'^0)}{4\pi|\boldsymbol{w}-\boldsymbol{w}'|}=\int\frac{d^4p}{(2\pi)^4}\frac{ie^{ip(w-w')}}{|\boldsymbol{p}|^2}$$

So we can combine the non-local interaction with the transverse photon propagator by defining a new photon propagator

$$G_F(p)_{\mu\nu}\equiv\begin{cases}\frac{i}{|\boldsymbol{p}|^2} & \mu,\nu=0\\ \frac{-i}{p^2-i\epsilon}\left(\delta_{ij}-\frac{p_ip_j}{|\boldsymbol{p}|^2}\right) & \mu=i\neq 0,\nu=j\neq 0\\ 0 \quad \text{otherwise}\end{cases}$$

With this propagator, the wavy photon line now carries a $\mu\nu = 0, 1, 2, 3$ index, with the extra $\mu = 0$ component taking care of the instantaneous interaction.
The Feynman rules for QED are:

1. For each Fermion propagator from y to x, $P = S_F(x-y)$
2. For each vector propagator, $P = G_F(x-y)$
3. For each vertex, $V = (ie_0\gamma^\mu)\int d^4w$
4. For each external point, $E = 1$
5. Divided by the symmetry factor

14.4.3 Lorentz Gauge

In Lorentz gauge,

$$\mathcal{H}_{int} = -e_0\overline{\Psi}\gamma^\mu\Psi A_\mu$$

The Feynman rules for QED in Lorentz gauge will be the same as that in Coulomb gauge expect for that the vector propagator will be

$$G_F(p)_{\mu\nu} = \frac{-i\eta_{\mu\nu}}{p^2 - i\epsilon} + i(1-\xi)\frac{p_\mu p_\nu}{(p^2 - i\epsilon)^2}$$

Especially, for Feynman gauge $\xi = 1$, we have

$$G_F(p)_{\mu\nu} = \frac{-i\eta_{\mu\nu}}{p^2 - i\epsilon}$$

14.5 Path integral quantization

14.5.1 Path integral formulation for free EM field

The correlation function is given by

$$\langle\Omega|TA_H(x_1)A_H(x_2)|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}A \exp\left[i\int_{-T}^{T} d^4x(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu})\right] A(x_1)A(x_2)}{\int \mathcal{D}A \exp\left[i\int_{-T}^{T} d^4x(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu})\right]}$$

The generating function is

$$Z[J] = \int \mathcal{D}A \exp\left[i\int d^4x(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}) + J^\mu A_\mu\right]$$

We can verify that

$$S = \int d^4x(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}) = \frac{1}{2}\int d^4x A_\mu(x)(\partial^2\eta^{\mu\nu} - \partial^\mu\partial^\nu)A_\nu(x)$$

Notice that $(\partial^2\eta^{\mu\nu} - \partial^\mu\partial^\nu)$ is singular, since for any $\alpha(x)$,

$$(\partial^2\eta^{\mu\nu} - \partial^\mu\partial^\nu)\partial_\mu\alpha(x) = 0$$

This difficulty is due to gauge invariance: $\alpha(x)$ is gauge equivalent to 0. The functional is badly defined because we are redundantly integrating over a continuous infinity of physically equivalent field configurations. To fix the problem, we would like to isolate the interesting part of the functional integral, which counts each physical configuration only once.
Let $G(A)$ be some function that we wish to set equal zero as a gauge-fixing condition. We could constrain the functional integral to cover only the configurations with $G(A) = 0$ by inserting a functional delta function, $\delta(G(A))$. To do so, we insert 1 in the path integral:

$$1 = \int \mathcal{D}\alpha(x)\delta(G(A(\alpha)))\det\left(\frac{\delta G}{\delta\alpha}\right)$$

where,

$$A_\mu(\alpha(x)) = A_\mu(x) + \frac{1}{e_0}\partial_\mu\alpha(x)$$

We set the gauge fixing function as $G(A) = \partial^\mu A_\mu - \omega(x)$, so $G(A(\alpha)) = \partial^\mu A_\mu + \frac{1}{e_0}\partial^2\alpha - \omega(x)$. It is obvious that $\det\left(\frac{\delta G}{\delta\alpha}\right)$ is equivalent to $\det(\partial^2)/e_0$, which is independent of A. So,

$$Z_0[A] = \det\left(\frac{\delta G}{\delta\alpha}\right)\int\mathcal{D}\alpha\int\mathcal{D}Ae^{iS[A]}\delta(G(A(\alpha)))$$

Now change variables from A to $A(\alpha)$. This is a simple shift, so $\mathcal{D}A = \mathcal{D}A(\alpha)$. Also, by gauge invariance, $S[A] = S[A(\alpha)]$. Since $A(\alpha)$ is now just a dummy integration variable, we can rename it bake to A, so

$$Z_0[A] = \det\left(\frac{\delta G}{\delta\alpha}\right)\int\mathcal{D}\alpha\int\mathcal{D}Ae^{iS[A]}\delta(\partial^\mu A_\mu - \omega(x))$$

Since the above equation is hold for any $\omega(x)$,so we have

$$\begin{aligned}
Z_0[A] &= N(\xi)\int\mathcal{D}\omega\exp\left[-i\int d^4x\frac{\omega^2}{2\xi}\right]\det(\partial^2)\int\mathcal{D}\alpha\int\mathcal{D}Ae^{iS[A]}\delta(\partial^\mu A_\mu - \omega(x)) \\
&= N(\xi)\frac{\det(\partial^2)}{e_0}\int\mathcal{D}\alpha\int\mathcal{D}Ae^{iS[A]}\exp\left[-i\int d^4x\frac{1}{2\xi}(\partial^\mu A_\mu)^2\right] \\
&= W(\xi)\int\mathcal{D}A\exp\left[\frac{i}{2}\int d^4xA_\mu(\partial^2\eta^{\mu\nu} - \partial^\mu\partial^\nu + \frac{1}{\xi}\partial^\mu\partial^\nu)A_\nu\right]
\end{aligned}$$

And we rewrite the generating function as

$$Z[J] = W(\xi)\int\mathcal{D}A\exp\left[\frac{i}{2}\int d^4xA_\mu(\partial^2\eta^{\mu\nu} - \partial^\mu\partial^\nu + \frac{1}{\xi}\partial^\mu\partial^\nu)A_\nu + J^\mu A_\mu\right]$$

Define

$$A'(x) = A(x) - i\int d^4yG_F(x-y)J(y)$$

Recall that

$$(\partial^2\eta^{\mu\nu} - \partial^\mu\partial^\nu + \frac{1}{\xi}\partial^\mu\partial^\nu)G_F(x-y)_{\nu\rho} = i\delta^\mu_\rho\delta(x-y)$$

We can derive that

$$Z[J]=Z_0[A]\exp\left[-\frac{1}{2}\int d^4xd^4yJ^\mu(x)G_F(x-y)_{\mu\nu}J^\nu(y)\right]$$

The two point correlation functions are

$$\langle 0|TA_H(x_1)A_H(x_2)|0\rangle = Z_0^{-1}\left(-i\frac{\delta}{\delta J(x_1)}\right)\left(-i\frac{\delta}{\delta J(x_2)}\right)Z[J]\Big|_{J=0}=G_F(x_1-x_2)$$

14.5.2 Perturbation theory for path integral quantization

We use QED theory as an example.

$$\mathcal{L}=-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}+\overline{\Psi}(i\partial\!\!\!/-m_0)\Psi+e_0\overline{\Psi}\gamma^\mu\Psi A_\mu$$

As we stated in section 4.1, the Lagrangian of QED is also invariant under a general gauge transformation. And we also notice that the measure $\mathcal{D}\Psi\mathcal{D}\overline{\Psi}$ is invariant under gauge transformation. By the similar method, we can show that

$$Z_0\equiv\int\mathcal{D}A\mathcal{D}\Psi\mathcal{D}\overline{\Psi}e^{iS[A,\Psi,\overline{\Psi}]}=W(\xi)\int\mathcal{D}A\mathcal{D}\Psi\mathcal{D}\overline{\Psi}e^{iS[A,\Psi,\overline{\Psi}]}\exp\left[-i\int d^4x\frac{1}{2\xi}(\partial^\mu A_\mu)^2\right]$$

So,define

$$\mathcal{L}_0\equiv-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}-\frac{1}{2\xi}(\partial^\mu A_\mu)^2+\overline{\Psi}(i\partial\!\!\!/-m_0)\Psi\quad\mathcal{L}_1\equiv e_0\overline{\Psi}\gamma^\mu\Psi A_\mu$$

$$\begin{aligned}
Z[J] &= W(\xi)\int\mathcal{D}A\mathcal{D}\overline{\Psi}\mathcal{D}\Psi e^{i\int d^4x[\mathcal{L}_0+\mathcal{L}_1+JA+\overline{\eta}\Psi+\overline{\Psi}\eta]}\\
&= W(\xi)e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)},\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)},-\frac{1}{i}\frac{\delta}{\delta\eta(x)})}\int\mathcal{D}\phi\mathcal{D}\overline{\Psi}\mathcal{D}Ae^{i\int d^4y[\mathcal{L}_0+JA+\overline{\eta}\Psi+\overline{\Psi}\eta]}\\
&\propto e^{i\int d^4x\mathcal{L}_1(\frac{1}{i}\frac{\delta}{\delta J(x)},\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)},-\frac{1}{i}\frac{\delta}{\delta\eta(x)})}\exp[-\int d^4yd^4z\frac{1}{2}J(y)G_F(y-z)J(z)+\overline{\eta}(y)S_F(y-z)\eta(z)]\\
&= \sum_{V=0}^{\infty}\frac{1}{V!}[ie_0\int d^4x(\frac{1}{i}\frac{\delta}{\delta J^\mu(x)}\cdot-\frac{1}{i}\frac{\delta}{\delta\eta(x)}\cdot\gamma^\mu\cdot\frac{1}{i}\frac{\delta}{\delta\overline{\eta}(x)})]^V\\
&\times\sum_{P_1=0}^{\infty}\frac{1}{P_1!}[-\frac{1}{2}\int d^4y_1d^4z_1J(y_1)G_F(y_1-z_1)J(z_1)]^{P_1}\\
&\times\sum_{P_2=0}^{\infty}\frac{1}{P_2!}[-\int d^4y_2d^4z_2\overline{\eta}(y_2)S_F(y_2-z_2)\eta(z_2)]^{P_2}
\end{aligned}$$

If we focus on a term with particular values of V, P_1 and P_2, then the number of surviving scalar sources is $E_1=2P_1-V$, the number of surviving fermion sources is $E_2=2P_2-2V$. We can introduce Feynman diagrams as in the ϕ^4 theory. In these diagrams, a wavy line segment stands for a vector propagator $G_F(x-y)$, a line with an arrow pointing from y to x for a fermion propagator $S_F(x-y)$, a filled circle at one end of a wavy line segment for a vector source $i\int d^4xJ(x)$, a filled circle at the start of a line with an arrow for a fermion source $i\int d^4x\eta(x)$, a filled circle at the end of a line with an arrow for a anti-fermion source $i\int d^4x\overline{\eta}(x)$, a vertex joining three line segments for $ie_0\gamma^\mu\int d^4x$.

14.5.3 Ward-Takahahsi identity (1)

The Noether current of the symmetry $\Psi \to e^{i\alpha}\Psi$ is $j^\mu = \overline{\Psi}\gamma^\mu\Psi$. Recall the conservation law in functional formalism

$$\langle \partial_\mu j^\mu(x)\phi(x_1)\cdots\phi(x_n)\rangle = \sum_{i=1}^{n}\langle \phi(x_1)\cdots(i\Delta\phi(x_i)\delta(x-x_i))\cdots\phi(x_n)\rangle$$

So, we can write the charge conservation law as

$$ie_0\partial_\mu\langle\Omega|Tj^\mu\Psi(x_1)\overline{\Psi}(x_2)|\Omega\rangle = -ie_0\delta(x-x_1)\langle\Omega|T\Psi(x_1)\overline{\Psi}(x_2)|\Omega\rangle + ie_0\delta(x-x_2)\langle\Omega|T\Psi(x_1)\overline{\Psi}(x_2)|\Omega\rangle$$

Notice that $i\langle\Omega|Tj^\mu\psi(x_1)\overline{\Psi}(x_2)|\Omega\rangle$ can be represented by

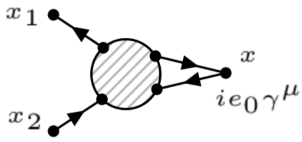


Figure 14.3: Feynman diagram representation of correlation function

From the diagram, we have

$$\langle A_\nu(y)\rangle = \int d^4x G_F(x-y)_{\mu\nu} ie\langle j^\mu(x)\rangle = \int \frac{d^4p}{(2\pi)^4} e^{-ipy} G_F(p)_{\mu\nu}\int d^4x e^{ipx} ie\langle j^\mu(x)\rangle$$

So,

$$\int d^4x\langle A_\nu(x)\rangle e^{ikx} = G_F(p)_{\mu\nu}\int d^4x e^{ikx} ie\langle j^\mu(x)\rangle$$

Compute the Fourier transformation of the equation of charge conservation by integrating

$$\int d^4x e^{ikx}\int d^4x_1 e^{-iqx_1}\int d^4x_2 e^{ipx_2}$$

We can get

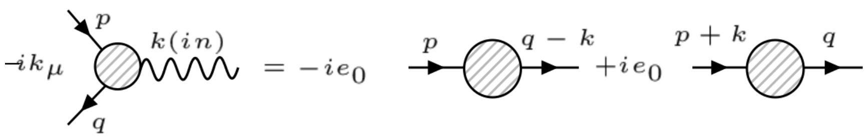


Figure 14.4: Feynman diagram representation of Ward identity

Note that in the diagram above, the external leg of photon will be cut-off, but external leg of fermion will remain. The above equation can be generated to the diagram with n external fermions. Another proof of Ward-Takahashi identity by calculating the Feynman diagram directly can be found in chapter 7.4 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*

14.6 Exact propagator of photon

14.6.1 Photon self-energy

The exact propagator of photon is

$$\mathcal{G}(x)_{\mu\nu} = \langle \Omega | T\{A_\mu(x) A_\nu(0)|\Omega\rangle_C$$

Its Fourier transformation can be represented by the following diagram.

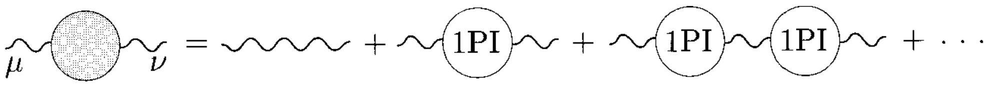


Figure 14.5: Feynman diagram representation of exact propagator of photon

Let us define $i\Pi^{\mu\nu}$ to be the sum of all 1-particle-irreducible insertions into the photon propagator. So, we have

$$\mathcal{G}(k) = G_F(k) + G_F(k)(i\Pi(k))G_F(k) + \cdots = G_F(k)\frac{1}{1 - i\Pi(k)G_F(k)}$$

Hence,

$$(i\mathcal{G})^{-1} = (iG_F)^{-1} - \Pi$$

Recall that

$$iG_F(p)_{\mu\nu} = \frac{\eta_{\mu\nu}}{k^2 - i\epsilon} - (1-\xi)\frac{k_\mu k_\nu}{(k^2 - i\epsilon)^2} = \frac{1}{k^2 - i\epsilon}(P^T_{\mu\nu} + \xi P^L_{\mu\nu})$$

Here,

$$P^T_{\mu\nu} \equiv \eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \quad P^L_{\mu\nu} \equiv \frac{k_\mu k_\nu}{k^2}$$

We can derive that

$$(iG_F)^{-1}(k)_{\mu\nu} = k^2(P^T_{\mu\nu} + \frac{1}{\xi}P^L_{\mu\nu})$$

We may also expand $i\Pi^{\mu\nu}$ as

$$\Pi^{\mu\nu} = P_T^{\mu\nu} f_T(k^2) + P_L^{\mu\nu} f_L(k^2) = \eta^{\mu\nu} f_T + \frac{k^\mu k^\nu}{k^2}(f_L - f_T)$$

Therefore,

$$(i\mathcal{G})^{-1}(k)_{\mu\nu} = (k^2 - f_T(k^2))P^T_{\mu\nu} + (\frac{k^2}{\xi} - f_L(k^2))P^L_{\mu\nu}$$

$$\mathcal{G}(k)_{\mu\nu} = \frac{-i}{k^2 - f_T(k^2)}P^T_{\mu\nu} + \frac{-i}{\frac{k^2}{\xi} - f_L(k^2)}P^L_{\mu\nu}$$

We observe that if $f_{T,L}(k^2 = 0) \neq 0$, a mass will be generated for the photon. Because $\Pi(k)$ comes from 1PI diagrams, it should not be singular at $k^2 = 0$, and so $f_L - f_T = O(k^2)$, as $k \to 0$. We will show that gauge invariance ensures that no mass is generated from the loop corrections

14.6.2 Ward identities(2)

We define the generating functional for connected diagrams

$$Z[J,\eta,\overline{\eta}] = e^{-iE[J,\eta,\overline{\eta}]}$$

For example,

$$\mathcal{G}(x-y)_{\mu\nu} = -i\frac{\delta^2 E[J,\eta,\overline{\eta}]}{\delta J^\mu(x)\delta J^\nu(y)}\bigg|_{J,\eta,\overline{\eta}=0}$$

Recall, for infinitesimal gauge transformations, $\delta A_\mu = \partial_\mu\lambda$, $\delta\Psi = ie_0\lambda\Psi$ and $\delta\overline{\Psi} = -ie_0\lambda\overline{\Psi}$. For a change of variables in the path integral, $Z[J,\eta,\overline{\eta}]$ will remain the same. Recall that

$$Z[J,\eta,\overline{\eta}] = W(\xi)\int\mathcal{D}A\mathcal{D}\overline{\Psi}\mathcal{D}\Psi e^{i\int d^4x[\mathcal{L}_0+\mathcal{L}_1+JA+\overline{\eta}\Psi+\overline{\Psi}\eta]}$$

The change of action is

$$\delta S = -\frac{1}{\xi}\int d^4x\partial_\mu A^\mu\partial^2\lambda + \int d^4xJ^\mu\partial_\mu\lambda + ie_0\overline{\eta}\Psi\lambda - ie_0\overline{\Psi}\eta\lambda$$

Hence, we must have

$$\int d^4x\lambda(x)W(\xi)\int\mathcal{D}A\mathcal{D}\overline{\Psi}\mathcal{D}\Psi e^{iS}\left[-\frac{1}{\xi}\partial^2\partial_\mu A^\mu - \partial_\mu J^\mu + ie_0(\overline{\eta}\Psi - \overline{\Psi}\eta)\right]$$

Since

$$\langle A_\mu(x)\rangle_{J,\eta,\overline{\eta}} = -\frac{\delta E}{\delta J^\mu} \quad \langle\Psi(x)\rangle_{J,\eta,\overline{\eta}} = -\frac{\delta E}{\delta\overline{\eta}} \quad \langle\overline{\Psi}(x)\rangle_{J,\eta,\overline{\eta}} = \frac{\delta E}{\delta\eta}$$

The above equation can be written as

$$\frac{1}{\xi}\partial^2\partial^\mu\frac{\delta E}{\delta J^\mu} - \partial_\mu J^\mu - ie_0\left[\overline{\eta}\frac{\delta E}{\delta\overline{\eta}} + \frac{\delta E}{\delta\eta}\eta\right] = 0$$

We can derive that

$$\frac{1}{\xi}\partial^2\partial^\mu\frac{\delta^2 E[J,\eta,\overline{\eta}]}{\delta J^\mu(x)\delta J^\nu(y)}\bigg|_{J,\eta,\overline{\eta}=0} - \partial_\nu\delta(x-y) = 0$$

that is,

$$\frac{i}{\xi}\partial^2\partial^\mu\mathcal{G}(x-y)_{\mu\nu} + \partial_\nu\delta(x-y) = 0$$

or, written in momentum-space,

$$-\frac{i}{\xi}k^2k^\mu\mathcal{G}(k)_{\mu\nu} + k_\nu = 0$$

So

$$-\frac{k^2}{k^2-\xi f_L(k^2)}k_\nu + k_\nu = 0$$

Which means $f_L(k^2) = 0$ and so, we have $f_T(k^2)\to O(k^2)$ as $k^2\to 0$. The exact propagator of photon is

$$\mathcal{G}(k)_{\mu\nu} = \frac{-i}{k^2(1-\pi(k^2))}P^T_{\mu\nu} + \frac{-i\xi}{k^2}P^L_{\mu\nu}$$

where $\pi(k^2)\equiv\frac{f_T(k^2)}{k^2}$

14.7 LSZ reduction formula

14.7.1 LSZ reduction formula and Feynman rules

Suppose that the probability for the quantum field to create or annihilate an exact one-particle eigenstate of H is Z_3, i.e.

$$\langle\Omega|A(0)|p,\lambda\rangle = \sqrt{Z_3}\epsilon_\lambda(p)$$

In Feynman gauge, because the norm of $|p,0\rangle$ is negative, the expansion of orthogonal complete set will be written as

$$\frac{d^3q}{(2\pi)^3}\frac{1}{2E_q}\eta^{\lambda\lambda'}|p,\lambda\rangle\langle p,\lambda'|$$

We have demonstrated that photon will remain massless when interacting with charged fermions. And recall that $\sum_\lambda \xi_\lambda(p)\xi^*_\lambda(p) = \eta_{\mu\nu}$. So, we can derive by the similar method in ϕ^4 theory that

$$\int d^4x\, e^{-ipx}\langle\Omega|TA_\mu(x)A_\nu(0)|\Omega\rangle_C = \frac{-iZ_3\eta_{\mu\nu}}{p^2-i\epsilon}+\cdots$$

The LSZ reduction formula for fermions would take the form as

Theorem 14.1 LSZ reduction formula

$$\begin{aligned} & \langle \boldsymbol{p}_1\cdots\boldsymbol{p}_n\,|S|\,\boldsymbol{k}_1\cdots\boldsymbol{k}_m\,\rangle \\ = \quad & \prod_1^n\int d^4x_i e^{-ip_ix_i}\prod_1^m\int d^4y_j e^{ik_jy_j} \\ \times \quad & \left(\frac{i}{\sqrt{Z_3}}\right)^{m+n} [p_1^2\epsilon^{*\mu_1}_{\lambda_1}(p_1)]\cdots[p_n^2\epsilon^{*\mu_n}_{\lambda_n}(p_n)][k_1^2\epsilon^{\nu_1}_{\lambda'_1}(k_1)]\cdots[k_m^2\epsilon^{\nu_m}_{\lambda'_m}(p_m)] \\ \times \quad & \langle\Omega|T\{A_{\mu_1}(x_1)\cdots A_{\mu_n}(x_n)A_{\nu_1}(y_1)\cdots A_{\nu_m}(y_m)\}|\Omega\rangle \end{aligned}$$

♣

The LSZ reduction formula in other gauge would give similar procedure for calculating scattering amplitude: Fourier transform the Green function in position space to momentum space, cut-off the external legs and multiply the polarization vector of asymptotic states. (Note there is still an extra factor $\sqrt{Z_3}^{m+n}$ to multiply, similar to the ϕ^4 theory).
Finally, we list the Feymann rules of QED in momentum space as follows:

1. For each incoming electron, draw a solid line with an arrow pointed towards the vertex, and label it with the electron's four-momentum, p_i.

2. For each outgoing electron, draw a solid line with an arrow pointed away from the vertex, and label it with the electron's four-momentum, p'_i.

3. For each incoming positron, draw a solid line with an arrow pointed away from the vertex, and label it with minus the positron's four-momentum, $-p_i$.

4. For each outgoing positron, draw a solid line with an arrow pointed towards the vertex, and label it with minus the positron's four-momentum, $-p'_i$.

5. For each incoming photon, draw a wavy line with an arrow pointed towards the vertex, and label it with the photon's four-momentum,k_i.

6. For each outgoing photon, draw a wavy line with an arrow pointed away from the vertex, and label it with the photon's four-momentum, k_i'.

7. The only allowed vertex joins two solid lines, one with an arrow pointing towards it and one with an arrow pointing away from it, and one wavy line. Using this vertex, join up all the external lines, including extra internal lines as needed. In this way, draw all possible diagrams that are topologically inequivalent.

8. Assign each internal line its own four-momentum. Think of the four-momenta as flowing along the arrows, and conserve four-momentum at each vertex.

9. The value of a diagram consists of the following factors:

   - for each incoming photon, $\epsilon_\lambda^\mu(k)$; for each outgoing photon, $\epsilon_\lambda^{*\mu}(k)$;
   - for each incoming electron, $u_r(\boldsymbol{k})$; for each outgoing electron, $\overline{u}_s(\boldsymbol{p})$;
   - for each incoming positron, $\overline{v}_{\overline{r}}(\overline{\boldsymbol{k}})$; for each outgoing positron, $v_{\overline{s}}(\overline{\boldsymbol{p}})$;
   - for each vertex, $ie_0\gamma^\mu$; for each internal photon, $G_F(p)$; for each internal fermion, $S_F(p)$.

10. Spinor indices are contracted by starting at one end of a fermion line: specifically, the end that has the arrow pointing away from the vertex. The factor associated with the external line is either $\bar{u}$ or $\bar{v}$. Go along the complete fermion line, following the arrows backwards, and write down (in order from left to right) the factors associated with the vertices and propagators that you encounter. The last factor is either a u or v. Repeat this procedure for the other fermion lines, if any. The vector index on each vertex is contracted with the vector index on either the photon propagator (if the attached photon line is internal) or the photon polarization vector (if the attached photon line is external).

11. The overall sign of a tree diagram is determined by drawing all contributing diagrams in a standard form: all fermion lines horizontal, with their arrows pointing from left to right, and with the left endpoints labeled in the same fixed order (from top to bottom); if the ordering of the labels on the right endpoints of the fermion lines in a given diagram is an even (odd) permutation of an arbitrarily chosen fixed ordering, then the sign of that diagram is positive (negative)

12. Each closed fermion loop contributes an extra minus sign.

13. Value of $i\mathcal{M}$ is given by a sum over the values of the contributing diagrams.

14. $\langle f|S|i\rangle = (Z_2)^{\frac{n_f}{2}}(Z_3)^{\frac{n_p}{2}}i\mathcal{M}\delta(\sum p_f - \sum p_i)$

14.7.2 Ward Takahashi identity (3)

Suppose the invariant matrix element for a process is $\mathcal{M}$, if we replace the polarization state vector ϵ_λ^μ (or $\epsilon_\lambda^{*\mu}$) of one incoming (or outgoing) photon with its momentum vector k^μ, we have

$$k^\mu \mathcal{M}_\mu = 0$$

Proof: Without losing generality, we can consider a physical process with a single incoming and outgoing fermion lines respectively. So, the ward identities states that

$$-ik_\mu F^\mu(k;p,q) = ie_0\left[F_0(p+k,q) - F_0(p,q-k)\right]$$

Here, F keeps the external fermion legs but cuts external photon lines. According to the LSZ reduction formula, from each diagram a contribution to an S matrix element by taking the coefficient of the product of poles

$$\left(\frac{-i}{\not{p}+m}\right)\left(\frac{-i}{\not{q}+m}\right)$$

But the terms on the right hand side contain one of these poles, but neither contains both poles. So they contribute nothing to S-matrix. So, we can have

$$k^\mu \mathcal{M}_\mu = 0$$

□

We also note that when calculating invariant matrix element, the main difference between different gauge are photon propagator. In Coulomb gauge, we have

$$G_F(p)_{\mu\nu} \equiv \begin{cases} \frac{i}{|\boldsymbol{p}|^2} & \mu,\nu = 0 \\ \frac{-i}{p^2-i\epsilon}\left(\delta_{ij} - \frac{p_i p_j}{|\boldsymbol{p}|^2}\right) & \mu = i \neq 0, \nu = j \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

In Lorentz gauge, we have

$$G_F(p)_{\mu\nu} = \frac{-i\eta_{\mu\nu}}{p^2 - i\epsilon} + i(1-\xi)\frac{p_\mu p_\nu}{(p^2-i\epsilon)^2}$$

We now argue that the lead to the same $\mathcal{M}$ element. For a general process, it can be represented as follows.

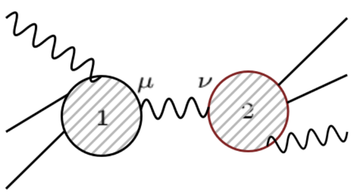


Figure 14.6: Feynman diagram representation of a QED process

The value of the diagram is

$$\mathcal{M}_1^\mu G_F(k)_{\mu\nu} \mathcal{M}_2^\nu$$

and $k_\mu \mathcal{M}_1^\mu = 0$, $k_\nu \mathcal{M}_2^\nu = 0$. So the factor ξ in Lorentz gauge is irrelevant to the value of $\mathcal{M}$. As for Coulomb gauge, denote $\mathcal{M}_1^\mu$ as α^μ, $\mathcal{M}_2^\mu$ as β^μ, so

$$\alpha^\mu G_F(k)_{\mu\nu}\beta^\nu = i\left(-\frac{\boldsymbol{\alpha}\cdot\boldsymbol{\beta}}{k^2} + \frac{(\boldsymbol{\alpha}\cdot\boldsymbol{k})(\boldsymbol{\beta}\cdot\boldsymbol{k})}{k^2\boldsymbol{k}^2} + \frac{\alpha^0\beta^0}{\boldsymbol{k}^2}\right)$$

Using the fact that $\boldsymbol{\alpha}\cdot\boldsymbol{k} + \alpha^0 k_0 = 0$, we can verify that

$$\alpha^\mu G_F(k)_{\mu\nu}\beta^\nu = \alpha^\mu\left(-\frac{i\eta_{\mu\nu}}{k^2}\right)\beta^\nu$$

So, the invariant matrix element is gauge invariant.

14.8 Renormalization

14.8.1 Renormalized quantum electrodynamics

The Lagrangian of QED is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\Psi}(i\partial\!\!\!/ - m_0)\Psi + e_0 j^\mu A_\mu,$$

Suppose the number of external photons is N_γ, the number of external fermions is N_e, the number of vertex is V. The superficial divergence is

$$D = 4 - N_\gamma - \frac{3}{2}N_e$$

Recall that C is a symmetry of QED, so $C|\Omega\rangle = |\Omega\rangle$. But j^μ changes sign under charge conjugation, $Cj^\mu(x)C^\dagger = -j^\mu(x)$, so its vacuum expectation value must vanish:

$$\langle\Omega|Tj^\mu(x)|\Omega\rangle = \langle\Omega|TC^\dagger Cj^\mu(x)C^\dagger C|\Omega\rangle = -\langle\Omega|Tj^\mu(x)|\Omega\rangle = 0$$

So, the amplitude with $N_\gamma = 1, N_e = 0$ will vanish. Similarly, we can verify that the diagram with $N_\gamma = 3, N_e = 0$ will also vanish.

As for the amplitude with $N_\gamma = 4$, the superficial divergence $D = 0$. So the only divergence must be of the form $\log\Lambda$. But the ward identity implies that

$$K^\mu \mathcal{M}_{\mu\nu\sigma\rho} = 0$$

So, the divergent terms must vanish.

If we neglect the vacuum term with $N_\gamma = 0, N_e = 0$, there are only three divergent amplitude terms left.

We need four counter terms to eliminate all the divergence. The Lagrangian can be written as

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_{ct}$$

Here,

$$\mathcal{L}_1 = -\frac{1}{4}F_r^{\mu\nu}F_{r\mu\nu} + i\overline{\Psi}_r\gamma^\mu\partial_\mu\Psi_r - m\overline{\Psi}_r\Psi_r + e\overline{\Psi}_r\gamma^\mu\Psi_r A_{r\mu}$$

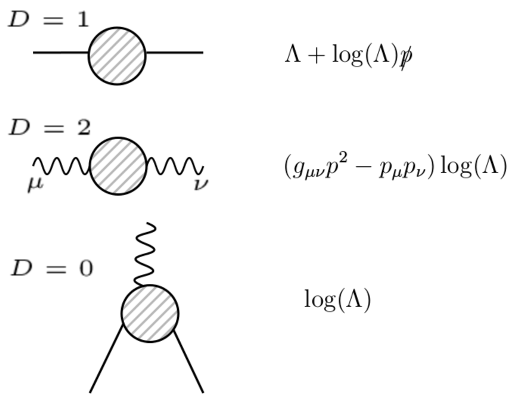


Figure 14.7: Feynman diagram representation of divergent amplitude in QED

$$\mathcal{L}_{ct} = -\frac{1}{4}\delta_3 F_r^{\mu\nu}F_{r\mu\nu} + i\delta_2\overline{\Psi}_r\gamma^\mu\partial_\mu\Psi_r - \delta_m\overline{\Psi}_r\Psi_r + e\delta_1\overline{\Psi}_r\gamma^\mu\Psi_r A_{r\mu}$$

$$A = \sqrt{Z_3}A_r \quad \Phi = \sqrt{Z_2}\Phi_r \quad \delta_3 = Z_3 - 1 \quad \delta_2 = Z_2 - 1$$

$$\delta_m = Z_2 m_0 - m \quad \delta_1 = Z_1 - 1 = (e_0/e)Z_2 Z_3^{\frac{1}{2}} - 1$$

The Feynman rules of counter terms are:

$$-\frac{1}{4}\delta_3 F_r^{\mu\nu}F_{r\mu\nu} \qquad -i(\eta^{\mu\nu}q^2 - q^\mu q^\nu)\delta_3$$

$$i\overline{\Psi}_r(\delta_2\not{\partial} - \delta_m)\Psi_r \qquad -i(\delta_2\not{p} + \delta_m)$$

$$e\delta_1\overline{\Psi}_r\gamma^\mu\Psi_r A_{r\mu} \qquad ie\gamma^\mu\delta_1$$

We also denote the renormalized 1PI component of exact propagator of photon as $i(\eta^{\mu\nu}q^2 - q^\mu q^\nu)\Pi_r(q^2)$, the renormalized 1PI component of exact propagator of fermion as $-i\Sigma_r(\not{p})$, the renormalized exact amputated photon-fermion-antifermion vertex as $ie\Gamma_r^\mu(p, p')$.
So, renormalized exact propagator of photon is

$$\mathcal{G}_r(q)_{\mu\nu} = \frac{-i}{q^2(1 - \Pi_r(q^2))}P_{\mu\nu}^T$$

renormalized exact propagator of fermion is

$$\mathcal{S}_r(p) = \frac{-i}{\not{p} + m + \Sigma_r(\not{p})}$$

The on-shell renormalization conditions are

$$\begin{aligned}
\Sigma_r(\not{p} = -m) &= 0 \\
\frac{d}{d\not{p}}\Sigma_r(\not{p})\Big|_{\not{p}=-m} &= 0 \\
\Pi_r(q^2 = 0) &= 0 \\
ie\Gamma_r^\mu(p = p', p^2 = -m^2) &= ie\gamma^\mu
\end{aligned}$$

Recall the ward identity, we have

$$ie\sqrt{Z_2}\partial_\mu\langle\Omega|Tj_r^\mu\Psi_r(x_1)\overline{\Psi}_r(x_2)|\Omega\rangle = -ie\delta(x-x_1)\langle\Omega|T\Psi_r(x_1)\overline{\Psi}_r(x_2)|\Omega\rangle + ie\delta(x-x_2)\langle\Omega|T\Psi_r(x_1)\overline{\Psi}_r(x_2)|\Omega$$

In momentum space, we have

$$-k_\mu Z_2 Z_1^{-1}\mathcal{S}_r(p+k)[ie\Gamma_r^\mu(p+k,p)]\mathcal{S}_r(p)\frac{1}{1-\Pi_r(k^2)} = e(\mathcal{S}_r(p+k)-\mathcal{S}_r(p))$$

So,

$$Z_2 Z_1^{-1}k_\mu[\Gamma_r^\mu(p+k,p)]\frac{1}{1-\Pi_r(k^2)} = \not{k} + \Sigma_r(\not{k}+\not{p}) - \Sigma_r(\not{p})$$

Since Γ_r, Π_r and Σ_r are all finite by renormalization, so Z_1/Z_2 must be finite. In $\overline{MS}$ renormalization scheme, we immediately get

$$Z_1 = Z_2$$

In on-shell renormalization scheme, taking the limit of $k \to 0$, we can also get that

$$Z_1 = Z_2$$

So, we know

$$e = \sqrt{Z_3}e_0$$

This means that the relation between the bare and renormalized electric charge depends only on the photon field strength renormalization, not on quantities particular to the fermions, leading to a universal electric charge which has the same value for all species.
In the following subsection, we would omit the subscript r unless it is necessary to emphasis the difference of bare field and renormalized fields.

14.8.2 One loop structure of QED

Photon propagator

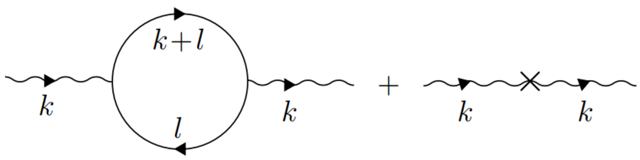


Figure 14.8: The one-loop and counterterm corrections to the photon propagator

$$\Pi(k^2) = -\frac{e^2}{\pi^2}\int_0^1 dx x(1-x)\left[\frac{1}{\epsilon} - \frac{1}{2}\ln(\frac{D}{\mu^2})\right] - \delta_3 + O(e^4)$$

where $D = x(1-x)k^2 + m^2 - i\epsilon$ and $\mu^2 = 4\pi e^{-\gamma}\tilde{\mu}^2$. Impose the OS renormalization condition $\Pi(0) = 0$, we have

$$\delta_3 = -\frac{e^2}{6\pi^2}\left[\frac{1}{\epsilon} - \ln(\frac{m}{\mu})\right] + O(e^4)$$

$$\Pi(k^2) = \frac{e^2}{2\pi^2}\int_0^1 dx x(1-x)\ln(\frac{D}{m^2}) + O(e^4)$$

Fermion propagator

The exact renormalized fermion propagator in OS renormalization can be written in Lehmann–Kallen form as

$$i\mathcal{S}(\not p) = \frac{1}{\not p + m - i\epsilon} + \int_{m_{th}}^{\infty} ds \frac{\rho_\Psi(s)}{\not p + \sqrt{s} - i\epsilon}$$

We see that the first term has a pole at $\not p = -m$ with residue one. This residue corresponds to the field normalization that is needed for the validity of the LSZ formula.

There is a problem, however: in quantum electrodynamics, the threshold mass m_{th} is m, corresponding to the contribution of a fermion and a zero energy photon. Thus the second term has a branch point at $\not p = -m$. The pole in the first term is therefore not isolated, and its residue is ill defined.

This is a reflection of an underlying infrared divergence, associated with the massless photon. To deal with it, we must impose an infrared cutoff that moves the branch point away from the pole. The most direct method is to change the denominator of the photon propagator from k^2 to $k^2 + m_\gamma^2$, where m_γ is a fictitious photon mass. Ultimately, we must deal with this issue by computing cross sections that take into account detector inefficiencies. In quantum electrodynamics, we must specify the lowest photon energy ω_{min} that can be detected. Only after computing cross sections with extra undetectable photons, and then summing over them, is it safe to take the limit $m_\gamma \to 0$.

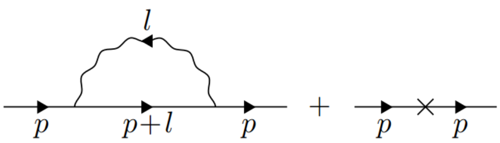


Figure 14.9: The one-loop and counterterm corrections to the fermion propagator

$$\Sigma(\not p) = \frac{e^2}{8\pi^2}\int_0^1 dx\,\big((2-\epsilon)(1-x)\not p + (4-\epsilon)m\big)\left[\frac{1}{\epsilon} - \frac{1}{2}\ln\left(\frac{D}{\mu^2}\right)\right] + \delta_2 \not p + \delta_m + O(e^4)$$

where $D = x(1-x)p^2 + xm^2 + (1-x)m_\gamma^2$. The fitness of $\Sigma(\not p)$ requires that

$$\delta_2 = -\frac{e^2}{8\pi^2}\left(\frac{1}{\epsilon} + \text{ finite}\right) + O(e^4)$$

$$\delta_m/m = -\frac{e^2}{2\pi^2}\left(\frac{1}{\epsilon} + \text{ finite}\right) + O(e^4)$$

Impose the OS renormalization condition $\Sigma(-m) = 0$ and $\Sigma'(-m) = 0$, we have

$$\Sigma(\not p) = -\frac{e^2}{8\pi^2}\int_0^1 dx\,\big((1-x)\not p + 2m\big)\ln\left(\frac{D}{D_0}\right) + \kappa_2(\not p + m) + O(e^4)$$

where $D_0 = x^2m^2 + (1-x)m_\gamma^2$ and $\kappa_2 = -2\ln(m/m_\gamma) + 1$.

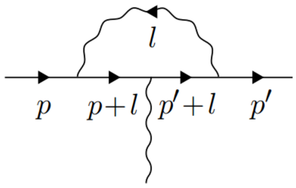


Figure 14.10: The one-loop correction to the photon-fermion-fermion vertex

Vertex

$$\Gamma^\mu(p,p') = (1+\delta_1)\gamma^\mu + \frac{e^2}{8\pi^2}\left[\left(\frac{1}{\epsilon} - 1 - \frac{1}{2}\int dF_3 \ln(D/\mu^2)\right)\gamma^\mu + \frac{1}{4}\int dF_3 \frac{\tilde{N}^\mu}{D}\right] + O(e^4)$$

where

$$\int dF_3 = 2\int_0^1 dx_1 dx_2 dx_3 \delta(x_1+x_2+x_3-1)$$
$$D = x_1(1-x_1)p^2 + x_2(1-x_2)p'^2 - 2x_1x_2 p\cdot p' + (x_1+x_2)m^2 + x_3 m_\gamma^2$$
$$\tilde{N}^\mu = \gamma_\nu[x_1\not{p} - (1-x_2)\not{p}' + m]\gamma^\mu[-(1-x_1)\not{p} + x_2\not{p}' + m]\gamma^\nu$$

Fitness of Γ^μ requires that

$$\delta_1 = -\frac{e^2}{8\pi^2}(\frac{1}{\epsilon} + \text{ finite }) + O(e^4)$$

Impose the OS renormalization condition $\Gamma_r^\mu(p=p', p^2=-m^2) = \gamma^\mu$, we have

$$\Gamma^\mu(p,p') = \gamma^\mu - \frac{e^2}{16\pi^2}\int dF_3\left[\left(\ln(D/D_0) + 2\kappa_1\right)\gamma^\mu - \frac{\tilde{N}^\mu}{2D}\right] + O(e^4)$$

where

$$D_0 = (1-x_3)^2 m^2 + x_3 m_\gamma^2 \quad \kappa_1 = -2\ln(m/m_\gamma) + \frac{5}{2}$$

Vertex function

Consider the process of electron-electron scattering In order to compute the contribution of these diagrams, we must evaluate $\overline{u}_{s'}(\boldsymbol{p}')\Gamma^\mu(p',p)u_s(\boldsymbol{p})$ with $p^2 = p'^2 = -m^2$, but with $q^2 = (p-p')^2$ arbitrary. Finally,we have

$$\overline{u}_{s'}(\boldsymbol{p}')\Gamma^\mu(p',p)u_s(\boldsymbol{p}) = \overline{u}'\left[F_1(q^2)\gamma^\mu - \frac{i}{m}F_2(q^2)S^{\mu\nu}q_\nu\right]u_s(\boldsymbol{p})$$

where we have defined the form factors

$$\begin{aligned} F_1(q^2) &= 1 - \frac{e^2}{16\pi^2}\int dF_3 \ln\left(1 + \frac{x_1x_2q^2/m^2}{(1-x_3)^2}\right) \\ &- \frac{e^2}{16\pi^2}\int dF_3 \frac{1-4x_3+x_3^2}{(1-x_3)^2 + x_3 m_\gamma^2/m^2} + \frac{(x_3+x_1x_2)q^2/m^2 - (1-4x_3+x_3^2)}{x_1x_2q^2/m^2 + (1-x_3)^2 + x_3m_\gamma^2/m^2} + O(e^4) \\ F_2(q^2) &= \frac{e^2}{8\pi^2}\int dF_3 \frac{x_3 - x_3^2}{x_1x_2q^2/m^2 + (1-x_3)^2} + O(e^4) \end{aligned}$$

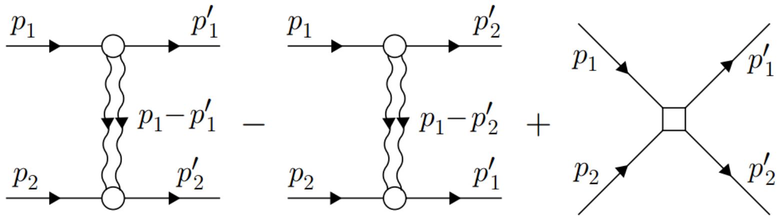


Figure 14.11: Diagrams for the exact electron-electron scattering amplitude. The vertices and photon propagator are exact; external lines stand for the usual u and $\bar{u}$ spinor factors, times the unit residue of the pole at $p^2 = -m^2$.

The loop corrections of QED are related to anomalous magnetic moment of electrons and lamb shift of energy level in hydrogen atom. We will not discuss these interesting phenomenon in this notes. Discussion on anomalous magnetic moment can be found in chapter 6.2 of *An introduction to quantum field theory (M.E.Peskin & D.V.Schroeder)*. You can also find some comments on lamb shift due to vacuum polarization in chapter 7.5 . And a detailed treatment of infrared divergence is available in chapter 6.1, 6.4 and 6.5.

14.8.3 Renormalization group

$\overline{MS}$ renormalization scheme

$$\Pi(k^2) = \frac{e^2}{2\pi^2}\int_0^1 dx\, x(1-x)\ln(\frac{x(1-x)k^2+m^2}{\mu^2}) + O(e^4)$$

$$Z_3 = 1+\delta_3 = 1 - \frac{e^2}{6\pi^2}\frac{1}{\epsilon} + O(e^4)$$

$$\Sigma(\not{p}) = -\frac{e^2}{8\pi^2}\int_0^1 dx \ln(\frac{x(1-x)p^2+xm^2+(1-x)m_\gamma^2}{\mu^2})\left((1-x)\not{p}+2m\right) + \frac{1}{2}\not{p} + m + O(e^4)$$

$$Z_2 = 1+\delta_2 = 1 - \frac{e^2}{8\pi^2}\frac{1}{\epsilon} + O(e^4)$$

$$Z_m = 1+\delta_m/m = 1 - \frac{e^2}{2\pi^2}\frac{1}{\epsilon} + O(e^4)$$

$$\Gamma^\mu(p,p') = \gamma^\mu + \frac{e^2}{8\pi^2}\left[-\left(1+\frac{1}{2}\int dF_3 \ln(D/\mu^2)\right)\gamma^\mu + \frac{1}{4}\int dF_3 \frac{\tilde{N}^\mu}{D}\right] + O(e^4)$$

$$D = x_1(1-x_1)p^2 + x_2(1-x_2)p'^2 - 2x_1x_2 p\cdot p' + (x_1+x_2)m^2 + x_3 m_\gamma^2$$

$$Z_1 = 1+\delta_1 = 1 - \frac{e^2}{8\pi^2}\frac{1}{\epsilon} + O(e^4)$$

Renormalization group

The Lagrangian of QED is

$$\mathcal{L} = -\frac{1}{4}F_{0\mu\nu}F_0^{\mu\nu} + \overline{\Psi}_0(i\not{\partial} - m_0)\Psi_0 + e_0\overline{\Psi}_0\gamma^\mu\Psi_0 A_{0\mu},$$

It can be written as

$$\mathcal{L} = -\frac{1}{4} Z_3 F_{\mu\nu} F^{\mu\nu} + \overline{\Psi}(i Z_2 \partial\!\!\!/ - Z_m m)\Psi + Z_1 e \overline{\Psi}\gamma^\mu \Psi A_\mu,$$

So,

$$\Psi_0 = Z_2^{1/2}\phi \quad m_0 = Z_2^{-1} Z_m m \quad A_0 = Z_3^{1/2} A \quad e_0 = Z_2^{-1} Z_1 Z_3^{-1/2} e\tilde{\mu}^{\epsilon/2} = Z_3^{-1/2} e\tilde{\mu}^{\epsilon/2}$$

After using dimensional regularization, the infinities coming from loop integrals take the form of inverse powers of ϵ. In the $\overline{\text{MS}}$ renormalization scheme, we choose the Zs to cancel off these powers of $1/\epsilon$, and nothing more. Therefore the Zs can be written as

$$Z_1 = 1 + \sum_{n=1}^{\infty} \frac{a_n(\lambda)}{\epsilon^n} \quad Z_2 = 1 + \sum_{n=1}^{\infty} \frac{b_n(\lambda)}{\epsilon^n}$$
$$Z_3 = 1 + \sum_{n=1}^{\infty} \frac{c_n(\lambda)}{\epsilon^n} \quad Z_m = 1 + \sum_{n=1}^{\infty} \frac{d_n(\lambda)}{\epsilon^n}$$

In QED, $a_1 = -\frac{e^2}{8\pi^2} + O(e^4)$, $b_1 = -\frac{e^2}{8\pi^2} + O(e^4)$,$c_1 = -\frac{e^2}{6\pi^2} + O(e^4)$,$d_1 = -\frac{e^2}{2\pi^2} + O(e^4)$
Remember that bare fields and parameters must be independent of μ.Define

$$E(e,\epsilon) \equiv \ln(Z_3^{-1/2}) = \sum_{n=1}^{\infty} \frac{E_n(e)}{\epsilon^n}$$

We can calculate $E_1 = -\frac{1}{2}c_1 = \frac{e^2}{12\pi^2} + O(e^4)$. As $\ln e_0 = E + \ln e + \frac{\epsilon}{2}\ln\tilde{\mu}$. From the independence of e_0, we can derive

$$\left(1 + \frac{eE_1'}{\epsilon} + \cdots\right)\frac{de}{d\ln\mu} + \frac{\epsilon}{2}e = 0$$

In a renormalizable theory, we should have

$$\frac{de}{d\ln\mu} = -\frac{\epsilon}{2}e + \beta(e)$$

So

$$\beta(e) = \frac{e^2}{2}E_1'(\lambda)$$

In QED, we have

$$\beta(e) = \frac{e^3}{12\pi^2} + O(e^5)$$

Define

$$M(e,\epsilon) \equiv \ln(Z_m Z_2^{-1}) = \sum_{n=1}^{\infty} \frac{M_n(\lambda)}{\epsilon^n}$$

We can calculate $M_1 = d_1 - b_1 = -\frac{3e^2}{8\pi^2} + O(e^4)$. As $\ln m_0 = M + \ln m$, define the anomalous dimension of the mass

$$\gamma_m(e) \equiv \frac{1}{m}\frac{dm}{d\ln\mu}$$

From the independence of m_0, we can derive

$$\gamma_m(e) = \frac{e}{2}M_1'$$

In QED theory, we have

$$\gamma_m(e) = -\frac{3e^2}{8\pi^2} + O(e^4)$$

Expand $\ln Z_2$ as

$$\ln Z_2 = \frac{b_1}{\epsilon} + \cdots$$

Define the anomalous dimension of the fermion field

$$\gamma_2(e) \equiv \frac{1}{2}\frac{d\ln Z_2}{d\ln\mu}$$

We can derive

$$\gamma_2(e) = -\frac{1}{4}eb_1'$$

In QED theory, we have

$$\gamma_2(e) = \frac{e^2}{16\pi^2} + O(e^4)$$

Expand $\ln Z_3$ as

$$\ln Z_3 = \frac{c_1}{\epsilon} + \cdots$$

Define the anomalous dimension of the EM field

$$\gamma_3(e) \equiv \frac{1}{2}\frac{d\ln Z_3}{d\ln\mu}$$

We can derive

$$\gamma_3(e) = -\frac{1}{4}ec_1'$$

In QED theory, we have

$$\gamma_3(e) = \frac{e^2}{12\pi^2} + O(e^4)$$