Particle Physics Phenomenology

Yue Zhang

Department of Physics
Carleton University
1125 Colonel By Dr, Ottawa
Ontario K1S 5B6, Canada

These are my personal lecture notes for graduate course PHYS 6601 at Carleton University. It covers the minimal set (in my opinion) of concepts and techniques you need to know about the Standard Model of elementary particles.

November 30, 2022

Contents

1 QFT and Symmetries

This section reviews some of the results discussed in course PHYS 5702 Relativistic Quantum Mechanics. We will put emphasis on the continuous internal symmetries of theories.

1.1 Classical field theory

1.1.1 Lagrangian formalism

A classical field is a function of space and time. A spin zero scalar field is denoted as $\phi(x)$, where x^μ is a Lorentz vector. The Lagrangian density is a function of ϕ and its first derivatives, $\mathcal{L}(\phi(x), \partial_\mu\phi(x))$. It could also depend on higher derivatives of ϕ, but we will not consider.

The Euler-Lagrange equation is derived for ϕ configurations that minimize the action

$$S[\phi] = \int d^4x \mathcal{L}(\phi(x), \partial_\mu\phi(x)) \ . \tag{1}$$

With the argument x integrated over, the resulting S is a functional of ϕ. A generic variation of $\phi(x) \to \phi(x) + \delta\phi(x)$ will cause S to change,

$$\delta S = S[\phi + \delta\phi] - S[\phi] = \int d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi)\right] \ . \tag{2}$$

Using $\delta(\partial_\mu\phi) = \partial_\mu(\delta\phi)$, and integrate by parts, we get

$$\delta S = \int d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right]\delta\phi(x) + \int d^4x\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) \ . \tag{3}$$

The last term is an integral over total derivative, and can be written as a surface integral at spacetime infinity. Requiring δS to vanish in the bulk with arbitrary $\delta\phi(x)$ leads to the equation of motion for field $\phi(x)$,

$$\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right) = 0 \ . \tag{4}$$

As the simplest example, we consider a free real scalar field theory with Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 \ . \tag{5}$$

Applying the above equation of motion, we get

$$(\Box + m^2)\phi(x) = 0 \ , \tag{6}$$

where $\Box = \partial_\mu\partial^\mu$. This is the Klein-Gordon equation for a free scalar field.

1.1.2 Hamiltonian formalism

In the Hamiltonian formalism, we first introduce the canonical momentum for the field ϕ, called π, defined as

$$\pi(x) = \frac{\delta L}{\delta\dot{\phi}} \ , \tag{7}$$

where $\dot{\phi} \equiv \partial\phi/\partial t$ and $L = \int d^3x\mathcal{L}$. Next, introduce the Hamiltonian

$$H[\phi, \pi] = \int d^3x\pi(x)\dot{\phi}(x) - L[\phi, \partial_\mu\phi] \ . \tag{8}$$

As shown by its arguments, H is a functional of ϕ and π. It is straightforward to show $\delta H/\delta\dot{\phi} = 0$ by using Eq. (7).

The mechanics of the classical field is dictated by the Hamilton's equations

$$\frac{\delta H}{\delta\phi} = -\dot{\pi} \ , \quad \frac{\delta H}{\delta\pi} = \dot{\phi} \ . \tag{9}$$

As an example, consider again the free real scalar field theory. The canonical momentum is $\pi(x) = \dot{\phi}(x)$. In turn, the Hamiltonian takes the form

$$H = \int d^3x \frac{1}{2}\left(\pi^2 + \nabla\phi\cdot\nabla\phi + m^2\phi^2\right) \ . \tag{10}$$

Next, we apply the first Hamilton's equation, by taking variation of H with respect to ϕ but holding π fixed,

$$\begin{aligned} \delta H &= \int d^3x \left[\nabla\phi\cdot\nabla(\delta\phi) + m^2\phi\delta\phi\right] \\ &= \int d^3x \left[\nabla\cdot((\nabla\phi)\delta\phi) - (\nabla^2\phi)\delta\phi + m^2\phi\delta\phi\right] \ . \end{aligned} \tag{11}$$

By dropping the surface integral term at space infinity, we find $\delta H/\delta\phi = (-\nabla^2 + m^2)\phi$. As a result, the first Hamilton's equation implies (also use $\dot{\pi} = \ddot{\phi}$)

$$\ddot{\phi} - \nabla^2\phi + m^2\phi = (\Box + m^2)\phi = 0 \ . \tag{12}$$

This is equivalent to Eq. (6).

For completeness, it is worth mentioning that the second Hamilton's equation simply gives $\pi = \dot{\phi}$, which is not a new information.

1.1.3 Symmetry and conserved current

We focus on internal symmetries of the Lagrangian. They correspond to transformations to the field ϕ that leaves $\mathcal{L}$ invariant, assuming that ϕ already satisfies the equation of motion.

Note this is a stronger requirement than having S invariant. More explicitly, with an infinitesimal variation, $\phi \to \phi + \delta\phi$, we have

$$\begin{aligned} \delta\mathcal{L} &= \mathcal{L}\left(\phi + \delta\phi, \partial_\mu(\phi + \delta\phi)\right) - \mathcal{L}(\phi, \partial_\mu\phi) \\ &= \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right]\delta\phi(x) + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) \ . \end{aligned} \tag{13}$$

Now the first term vanishes because it is proportional to the Euler-Lagrange equation. For this to be a symmetry transformation that leaves $\mathcal{L}$ invariant, we need

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) = 0 \ . \tag{14}$$

It is useful to call $\delta\phi \equiv \varepsilon F(\phi)$ by factorizing out the infinitesimal parameter ε. With this, we can define the current related to the above symmetry transformation

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} F(\phi) \ . \tag{15}$$

Eq. (14) tells that J^μ is a conserved current (Nöther's current)

$$\partial_\mu J^\mu = 0 \ . \tag{16}$$

The corresponding conserved charge is

$$Q = \int d^3x J^0(x) \ , \tag{17}$$

where the integral goes over the whole space volume. It is conserved because $\dot{Q} = \int d^3x \nabla \cdot \vec{J} = \oint d\vec{\mathcal{S}} \cdot \vec{J}$. Assuming the field value vanishes quickly at space infinity (thus $\vec{J}$ also vanishes there), we get $\dot{Q} = 0$.

For every symmetry, there is a conserved current. The number of symmetries is counted by the number of independent infinitesimal parameters you can find.

As a simple example, we consider the theory of a free complex scalar field $\Phi(x)$. The Lagrangian is

$$\mathcal{L} = \partial_\mu \Phi(x) \partial^\mu \Phi(x)^* - m^2 \Phi(x) \Phi(x)^* \ . \tag{18}$$

By treating Φ and Φ^* as independent degrees of freedom, we can derive the same Klein-Gordon equation for Φ.

This Lagrangian has a continuous internal symmetry. It is invariant under

$$\Phi \to \Phi e^{i\varepsilon} \ , \quad \Phi^* \to \Phi^* e^{-i\varepsilon} \ . \tag{19}$$

The corresponding infinitesimal transformation is $\delta\Phi = i\varepsilon\Phi$, $\delta\Phi^* = -i\varepsilon\Phi^*$. Applying Eq. (15), and again treating Φ, Φ^* as two independent fields, we obtain the Nöther's current

$$J^\mu = -i(\Phi \partial^\mu \Phi^* - \Phi^* \partial^\mu \Phi) \ . \tag{20}$$

It is straightforward to check this current is conserved. You just need to apply the equations of motion for Φ and Φ^*.

1.2 Quantization of free fields

1.2.1 Real scalar

As already mentioned, the Lagrangian for a free real scalar field takes the form

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 \ . \tag{21}$$

A quantized scalar field is an operator which can create or annihilate single particle state at spacetime position x.

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} \left(a_{\vec{p}} e^{-ip\cdot x} + a^\dagger_{\vec{p}} e^{ip\cdot x} \right) \ , \tag{22}$$

where $E_{\vec{p}} = \sqrt{|\vec{p}|^2 + m^2}$, $p \cdot x = E_{\vec{p}} t - \vec{p} \cdot \vec{x}$, and the creation/annihilation operators satisfy the commutation relation,

$$\left[a_{\vec{p}}, a^\dagger_{\vec{p}'}\right] = 2E_{\vec{p}}(2\pi)^3 \delta^3(\vec{p} - \vec{p}') \ . \tag{23}$$

This ensures the desired commutation relation for ϕ field and its canonical momentum $\pi = \dot{\phi}$,

$$[\phi(x), \pi(y)]_{x^0=y^0} = i\delta^3(\vec{x} - \vec{y}) \ . \tag{24}$$

1.2.2 Complex scalar

A complex scalar field is can be written as a linear combination of two real scalar fields with equal mass

$$\Phi(x) = \frac{1}{\sqrt{2}} \left(\phi_1(x) + i\phi_2(x)\right) \ . \tag{25}$$

The Lagrangian for Φ is then

$$\mathcal{L}(\Phi) = \mathcal{L}(\phi_1) + \mathcal{L}(\phi_2) = \partial_\mu \Phi \partial^\mu \Phi^\dagger - m^2 \Phi \Phi^\dagger \ . \tag{26}$$

From now on we use $\dagger$ instead of $*$ for hermitian conjugate because quantum fields are operators.

Using the expansion Eq. (22), we can write $\Phi(x)$ as

$$\Phi(x) = \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} \left(a_{\vec{p}} e^{-ip\cdot x} + b^\dagger_{\vec{p}} e^{ip\cdot x}\right) \ , \tag{27}$$

where

$$\begin{aligned} a_{\vec{p}} &= \frac{1}{\sqrt{2}} \left(a_{1\vec{p}} + i a_{2\vec{p}}\right) \ , \quad a^\dagger_{\vec{p}} = \frac{1}{\sqrt{2}} \left(a^\dagger_{1\vec{p}} - i a^\dagger_{2\vec{p}}\right) \ , \\ b_{\vec{p}} &= \frac{1}{\sqrt{2}} \left(a_{1\vec{p}} - i a_{2\vec{p}}\right) \ , \quad b^\dagger_{\vec{p}} = \frac{1}{\sqrt{2}} \left(a^\dagger_{1\vec{p}} + i a^\dagger_{2\vec{p}}\right) \ . \end{aligned} \tag{28}$$

The operators $a, a^\dagger$ and $b, b^\dagger$ are used to annihilate or create a particle or the antiparticle, and they satisfies the same commutation relations as Eq. (23), respectively.

1.2.3 Antiparticle and causality

Causality states that for two points that are space-like separated, two operators (corresponding to experimental actions) must commutate with each other,

$$\left[\hat{O}_1(x), \hat{O}_2(y)\right] = 0 \ , \quad \text{if } (x-y)^2 < 0 \ . \tag{29}$$

In QFT, the operators are made of fields. To illustrate the point, we consider the simplest case where $\hat{O}_1(x) = \Phi(x)$, $\hat{O}_2(y) = \Phi^\dagger(y)$, and evaluate the matrix element of their commutator between vacuum states.

We first act $\Phi(x)$ and $\Phi^\dagger(y)$ on the vacuum ket state $|0\rangle$. Using Eq. (27), we get

$$\begin{aligned}\Phi(x)|0\rangle &= \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} e^{ip\cdot x} b^\dagger_{\vec{p}}|0\rangle = \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} e^{ip\cdot x}|b,\vec{p}\rangle \ , \\ \Phi(y)^\dagger|0\rangle &= \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} e^{ip\cdot y} a^\dagger_{\vec{p}}|0\rangle = \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} e^{ip\cdot y}|a,\vec{p}\rangle \ .\end{aligned} \tag{30}$$

Physically, we can interpret them as creating an antiparticle (particle) at spacetime point x (y).

Next, we consider the corresponding bra states

$$\begin{aligned}\langle 0|\Phi(y)^\dagger &= \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} e^{-ip\cdot y}\langle 0|b_{\vec{p}} = \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} e^{-ip\cdot y}\langle b,\vec{p}| \ , \\ \langle 0|\Phi(x) &= \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} e^{-ip\cdot x}\langle 0|a_{\vec{p}} = \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} e^{-ip\cdot x}\langle a,\vec{p}| \ .\end{aligned} \tag{31}$$

which corresponds to a single antiparticle (particle) state at spacetime point y (x).

Putting them together, we obtain

$$\begin{aligned}\langle 0|\Phi(y)^\dagger\Phi(x)|0\rangle &= \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} e^{ip\cdot(x-y)} \ , \\ \langle 0|\Phi(x)\Phi(y)^\dagger|0\rangle &= \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} e^{-ip\cdot(x-y)} \ .\end{aligned} \tag{32}$$

Diagrammatically, they correspond to an antiparticle (particle) travels from x to y (y to x), as shown below.

The matrix element of the commutator is then

$$\langle 0|[\Phi(x),\Phi(y)^\dagger]|0\rangle = \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} \left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right) \ . \tag{33}$$

This is a Lorentz invariant quantity, thus we can evaluate it in any reference frame and get the same result.

For space-like separated two points, it is always possible to find a reference frame such that $x^0 = y^0$. In this frame, we have

$$\langle 0|[\Phi(x),\Phi(y)^\dagger]|0\rangle = \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} \left(e^{i\vec{p}\cdot\vec{r}} - e^{-i\vec{p}\cdot\vec{r}}\right) \ , \tag{34}$$

where $\vec{r} = \vec{x} - \vec{y}$. It is then straightforward to show this equals 0, by defining $\vec{p} \to -\vec{p}$ for one of the terms.

The lesson we learn here is in QDT, the propagation of a particle is equivalent to the propagation of its antiparticle in the opposite spacetime direction. Had we quantized the field Φ without involving the antiparticle operators, such cancelation would not occur. The presence of antiparticle guarantees causality to work in QFT. Every particle has an antiparticle, with equal mass.

1.2.4 Dirac fermion

The Lagrangian for a free Dirac fermion field takes the form

$$\mathcal{L}(\psi, \partial_\mu \psi, \bar{\psi}, \partial_\mu \bar{\psi}) = \bar{\psi} i \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi \ . \tag{35}$$

Because there is no $\partial_\mu \bar{\psi}$ term, the equation of motion is most conveniently derived by taking derivative with respect to $\bar{\psi}$, which leads to

$$(i \gamma^\mu \partial_\mu - m) \psi = 0 \ . \tag{36}$$

This is the Dirac equation. The field ψ is a four component spinor.

In this note, we will use the following convention for γ matrices,

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \ , \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \ , \quad \gamma_5 = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \ , \tag{37}$$

where $\mathbb{1}$ is a 2×2 unit matrix, and σ_i are the Pauli matrices. A useful anti-commutation relation: $\gamma_5 \gamma^\mu = -\gamma^\mu \gamma_5$. In three space dimensions, the γ_5 matrix is block diagonal, which allows us to define the projection operators

$$\mathbb{P}_L = \frac{1 - \gamma_5}{2} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix} \ , \quad \mathbb{P}_R = \frac{1 + \gamma_5}{2} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1} \end{pmatrix} \ . \tag{38}$$

In the Fourier space, the field operator can be expanded as

$$\psi(x) = \sum_s \int \frac{d^3 p}{(2\pi)^3 2 E_{\vec{p}}} \left(u(\vec{p}, s) a_{\vec{p},s} e^{-ip \cdot x} + v(\vec{p}, s) b^\dagger_{\vec{p},s} e^{ip \cdot x} \right) \ , \tag{39}$$

where the index s goes over two discrete spin orientations of the fermion. Here the particle/antiparticle creation and annihilation operators satisfy the anti-commutation relation,

$$\left\{ a_{\vec{p},s}, a^\dagger_{\vec{p}',s'} \right\} = \left\{ b_{\vec{p},s}, b^\dagger_{\vec{p}',s'} \right\} = 2 E_{\vec{p}} (2\pi)^3 \delta^3 (\vec{p} - \vec{p}') \delta_{ss'} \ . \tag{40}$$

Plugging Eq. (39) into the Dirac equation, we get

$$\sum_s \int \frac{d^3 p}{(2\pi)^3 2 E_{\vec{p}}} \left[(\not{p} - m) u(\vec{p}, s) a_{\vec{p},s} e^{-ip \cdot x} + (\not{p} + m) v(\vec{p}, s) b^\dagger_{\vec{p},s} e^{ip \cdot x} \right] = 0 \ . \tag{41}$$

For it to hold for any value of x, we get two equations for the spinors u and v

$$(\not{p} - m)u(\vec{p}, s) = 0 \ , \quad (\not{p} + m)v(\vec{p}, s) = 0 \ . \tag{42}$$

With the above explicit forms of γ matrices, we can solve for u and v.

Here we work out the u solution explicitly, by first defining

$$u(\vec{p}, s) = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \ . \tag{43}$$

The first equation of Eq. (42) reads

$$\begin{pmatrix} -m\mathbb{1} & E\mathbb{1} - \vec{p}\cdot\vec{\sigma} \\ E\mathbb{1} + \vec{p}\cdot\vec{\sigma} & -m\mathbb{1} \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = 0 \ , \tag{44}$$

or

$$\begin{cases} -m\xi + E\eta - \vec{p}\cdot\vec{\sigma}\eta = 0 \\ E\xi + \vec{p}\cdot\vec{\sigma}\xi - m\eta = 0 \end{cases} \tag{45}$$

Adding and subtracting the two equations in Eq. (45) lead to

$$\begin{cases} (E - m)(\xi + \eta) = \vec{p}\cdot\vec{\sigma}(\xi - \eta) \\ (E + m)(\xi - \eta) = \vec{p}\cdot\vec{\sigma}(\xi + \eta) \end{cases} \tag{46}$$

Next, define

$$\xi + \eta = 2A\chi_s \ , \tag{47}$$

where A is an overall normalization factor. In general, the two component spinor χ_s is a linear combination of two unit vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \ , \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ , \tag{48}$$

and satisfies $\chi_s^\dagger \chi_s = 1$. Using the second equation of Eq. (46), we can find

$$\xi - \eta = 2A\frac{\vec{p}\cdot\vec{\sigma}}{E + m}\chi_s \ . \tag{49}$$

Eqs. (47) and (49) allows us to solve ξ and η, and in turn the four spinor $u(\vec{p}, s)$,

$$u(\vec{p}, s) = \sqrt{\frac{E + m}{2}} \begin{pmatrix} \left(\mathbb{1} + \frac{\vec{p}\cdot\vec{\sigma}}{E+m}\right)\chi_s \\ \left(\mathbb{1} - \frac{\vec{p}\cdot\vec{\sigma}}{E+m}\right)\chi_s \end{pmatrix} \ . \tag{50}$$

Similarly, we can also find

$$v(\vec{p}, s) = \sqrt{\frac{E + m}{2}} \begin{pmatrix} \left(\mathbb{1} + \frac{\vec{p}\cdot\vec{\sigma}}{E+m}\right)\chi_s \\ -\left(\mathbb{1} - \frac{\vec{p}\cdot\vec{\sigma}}{E+m}\right)\chi_s \end{pmatrix} \ . \tag{51}$$

The normalization factor A is found from the condition $\bar{u}(\vec{p},s)u(\vec{p},s') = \bar{v}(\vec{p},s)v(\vec{p},s') = 2m\delta_{ss'}$. This guarantees the desired anti-commutation relation between ψ and its canonical momentum $\pi = \bar{\psi}i\gamma^0 = i\psi^\dagger$,

$$\left\{\psi(x), i\psi^\dagger(y)\right\}_{x^0=y^0} = i\delta^3(\vec{x}-\vec{y}) \ . \tag{52}$$

It is straightforward to verify two very useful identities

$$\begin{aligned} \sum_s u(\vec{p},s)\bar{u}(\vec{p},s) &= \not{p} + m \ , \\ \sum_s v(\vec{p},s)\bar{v}(\vec{p},s) &= \not{p} - m \ . \end{aligned} \tag{53}$$

1.2.5 Helicity eigenstates

Define the helicity operator (here we include a factor of 2 for convenience)

$$\hat{h} = \frac{2\vec{p}\cdot\vec{\Sigma}}{|\vec{p}|} = \frac{1}{|\vec{p}|}\begin{pmatrix} \vec{p}\cdot\vec{\sigma} & 0 \\ 0 & \vec{p}\cdot\vec{\sigma} \end{pmatrix} \tag{54}$$

The eigenvalues of this matrix is $\lambda = \pm 1$. One can verify that the helicity operator commutates with the quantum mechanics Hamiltonian. Thus helicity is a conserved quantity. It is useful to find the corresponding helicity eigenstates satisfying

$$\hat{h}u(\vec{p},s) = \pm u(\vec{p},s) \ , \quad \hat{h}v(\vec{p},s) = \pm v(\vec{p},s) \ . \tag{55}$$

Plugging in the above matrix forms for u and v, we find that helicity eigenstates require

$$\frac{\vec{p}\cdot\vec{\sigma}}{|\vec{p}|}\chi_s = \pm\chi_s \ . \tag{56}$$

This is a simple linear algebra problem. The answers are

$$\begin{aligned} \chi_s &= \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}e^{i\phi} \end{pmatrix} \equiv \chi_+ \ , \quad (\text{for } \lambda = +1) \ , \\ \chi_s &= \begin{pmatrix} \sin\frac{\theta}{2}e^{-i\phi} \\ -\cos\frac{\theta}{2} \end{pmatrix} \equiv \chi_- \ , \quad (\text{for } \lambda = -1) \ . \end{aligned} \tag{57}$$

The u, v spinors constructed using $\chi_\pm$ are helicity eigenstates. They are the building blocks of helicity amplitudes in QFT.

An interesting limit is the case of massless or ultra-relativistic fermions, where $m \ll E \simeq |\vec{p}|$. The helicity eigenstate u, v spinors take the following asymptotic forms,

$$\begin{aligned} u(\vec{p},+) &= \sqrt{2E}\begin{pmatrix} \chi_+ \\ 0 \end{pmatrix} \ , \quad u(\vec{p},-) = \sqrt{2E}\begin{pmatrix} 0 \\ \chi_- \end{pmatrix} \ , \\ v(\vec{p},+) &= \sqrt{2E}\begin{pmatrix} \chi_+ \\ 0 \end{pmatrix} \ , \quad v(\vec{p},-) = \sqrt{2E}\begin{pmatrix} 0 \\ -\chi_- \end{pmatrix} \ . \end{aligned} \tag{58}$$

In this limit, they are also eigenstates of the chirality operators $\mathbb{P}_{L,R}$.

1.2.6 Majorana fermion

A Majorana fermion is a special case which is its own antiparticle. We first introduce the charge-conjugation transformation to a fermion field,

$$\psi^c(x) \equiv C\bar{\psi}(x)^T \ , \quad C = -i\gamma^2\gamma^0 \ . \tag{59}$$

$\psi^c(x)$ is the antiparticle field. It could be slightly simplified and written as $i\gamma^2\psi(x)^*$. The use of $*$ is a bit sloppy here. One should keep in mind that it works as a $\dagger$ when acting on the creation or annihilation operator.

With the explicit form of γ^2 defined in Eq. (37), if ψ is written as

$$\psi = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \ , \tag{60}$$

the corresponding charge conjugation field is then

$$\psi^c = \begin{pmatrix} 0 & -i\sigma_2 \\ i\sigma_2 & 0 \end{pmatrix} \begin{pmatrix} \xi^* \\ \eta^* \end{pmatrix} = \begin{pmatrix} -i\sigma_2\eta^* \\ i\sigma_2\xi^* \end{pmatrix} \ , \tag{61}$$

Now we are ready to quantitatively define that a Majorana fermion ($\psi = \psi^c$) satisfies the condition

$$\eta = i\sigma_2\xi^* \ . \tag{62}$$

We can still expand the Majorana fermion field in the Fourier space using Eq. (39), but $\psi = \psi^c$ implies the following relations,

$$\begin{aligned} a_{\vec{p},s} = b_{\vec{p},s} \ , \quad a^\dagger_{\vec{p},s} = b^\dagger_{\vec{p},s} \ , \\ u(\vec{p},s) = -i\gamma^2 v(\vec{p},s)^* \ , \quad v(\vec{p},s) = -i\gamma^2 u(\vec{p},s)^* \ . \end{aligned} \tag{63}$$

If u takes the forms of helicity eigenstates,

$$u(\vec{p},\pm) = \sqrt{\frac{E+m}{2}} \begin{pmatrix} \left(\mathbb{1} + \frac{\vec{p}\cdot\vec{\sigma}}{E+m}\right)\chi_\pm \\ \left(\mathbb{1} - \frac{\vec{p}\cdot\vec{\sigma}}{E+m}\right)\chi_\pm \end{pmatrix} \ , \tag{64}$$

then the corresponding v must be

$$v(\vec{p},\pm) = \sqrt{\frac{E+m}{2}} \begin{pmatrix} -i\sigma_2\left(\mathbb{1} - \frac{\vec{p}\cdot\vec{\sigma}^*}{E+m}\right)\chi^*_\pm \\ i\sigma_2\left(\mathbb{1} + \frac{\vec{p}\cdot\vec{\sigma}^*}{E+m}\right)\chi^*_\pm \end{pmatrix} = \pm\sqrt{\frac{E+m}{2}} \begin{pmatrix} \left(\mathbb{1} + \frac{\vec{p}\cdot\vec{\sigma}}{E+m}\right)\chi_\mp \\ \left(\mathbb{1} - \frac{\vec{p}\cdot\vec{\sigma}}{E+m}\right)\chi_\mp \end{pmatrix} \ . \tag{65}$$

In the second step, we have used the identity $\sigma_2\vec{\sigma}^* = -\vec{\sigma}\sigma_2$, and the relation $i\sigma_2\chi^*_\pm = \pm\chi_\mp$ (see Eqs. (57)).

A vector current made of Majorana fermion must vanish,

$$\bar{\psi}\gamma^\mu\psi = 0 \ , \quad \text{if } \psi = \psi^c \ . \tag{66}$$

Physically, this means we cannot tell the net particle number, usually defined as number of particles minus number of antiparticles in the system, because

we cannot distinguish them. On the other hand, the axial current counterpart $\bar{\psi}\gamma^\mu\gamma^5\psi$ is nonzero.

The Lagrangian for a Majorana fermion takes the form

$$\mathcal{L} = \frac{1}{2}\left(\bar{\psi}i\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi\right) \ . \tag{67}$$

The factor $\frac{1}{2}$ has the same origin as the Lagrangian for a real scalar field.

1.2.7 Massive vector boson

A free massive vector boson A_μ is described by the Proca Lagrangian,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2A_\mu A^\mu = -\frac{1}{2}\partial_\mu A_\nu\partial^\mu A^\nu + \frac{1}{2}\partial_\mu A_\nu\partial^\nu A^\mu + \frac{1}{2}m^2A_\mu A^\mu \ , \tag{68}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor. Here A_μ is a real field.

The Euler-Lagrange equation takes the form

$$\frac{\partial\mathcal{L}}{\partial A_\mu} = \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial\partial_\nu A_\mu}\right) \ , \quad \Rightarrow \quad (\Box + m^2)A_\mu - \partial_\mu\partial_\nu A^\nu = 0 \ . \tag{69}$$

Apply another derivative ∂^μ to the above equation, we get

$$m^2\partial^\mu A_\mu = 0 \ . \tag{70}$$

Therefore, if $m \neq 0$, we must have $\partial^\mu A_\mu = 0$. This is a constraint. It reduces the degrees of freedom of A_μ to three, corresponding to three possible polarizations (see below).

With the above constraint, the equation of motion can be simplified to

$$(\Box + m^2)A_\mu = 0 \ . \tag{71}$$

Quantizing the massive vector field leads to the following form in the Fourier space,

$$A_\mu(x) = \sum_{\lambda=0,\pm1}\int\frac{d^3p}{(2\pi)^3 2E_{\vec{p}}}\left(\varepsilon_\mu(\vec{p},\lambda)a_{\vec{p},\lambda}e^{-ip\cdot x} + \varepsilon^*_\mu(\vec{p},\lambda)a^\dagger_{\vec{p},\lambda}e^{ip\cdot x}\right) \ , \tag{72}$$

where the creating and annihilation operators satisfy the regular commutation relation

$$\left[a_{\vec{p},\lambda}, a^\dagger_{\vec{p}',\lambda'}\right] = 2E_{\vec{p}}(2\pi)^3\delta^3(\vec{p}-\vec{p}')\delta_{\lambda\lambda'} \ . \tag{73}$$

There are three types of polarization vectors for a massive vector boson, corresponding to $\lambda = 0, \pm1$. They satisfy

$$\varepsilon_\mu(\vec{p},\lambda)\varepsilon^{\mu*}(\vec{p},\lambda') = -\delta_{\lambda\lambda'} \ , \quad p^\mu\varepsilon_\mu(\vec{p},\lambda) = 0 \ . \tag{74}$$

In a special case, where the particle travels along the $\hat{z}$ axis, $p^\mu = (E, 0, 0, p)$, we can write explicit forms of ε_μ

$$\begin{aligned} \varepsilon^\mu(\vec{p}, \pm 1) &= \frac{1}{\sqrt{2}} (0, 1, \pm i, 0) \ , \quad \text{(transverse modes)} \ , \\ \varepsilon^\mu(\vec{p}, 0) &= \frac{1}{m} (p, 0, 0, E) \ , \quad \text{(longitudinal mode)} \ . \end{aligned} \tag{75}$$

In calculations, we often need to sum over polarizations of a vector boson, in the form

$$\sum_{\lambda=0,\pm 1} \varepsilon_\mu(\vec{p}, \lambda) \varepsilon_\nu(\vec{p}, \lambda) = -g_{\mu\nu} + \frac{p_\mu p_\nu}{m^2} \ . \tag{76}$$

1.2.8 Massless vector boson

In this note, we take a simplified way to discuss the quantization of massless vector boson, without resorting to path integrals. The trick here is to treat a vector boson that is massless as one with a tiny but nonzero mass. Physically, this makes sense because our experiments never tell that the photon is strictly massless, but can only set more and more stringent upper bound on the photon mass. Mathematically, this approach does not looks great because the vector boson mass breaks the apparent gauge invariance (that is what we need massless vector bosons for, in many cases). Nonetheless, we proceed with the following argument.

Without loss of generality, we consider the vector boson travels along the $\hat{z}$ axis, with $p^\mu = (E, 0, 0, p)$. The longitudinal polarization vector is $\varepsilon^\mu(\vec{p}, 0) = (p, 0, 0, E)/m$. In the limit $m \to 0$, we have $E \to p$. As a result, the two vectors p_μ and $\varepsilon_\mu(\vec{p}, 0)$ are in parallel. Do not worry about the denominator.

In this case, we can remove the longitudinal component perform a transformation to the vector field A_μ,

$$A_\mu(x) \to A_\mu(x) + \partial_\mu \omega(x) \ . \tag{77}$$

In the momentum space, this corresponds to

$$\tilde{A}_\mu(p) \to \tilde{A}_\mu(p) + p_\mu \tilde{\omega}(p) \ . \tag{78}$$

Because $p_\mu \propto \varepsilon_\mu(\vec{p}, 0)$ in the $m \to 0$ limit, we can always remove all the parts in A_μ that are proportional to $\varepsilon_\mu(\vec{p}, 0)$.

An important observation here is that we are always allowed to do so because without a mass term, the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \ , \tag{79}$$

is invariant under the above transformation, Eq. (77). It means we can always perform the proper field transformation and work in a "basis" where the vector boson only has transverse degrees of freedom. It indicates the longitudinal

polarization component of a massless vector boson is not physical. In such a basis, the massless vector field expands as

$$A_\mu(x) = \sum_{\lambda=\pm1} \int \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} \left(\varepsilon_\mu(\vec{p},\lambda) a_{\vec{p},\lambda} e^{-ip\cdot x} + \varepsilon^*_\mu(\vec{p},\lambda) a^\dagger_{\vec{p},\lambda} e^{ip\cdot x} \right) . \tag{80}$$

As we learned from classical E&M, the sum over transverse polarizations of photon field satisfy

$$\sum_{\lambda=0,\pm1} \varepsilon^i(\vec{p},\lambda)\varepsilon^j(\vec{p},\lambda) = \delta_{ij} + \frac{p^i p^j}{|\vec{p}|^2} . \tag{81}$$

In the notation of four component vectors, it becomes

$$\sum_{\lambda=\pm1} \varepsilon^\mu(\vec{p},\lambda)\varepsilon^\nu(\vec{p},\lambda) = -g^{\mu\nu} + \frac{p^\mu p_\nu}{|\vec{p}|^2} . \tag{82}$$

The last term is ugly. Fortunately, we can drop such a term in most calculations because a massless vector boson must couple to a conserved current. This corresponds to the Ward identity (see Sec. 2.4.1 for a concrete example). At the Lagrangian level, the interaction term reads

$$\mathcal{L}_{\rm int} \sim A_\mu J^\mu , \tag{83}$$

where J^μ satisfies $\partial_\mu J^\mu = 0$. At quantum level, the vector boson cannot stay massless if the current is not conserved. When computing amplitude squares, after the polarization sum, A_μ (or ε_μ) would turn into a p_μ if it contributes through the second term of Eq. (82). Current conservation implies this contribution is zero. Thus, effectively, we can always work with

$$\sum_{\lambda=\pm1} \varepsilon^\mu(\vec{p},\lambda)\varepsilon^\nu(\vec{p},\lambda) = -g^{\mu\nu} , \tag{84}$$

for a massless vector boson.

1.3 Global and gauge symmetries

We focus on internal continuous symmetries.

1.3.1 $U(1)$ global symmetry

A $U(1)$ global symmetry usually refers to a constant phase redefinition of a field. Take the complex scalar Lagrangian, Eq. (26), as the first example. It is invariant under the field transformations

$$\Phi(x) \to \Phi(x) e^{i\theta} , \quad \Phi(x)^\dagger \to \Phi(x)^\dagger e^{-i\theta} , \tag{85}$$

where θ does not depend on x thus derivatives on the field does not act on θ. This is clearly true because every term contains a $\Phi(x)$ and a $\Phi(x)^\dagger$, allowing the

phase factors to cancel after the above transformations. Such a transformation thus corresponds to a symmetry of the Lagrangian.

The infinitesimal transformations are $\Phi \to \Phi + i\varepsilon\Phi$, $\Phi^\dagger \to \Phi^* - i\varepsilon\Phi^\dagger$. Using Eq. (15), the conserved current for the symmetry is

$$J^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)} i\Phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^\dagger)}(-i\Phi^\dagger) = i(\Phi\partial^\mu\Phi^\dagger - \Phi^\dagger\partial^\mu\Phi) \ . \tag{86}$$

Using equation of motion, it is straightforward to show $\partial_\mu J^\mu = 0$. The corresponding conserved charge is

$$Q = \int d^3x J^0(x) = i\int d^3x(\Phi\dot{\Phi}^\dagger - \Phi^\dagger\dot{\Phi}) \ . \tag{87}$$

As the second example, we consider the Dirac fermion Lagrangian, Eq. (35). It is invariant under the field transformations

$$\begin{aligned} &\psi(x) \to \psi(x)e^{i\theta} \ , \quad \bar{\psi}(x) \to \bar{\psi}(x)e^{-i\theta} \ , \\ &\psi(x) \to \psi(x) + i\varepsilon\psi \ , \quad \bar{\psi}(x) \to \bar{\psi}(x) - i\varepsilon\bar{\psi} \ , \quad \text{(infinitesimal)}. \end{aligned} \tag{88}$$

The conserved current is

$$J^\mu = i\bar{\psi}\gamma^\mu(i\psi) = -\bar{\psi}\gamma^\mu\psi \ . \tag{89}$$

This is a vector current, which always stays conserved after including quantum corrections. The conserved charge is

$$Q = -\int d^3x\bar{\psi}\gamma^0\psi = -\int d^3x\psi^\dagger\psi(x) \ . \tag{90}$$

$\psi^\dagger\psi$ corresponds to a number density. The overall sign is unimportant.

1.3.2 Symmetries of QED

Next, we consider an interacting theory – the QED. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}\gamma^\mu(i\partial_\mu - eA_\mu)\psi - m\bar{\psi}\psi \ . \tag{91}$$

Like before, the fermion part of the theory has a $U(1)$ global symmetry, under the transformation Eq. (88).

Thanks to the presence of more fields, the theory of QED accommodates a different type of symmetry, where the transformations act jointly on the fermion and vector boson fields,

$$\begin{aligned} \psi(x) &\to \psi(x)e^{-i\omega(x)} \ , \\ \bar{\psi}(x) &\to \bar{\psi}(x)e^{i\omega(x)} \ , \\ A_\mu(x) &\to A_\mu(x) + \partial_\mu\omega(x) \ . \end{aligned} \tag{92}$$

Here, the fermion field changes by a phase factor $e^{-i\omega(x)}$ where $\omega(x)$ is spacetime dependent. It corresponds to a local transformation. It is straightforward to check that

$$\begin{aligned} \bar{\psi} i\partial\!\!\!/ \psi &\to \bar{\psi} i\partial\!\!\!/ \psi + e\bar{\psi}\gamma^\mu \psi \partial_\mu \omega \ , \\ -e\bar{\psi}\gamma^\mu \psi A_\mu &\to -e\bar{\psi}\gamma^\mu \psi A_\mu - e\bar{\psi}\gamma^\mu \partial_\mu \omega \ , \end{aligned} \tag{93}$$

and the fermion mass term is simply invariant. Adding all the pieces together, the whole Lagrangian remains invariant. This is a $U(1)$ gauge symmetry.

QED is a gauge theory. The photon is the gauge boson. It couples to a conserved vector current $\bar{\psi}\gamma^\mu\psi$. e is the gauge coupling. The minus sign shows up in front of e because the electron has a negative electric charge.

1.3.3 $SU(2)$ and $SU(3)$ groups

Next we discuss symmetries larger than $U(1)$. This usually requires more fields. We start from the global symmetries, by consider a free theory of two degenerate fermion fields,

$$\begin{aligned} \mathcal{L} &= \bar{\psi}_1 i\partial\!\!\!/ \psi_1 - m\bar{\psi}_1\psi_1 + \bar{\psi}_2 i\partial\!\!\!/ \psi_2 - m\bar{\psi}_2\psi_2 \\ &= \begin{pmatrix} \bar{\psi}_1 & \bar{\psi}_2 \end{pmatrix} i\partial\!\!\!/ \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} - \begin{pmatrix} \bar{\psi}_1 & \bar{\psi}_2 \end{pmatrix} \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \ . \end{aligned} \tag{94}$$

In the second step we rewrite the Lagrangian using matrices. It is useful to introduce notations for the fermion doublet and mass matrix

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \ , \quad \bar{\Psi} = \begin{pmatrix} \bar{\psi}_1 & \bar{\psi}_2 \end{pmatrix} \ , \quad \mathbb{M} = \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \ . \tag{95}$$

Such a Lagrangian is invariant under the global symmetry transformations

$$\Psi \to e^{\frac{i}{2}(\alpha_0 + \vec{\alpha}\cdot\vec{\sigma})}\Psi \ , \quad \bar{\Psi} \to \bar{\Psi} e^{-\frac{i}{2}(\alpha_0 + \vec{\alpha}\cdot\vec{\sigma})} \ , \tag{96}$$

where $\alpha_{0,1,2,3}$ are real parameters. They are not x dependent. In particular, the mass term is invariant because $\mathbb{M}$ is proportional to a 2×2 unit matrix.

The symmetry supported by parameter α_0 corresponds to a common phase rotation of both $\psi_{1,2}$ fields, which is a $U(1)$ symmetry.

The symmetries supported by the three parameters $\vec{\alpha}$ correspond to an $SU(2)$ group. Under the $SU(2)$ rotation, $\psi_{1,2}$ can transform into each other. The $SU(2)$ symmetry has three generators, which are related to the Pauli matrix,

$$T^a = \frac{1}{2}\sigma_a \ , \quad (a = 1,2,3) \ . \tag{97}$$

The total global symmetry of the theory is $SU(2)\times U(1)$.

We can repeat the above discussion for a theory with three degenerate fermions. In that case, the total global symmetry would be $SU(3)\times U(1)$. The generators of the $SU(3)$ group are related to the Gell-Mann matrices

$$T^a = \frac{1}{2}\lambda_a \ , \quad (a = 1,\ldots,8) \ , \tag{98}$$

where

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ , \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ , \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ ,$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \ , \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \ , \tag{99}$$
$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \ , \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \ , \quad \lambda_8 = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \ .$$

The above model building procedure can be generalized to N copies of fermions, where we will have an $SU(N) \times U(1)$ global symmetry. An $SU(N)$ group has $N^2 - 1$ generators. All of them are Hermitian matrices. They satisfy the Lie algebra

$$\left[T^a, T^b\right] = i f^{abc} T^c \ , \tag{100}$$

where f^{abc} is the antisymmetric group structure constant. For the case of $SU(2)$, $f^{abc} = \varepsilon^{abc}$. In general, the structure constant satisfies the Jacobean identity.

$$f^{acb} f^{bde} + f^{adb} f^{bec} + f^{aeb} f^{bcd} = 0 \ . \tag{101}$$

Beyond $SU(2)$, the anti-commutator of two different generators do not always vanish,

$$\left\{T^a, T^b\right\} = \frac{1}{N} \delta^{ab} + d^{abc} T^c \ , \tag{102}$$

where d^{abc} is the symmetric group structure constant and another signature of the $SU(N)$ group, besides f^{abc}.

For fundamental representation of the $SU(N)$ group, the generators satisfy

$$\begin{aligned} &\mathrm{Tr}(T^a) = 0 \ , \quad \mathrm{Tr}(T^a T^b) = \frac{1}{2} \delta_{ab} \ , \\ &\sum_{a=1}^{N^2-1} (T^a)^2 = \frac{N^2 - 1}{2N} \mathbb{1} \ , \\ &\sum_{a=1}^{N^2-1} T^a_{ij} T^a_{kl} = \frac{1}{2} \left(\delta_{il} \delta_{kj} - \frac{1}{N} \delta_{ij} \delta_{kl} \right) \ . \end{aligned} \tag{103}$$

1.3.4 Gauged non-abelian symmetry

Here, we work in the general case of $SU(N)$ gauge symmetry. Introducing N degenerate fermions, the Lagrangian is

$$\mathcal{L} = -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} + \bar{\Psi} \gamma^\mu \left(i \partial_\mu + g A^a_\mu T^a \right) \Psi - m \bar{\Psi} \Psi \ , \tag{104}$$

where Ψ is a column matrix of N fermions, ψ_i, $(i = 1, \ldots, N)$. In the presence of fermions, this is also called the Yang-Mills theory.

The Yang-Mill Lagrangian is invariant under the joint transformations

$$\begin{aligned} \Psi(x) &\to e^{ig\omega^a(x)T^a}\Psi(x) \ , \\ \bar{\Psi}(x) &\to \bar{\Psi}(x)e^{-ig\omega^a(x)T^a} \ , \\ A^a_\mu(x) &\to A^a_\mu(x) + \partial_\mu\omega^a(x) - gf^{abc}\omega^b(x)A^c_\mu(x) \ . \end{aligned} \tag{105}$$

For simplicity, we will just show the above Lagrangian is invariant under infinitesimal transformations, in the limit $\omega^a \ll 1$,

$$e^{\pm ig\omega^a(x)T^a} \simeq 1 \pm ig\omega^a(x)T^a \ . \tag{106}$$

First, under the above transformation, the mass terms is invariant up to first order in ω,

$$\bar{\Psi}\Psi \to \bar{\Psi}(1 - ig\omega^a(x)T^a)(1 + ig\omega^a(x)T^a)\Psi = \bar{\Psi}\Psi + \mathcal{O}(\omega^2) \ . \tag{107}$$

The fermion kinetic term becomes

$$\begin{aligned} \bar{\Psi}i\not{\partial}\Psi &\to \bar{\Psi}(1 - ig\omega^a(x)T^a)i\not{\partial}(1 + ig\omega^a(x)T^a)\Psi \\ &= \bar{\Psi}i\not{\partial}\Psi - g\bar{\Psi}\gamma^\mu T^a\Psi\partial_\mu\omega + \mathcal{O}(\omega^2) \ , \end{aligned} \tag{108}$$

which is not invariant at first order in ω by itself.

We have to resort to the fermion gauge interaction term,

$$\begin{aligned} g\bar{\Psi}\gamma^\mu A^a_\mu T^a\Psi &\to g\bar{\Psi}(1 - ig\omega^bT^b)\gamma^\mu T^a(1 + ig\omega^cT^c)\Psi\left(A^a_\mu + \partial_\mu\omega^a - gf^{abc}\omega^bA^c_\mu\right) \\ &= g\bar{\Psi}\gamma^\mu A^a_\mu T^a\Psi + ig^2\Psi\gamma^\mu[T^a, T^b]\Psi\omega^bA^a_\mu \\ &\qquad + g\bar{\Psi}\gamma^\mu T^a\Psi\left(\partial_\mu\omega^a - gf^{abc}\omega^bA^c_\mu\right) \\ &= g\bar{\Psi}\gamma^\mu A^a_\mu T^a\Psi - g^2f^{abc}\Psi\gamma^\mu T^c\Psi\omega^bA^a_\mu \\ &\qquad + g\bar{\Psi}\gamma^\mu T^a\Psi\partial_\mu\omega^a - g^2f^{abc}\bar{\Psi}\gamma^\mu T^a\Psi\omega^bA^c_\mu \\ &= g\bar{\Psi}\gamma^\mu A^a_\mu T^a\Psi + g\bar{\Psi}\gamma^\mu T^a\Psi\partial_\mu\omega^a + \mathcal{O}(\omega^2) \ . \end{aligned} \tag{109}$$

In the last step, a cancelation between two terms occur due to the total anti-symmetric nature of f^{abc}.

Nicely, the sum of the two terms on the left-hand side of Eqs. (108) and (109) are invariant up to first order in ω.

The last thing to work out is the gauge kinetic term. We need to specify the form of $F^a_{\mu\nu}$,

$$F^a_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu \ . \tag{110}$$

Under the gauge transformation Eq. (105), we have, up to first order in ω,

$$\begin{aligned} \delta(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu) &= -gf^{abc}\left[\partial_\mu(\omega^bA^c_\nu) - \partial_\nu(\omega^bA^c_\mu)\right] \\ &= -gf^{abc}\omega^b(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu) - gf^{abc}\left[(\partial_\mu\omega^b)A^c_\nu) - (\partial_\nu\omega^b)A^c_\mu\right] \ , \end{aligned} \tag{111}$$

and

$$\begin{aligned}\delta(gf^{abc}A^b_\mu A^c_\nu) &= gf^{abc}\left[(\partial_\mu\omega^b)A^c_\nu) - (\partial_\nu\omega^b)A^c_\mu\right] \\ &\quad - g^2 f^{abc}f^{bde}\omega^d A^e_\mu A^c_\nu - g^2 f^{abc}f^{cfg}A^b_\mu\omega^f A^g_\nu \ .\end{aligned} \tag{112}$$

In the last lines, the first term of Eq. (212) cancels the second term of Eq. (111). Adding the rest, we get

$$\begin{aligned}\delta F^a_{\mu\nu} &= -gf^{abc}\omega^b(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu) \\ &\quad - g^2 f^{abc}f^{bde}\omega^d A^e_\mu A^c_\nu - g^2 f^{abc}f^{cfg}A^b_\mu\omega^f A^g_\nu \\ &\quad \text{(use } f^{abc} = -f^{acb} \text{ in the first term of second line} \\ &\quad \text{(rename indices } b\to e, c\to b, f\to d, g\to c \text{ in the last term)} \\ &= -gf^{abc}\omega^b(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu) \\ &\quad + g^2 f^{acb}f^{bde}\omega^d A^e_\mu A^c_\nu - g^2 f^{aeb}f^{bdc}\omega^d A^e_\mu A^c_\nu \\ &\quad \text{(use } f^{bdc} = -f^{bcd} \text{ in the last term), (PS: I hate doing this too..)} \\ &= -gf^{abc}\omega^b(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu) \\ &\quad + g^2(f^{acb}f^{bde} + f^{aeb}f^{bcd})\omega^d A^e_\mu A^c_\nu \ .\end{aligned} \tag{113}$$

Finally, for the last line, we use the Jacobean identity Eq. (101) and find

$$\begin{aligned}\delta F^a_{\mu\nu} &= -gf^{abc}\omega^b(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu) - g^2 f^{adb}f^{bec}\omega^d A^e_\mu A^c_\nu \\ &\quad \text{(rename indices } e\to m, c\to n \text{ and then } b\to c, d\to b \text{ in the last term)} \\ &= -gf^{abc}\omega^b(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu) - g^2 f^{abc}f^{cmn}\omega^b A^m_\mu A^n_\nu \\ &= -gf^{abc}\omega^b\left(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu + gf^{cmn}A^m_\mu A^n_\nu\right) \\ &= -gf^{abc}\omega^b F^c_{\mu\nu} \ .\end{aligned} \tag{114}$$

For non-abelian gauge boson, the field strength tensor is not invariant under the gauge transformation, Eq. (105). However, the square is

$$\begin{aligned}\delta\left(F^a_{\mu\nu}F^{a\mu\nu}\right) &= (-gf^{abc}\omega^b F^c_{\mu\nu})F^{a\mu\nu} + F^a_{\mu\nu}(-gf^{abc}\omega^b F^{c\mu\nu}) \\ &= -gf^{abc}\omega^b(F^c_{\mu\nu}F^{a\mu\nu} + F^a_{\mu\nu}F^{c\mu\nu}) \ ,\end{aligned} \tag{115}$$

which vanishes because the indices a and c are antisymmetric due to f^{abc}, but are symmetric inside the bracket. This proves that the gauge kinetic term is also gauge invariant.

Thus, the Yang-Mills Lagrangian, Eq. (104), is indeed gauge invariant. We have shown this with the infinitesimal field transformations, and find this conclusion holds up to first order in the infinitesimally small parameter ω.

1.3.5 Vectorlike and chiral gauge theories

With the projection operators $\mathbb{P}_{L,R}$, we can separate a Dirac fermion into left- and right-handed components,

$$\psi = \mathbb{P}_L\psi + \mathbb{P}_R\psi = \psi_L + \psi_R \ . \tag{116}$$

In the language of Ψ_L, Ψ_R, the Yang-Mills Lagrangian Eq. (104) can be written as

$$\begin{aligned}\mathcal{L} = &-\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \bar{\Psi}_L\gamma^\mu\left(i\partial_\mu + gA^a_\mu T^a\right)\Psi_L + \bar{\Psi}_R\gamma^\mu\left(i\partial_\mu + gA^a_\mu T^a\right)\Psi_R \\ &- m\bar{\Psi}_L\Psi_R - m\Psi_R\Psi_L \ ,\end{aligned} \tag{117}$$

In this model, under the $SU(N)$ gauge transformation, both Ψ_L, Ψ_R transform the same way as Ψ in Eq (105). A theory that treats left- and right-handed fermions equally is called vectorlike. In such a theory, Dirac masses for fermions are allowed.

The opposite to vectorlike are the chiral theories, where left- and right-handed fermions are in different representations under the gauge group. An extreme example is having only Ψ_L charged under the $SU(N)$ but Ψ_R being gauge singlet. In this case, the gauge invariant Lagrangian reads (for now let's put anomalies aside)

$$\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \bar{\Psi}_L\gamma^\mu\left(i\partial_\mu + gA^a_\mu T^a\right)\Psi_L + \bar{\Psi}_R\gamma^\mu i\partial_\mu\Psi_R \ , \tag{118}$$

where the right-hande fermion is a gauge singlet and only has normal derivative. The fermion mass terms are also forbidden because left and right are not compatible anymore. This Lagrangian is invariant under the gauge transformations similar to Eq (105), but only Ψ_L transforms and Ψ_R stays invariant. We can call the gauge symmetry $SU(N)_L$.

In the Standard Model, we will see that the gauge symmetry for strong interaction $SU(3)_c$ is vectorlike, whereas the symmetries for weak interaction $SU(2)_L \times U(1)_Y$ are chiral. Nature is diverse.

1.4 Spontaneous symmetry breaking

Symmetries are nice. Breaking them can be more fun, if we do so in an organized way (not randomly breaking it). In this subsection, we discuss spontaneous breaking of global and gauge symmetries. Spontaneous symmetry breaking (SSB) is a phenomenon that occurs in nature.

For simplicity, we work with the $U(1)$ symmetry and derive some important concepts that can be generalized to the SSB of higher symmetries.

1.4.1 Linear σ Model and Goldstone theorem

First consider the Lagrangian for a complex scalar field with a non-trivial potential,

$$\mathcal{L} = \partial_\mu\Phi(x)\partial^\mu\Phi(x)^\dagger - V(|\Phi(x)|^2) \ , \tag{119}$$

where $|\Phi(x)|^2 = \Phi(x)\Phi(x)^\dagger$, and the potential takes the following form

$$V(|\Phi(x)|^2) = \lambda|\Phi(x)|^4 + \mu^2|\Phi(x)|^2 \ . \tag{120}$$

Clearly this Lagrangian features a $U(1)$ global symmetry, under which

$$\Phi(x) \to \Phi(x) e^{i\theta} \ , \quad \Phi(x)^\dagger \to \Phi(x)^\dagger e^{-i\theta} \ . \tag{121}$$

The exact form of the potential will dictate the fate such this symmetry.

First, for the potential to be bounded from below so that the theory has a well-defined vacuum energy, we always need $\lambda \geq 0$. In the marginal case $\lambda = 0$, we must have $\mu^2 > 0$. However, if $\lambda > 0$, μ^2 can take either sign.

$\lambda > 0$, $\mu^2 > 0$ correspond to the boring case, where the potential is minimized with

$$\langle \Phi(x) \rangle = 0 \ , \tag{122}$$

where $\langle \dots \rangle$ means taking the vacuum expectation value (VEV) or vacuum condensate of a field. The minimization of the potential must already be done at classical level, before talking about quantum fluctuations.

On the other hand, if $\lambda > 0$, $\mu^2 < 0$, the potential is minimized with a nonzero VEV of the scalar field,

$$\langle |\Phi(x)|^2 \rangle = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2} \ , \tag{123}$$

where v is a real positive parameter with same dimension as mass. For the VEV of $\langle \Phi(x) \rangle$, we have the freedom to choose its phase. In general

$$\langle \Phi(x) \rangle = \frac{v}{\sqrt{2}} e^{i\delta} \ . \tag{124}$$

The freedom of choosing any value δ corresponds to a circle in the complex plane of $\Phi(x)$. Mathematically, it also corresponds to the coset space of the broken symmetry [1]. Hereafter, we will use the freedom and always choose $\delta = 0$, so that

$$\langle \Phi(x) \rangle = \frac{v}{\sqrt{2}} \ . \tag{125}$$

After making this decision, we find such a VEV is no longer invariant under the $U(1)$ transformation, indicating the symmetry is broken. Breaking a symmetry through the VEV of a scalar field is called spontaneous.

It is worth emphasizing the significance of the above minimization of potential. We are not simply minimizing V as a function of some variable. Instead, V is a function of $\Phi(x)$. The latter is a field that fills the whole spacetime. In fact, we are talking about the minimizing procedure for every point in the spacetime. In the above example, they have a common preferred vacuum condensate, Eq. (125). At every point in spacetime, the field value is $v/\sqrt{2}$.

[1] (This footnote could appear vague to readers other than myself.) It is fun to consider a case of bigger symmetry breaking, $SO(3) \to U(1)$. This is the Georgi-Glashow model. The group manifold for $SO(3)$ is the surface of a sphere. The group manifold of $U(1)$ is a great circle that belongs to the surface. The coset space corresponds to choosing the orientation of a great circle on the surface of a sphere. In field theory, at every $\vec{x}$, such an orientation (in group theory space) can be mapped to actually directions in three dimensional space. It leads to a non-trivial topology, and magnetic monopoles.

We proceed with the case of $\lambda > 0$, $\mu^2 < 0$, with SSB. Now we are ready to discuss the quantum field, which describes excitation around the vacuum of the system,

$$\Phi(x) = \langle \Phi(x) \rangle + \text{excitations} = \frac{v + S(x) + iG(x)}{\sqrt{2}} \ , \tag{126}$$

where $S(x)$ and $G(x)$ are excitations along the real and imaginary axis in the complex plane of Φ, respectively. Both are real scalar and they are quantum fields. The way we organize the excitations here is called linear realization. Plugging this expansion back to the Lagrangian, Eq. (119), we get

$$\mathcal{L} = \frac{1}{2}\partial_\mu S \partial^\mu S + \frac{1}{2}\partial_\mu G \partial^\mu G - \lambda v^2 S^2 - \frac{\lambda}{4}(S^2+G^2)^2 - \lambda v S(S^2+G^2) + \text{constants} \ . \tag{127}$$

The first two terms are regular kinetic terms for the real scalar fields. The third term is quadratic in fields and corresponds to a mass term for S, with $m_S^2 = 2\lambda v^2$. The fourth and fifth terms always involve more than two fields in each term and correspond to interaction terms. The constant term does not depend on S or G and corresponds to the vacuum energy.

Interestingly, there is no mass term for G. It is a massless scalar after SSB. This is not an accident, but serves as an example of the Goldstone theorem. Spontaneously breaking a global symmetry yields a Goldstone boson.

The Goldstone theorem can be generalized to the breaking of higher symmetries. Every broken symmetry leads to a massless Goldstone boson. The number of broken symmetry generators is equal to the number of resulting Goldstone bosons.

As another generalization, the Lagrangian Eq. (119) we started with has an exact $U(1)$ global symmetry. In general, global symmetries are allowed to be explicitly broken at the Lagrangian level. If we do so by adding a tiny explicit $U(1)$ symmetry breaking term, such as a $\mu'^2\Phi^2 + \text{h.c.}$ in the potential, the resulting G will not be massless, but has a small mass, proportional to μ'. In that case, G is called a pseudo-Goldstone boson. The QCD axion is a good example of this.

1.4.2 Nonlinear realization

We can repeat the above discussion by using a different way

$$\Phi = \frac{v + S(x)}{\sqrt{2}} e^{iG(x)/v} \ . \tag{128}$$

where $S(x)$ and $G(x)$ are excitations along the radial and angular directions, respectively. The resulting Lagrangian now takes the form

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}\partial_\mu S \partial^\mu S + \frac{(v+S)^2}{2}\partial_\mu \left[e^{iG(x)/v}\right] \partial^\mu \left[e^{-iG(x)/v}\right] - V \\ &= \frac{1}{2}\partial_\mu S \partial^\mu S + \frac{1}{2}\partial_\mu G \partial^\mu G + \left(\frac{S^2}{2v^2} + \frac{S}{v}\right)\partial_\mu G \partial^\mu G - \frac{\lambda}{4}(S^2+2vS)^2 \ . \end{aligned} \tag{129}$$

Again, G is the massless Goldstone boson. In this realization, the Goldstone boson is always derivatively coupled.

Remarkably, the linear and nonlinear theories of Goldstone are equivalent to each other at low energies, with the same S matrix elements. We will elaborate on this in Section 1.5.

1.4.3 Hidden symmetry: scalar QED example

Next, we move on to discuss the spontaneous breaking of $U(1)$ gauge symmetry. The model we consider is called scalar QED,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + D_\mu\Phi(x)\left(D^\mu\Phi(x)\right)^* - V(|\Phi(x)|^2) \ , \tag{130}$$

where the covariant derivative is

$$D_\mu = \partial_\mu - igqA_\mu(x) \ , \tag{131}$$

where q is the $U(1)$ charge of Φ. Before SSB, the Lagrangian is invariant under the gauge transformations

$$\Phi(x) \to \Phi(x)e^{igq\omega(x)} \ , \quad A_\mu(x) \to A_\mu(x) + \partial_\mu\omega(x) \ . \tag{132}$$

Now if $\lambda > 0$, $\mu^2 < 0$ in the scalar potential, SSB will occur. The VEV of Φ is the same as Eq. (125) and is not invariant under the above gauge transformation. To work out the resulting particle spectrum, it is most convenient to work with the nonlinear realization Eq. (128). The Lagrangian in terms of S, G, A_μ is

$$\begin{aligned}\mathcal{L} = &-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\partial_\mu S\partial^\mu S + \frac{1}{2}g^2q^2v^2\left(1+\frac{S}{v}\right)^2\left(A_\mu - \frac{\partial_\mu G}{gqv}\right)\left(A^\mu - \frac{\partial^\mu G}{gqv}\right) \\ &- \frac{\lambda}{4}(S^2+2vS)^2 \ .\end{aligned} \tag{133}$$

Here is a remarkable observation. We can completely remove the G field from the above Lagrangian by redefining the vector field

$$A'_\mu(x) = A_\mu(x) - \frac{1}{gqv}\partial_\mu G(x) \ . \tag{134}$$

The Lagrangian becomes

$$\mathcal{L} = -\frac{1}{4}F'_{\mu\nu}F'^{\mu\nu} + \frac{1}{2}\partial_\mu S\partial^\mu S + \frac{1}{2}g^2q^2v^2\left(1+\frac{S}{v}\right)^2 A'_\mu A'^\mu - \frac{\lambda}{4}(S^2+2vS)^2 \ , \tag{135}$$

where we used $F'_{\mu\nu} = F_{\mu\nu}$. This Lagrangian describes a massive real scalar boson S and a massive vector boson A'_μ, and their interactions. The masses are

$$m_{A'} = gqv \ , \quad m_S^2 = 2\lambda v^2 \ . \tag{136}$$

With the SSB of a gauge symmetry, the corresponding gauge boson becomes massive. This is called the Higgs mechanism.

The choice of vector field redefinition to get rid of G is called the unitary gauge. The price is to make the gauge boson massive. In the unitary gauge, there is G field to talk about, it is a would-be Goldstone boson. It would be a true Goldstone boson if the gauge symmetry was global.

1.5 Soft pion limit in the σ model

We close this section by discussing a fun subject, the low energy limit of the σ model and the equivalence of linear and nonlinear realizations. The Lagrangian is

$$\mathcal{L} = \partial_\mu \Phi \partial^\mu \Phi^\dagger - \lambda \left(|\Phi|^2 - v^2/2 \right)^2 \ . \tag{137}$$

The scalar potential is the same as that in Eq. (120), up to shift in the vacuum energy.

1.5.1 Linear σ Model

Define

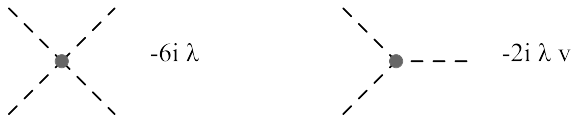
$$\Phi = \frac{1}{\sqrt{2}} \left(v + S + iG \right) \ . \tag{138}$$

The above Lagrangian can be written in terms of the S and G fields,

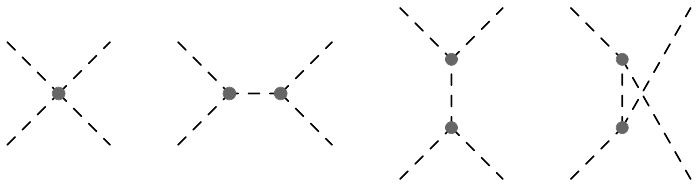
$$\mathcal{L} = \frac{1}{2} \partial_\mu S \partial^\mu S + \frac{1}{2} \partial_\mu G \partial^\mu G - \lambda v^2 S^2 - \frac{\lambda}{4} (S^2 + G^2)^2 - \lambda v S (S^2 + G^2) \ . \tag{139}$$

The mass of the "Higgs" boson S is given by $m_S^2 = 2\lambda v^2$. The goldstone boson G is massless.

Consider the very low energy scattering process $GG \to GG$, with center-of-mass energy well below m_S. The interaction Feynman rules relevant for such a process are:



There are four Feynman diagrams for the $GG \to GG$ scattering:



Applying the above Feynman rules, we obtain the scattering amplitude

$$\begin{aligned} i\mathcal{M} = &-6i\lambda + (-2i\lambda v)\frac{i}{s-m_S^2}(-2i\lambda v) + (-2i\lambda v)\frac{i}{t-m_S^2}(-2i\lambda v) \\ &+ (-2i\lambda v)\frac{i}{u-m_S^2}(-2i\lambda v) \ , \end{aligned} \tag{140}$$

where $s=(p_1+p_2)^2, t=(p_1-p_3)^2, u=(p_2-p_3)^2$ are the Mandelstam variables for $2\to 2$ scattering.

In the low energy scattering limit, $s, |t|, |u|$ are all much smaller than m_S^2. We Taylor expand the above S propagator up to second order [using $(\varepsilon - x)^{-1} \simeq -1/x - \varepsilon/x^2 - \varepsilon^2/x^3$],

$$i\mathcal{M} \simeq -6i\lambda - 4i\lambda^2 v^2 \left[-\frac{3}{m_S^2} - \frac{s+t+u}{m_S^4} - \frac{s^2+t^2+u^2}{m_S^6} \right] \ . \tag{141}$$

Using $m_S^2 = 2\lambda v^2$ and $s+t+u=0$ (note the Goldstone bosons are massless), we find

$$i\mathcal{M} \simeq \frac{i(s^2+t^2+u^2)}{2\lambda v^4} \ . \tag{142}$$

1.5.2 Non-linear σ Model

In this case, we define

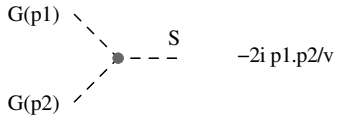
$$\Phi = \frac{1}{\sqrt{2}}\,(v+S)\,e^{iG/v} \ . \tag{143}$$

The original Lagrangian can be written as

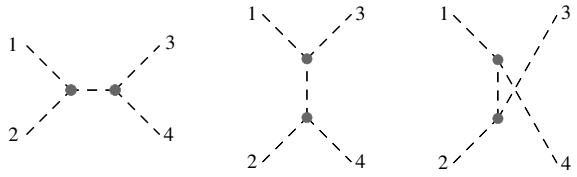
$$\mathcal{L} = \frac{1}{2}\partial_\mu S \partial^\mu S + \frac{1}{2}\partial_\mu G \partial^\mu G - \lambda v^2 S^2 + \left(\frac{S^2}{2v^2} + \frac{S}{v}\right)\partial_\mu G \partial^\mu G - \frac{\lambda}{4}(S^2+2vS)^2 \ . \tag{144}$$

Still, the mass of the radial mode S is $m_S^2 = 2\lambda v^2$ and the goldstone boson G remains massless. Here the Goldstone boson is always derivatively coupled in interactions. And there is no contact $(\partial G)^4$ term.

The only Feynman rule relevant for the $GG \to GG$ scattering is



The corresponding Feynman diagrams are



The sum of three scattering amplitudes are

$$\begin{aligned} i\mathcal{M} &= \frac{-2ip_1\cdot p_2}{v}\frac{i}{s-m_S^2}\frac{-2ip_3\cdot p_4}{v} + \frac{-2ip_1\cdot p_3}{v}\frac{i}{t-m_S^2}\frac{-2ip_2\cdot p_4}{v} \\ &\quad + \frac{-2ip_1\cdot p_4}{v}\frac{i}{u-m_S^2}\frac{-2ip_2\cdot p_3}{v} \\ &= -\frac{i}{v^2}\left[\frac{s^2}{s-m_S^2}+\frac{t^2}{t-m_S^2}+\frac{u^2}{u-m_S^2}\right] , \end{aligned} \tag{145}$$

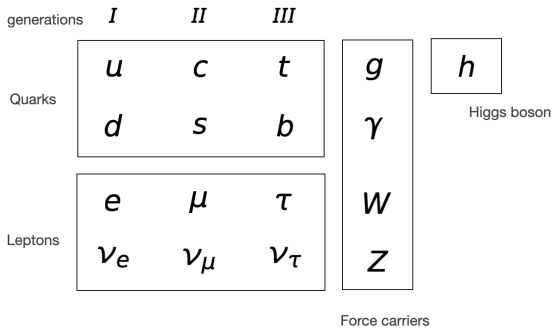
where we used the relations, $s = 2p_1\cdot p_2 = 2p_3\cdot p_4$, $t = -p_1\cdot p_3 = -p_2\cdot p_4$, $u = -2p_1\cdot p_4 = -2p_2\cdot p_3$, which are valid for massless Goldstone bosons.

For low energy scattering, we expand the S propagators to zeroth order, and get

$$i\mathcal{M} \simeq \frac{i(s^2+t^2+u^2)}{v^2m_S^2} = \frac{i(s^2+t^2+u^2)}{2\lambda v^4} . \tag{146}$$

Clearly, Eq. (10) and Eq. (6) give the same result. This is an example showing how the *Soft Theorem* works. At very low energies, the excitations along the imaginary axis and in the angular direction asymptot to each other. As a result, all S-matrix elements involving the goldstone boson are identical.

2 The Standard Model Lagrangian



All particles in the above table are physical mass eigenstates.

2.1 The Lagrangian

We start from the "weak basis" before electroweak symmetry breaking. The particles or fields are not in the mass basis yet.

The gauge groups of the Standard Model (SM) are

$$SU(3)_c \times SU(2)_L \times U(1)_Y \ , \tag{147}$$

that give strong and electroweak interactions.

The $SU(3)$ gauge group has 8 generators, each given by half of a Gell-Mann matrix. Correspondingly, there are 8 gauge bosons, called gluons, G_μ^α, with $\alpha = 1, \cdots, 8$.

The $SU(2)_L$ gauge group has 3 generators, each given by half of a Pauli matrix. The corresponding gauge bosons are denoted by W_μ^a, with $a = 1, 2, 3$.

The $U(1)_Y$ is called the hypercharge gauge group, with one gauge boson B_μ.

For the fermion sector, because the SM is chiral, all fermions are separated into left-handed and right-handed fields, excepted for neutrinos which are only left-handed. They organize into fundamental representations or gauge singlet under the above gauge groups.

We introduce some convenions for indices:

- Lorentz indices are μ, ν, ρ, σ.
- Color indices for gluons are $\alpha, \beta, \gamma \cdots = 1, \cdots, 8$.
- Color indices for quarks are $i, j, k \cdots = 1, 2, 3$.
- Generation indices are $m, n = 1, 2, 3$.
- $SU(2)_L$ indices of fermions are kept implicit. We write them as matrices.

There are three generation (copies) of fermions. For each generation, left-handed quarks and leptons form $SU(2)_L$ doublets

$$Q_{Lmi} = \begin{pmatrix} u_{Lmi} \\ d_{Lmi} \end{pmatrix} \ , \quad (Y = 1/6) \ , \quad L_{Lm} = \begin{pmatrix} \nu_{Lm} \\ e_{Lm} \end{pmatrix} \ , \quad (Y = -1/2) \ , \tag{148}$$

whereas their right-handed counterparts are singlet under the $SU(2)_L$,

$$u_{Rmi} \ , \quad (Y = 2/3) \ , \quad d_{Rmi} \ , \quad (Y = -1/3) \ , \quad e_{Rm} \ , \quad (Y = -1) \ . \tag{149}$$

The corresponding value of hypercharge is given by Y in bracket behind each field above.

In addition, the SM also contains one $SU(2)_L$ doublet scalar field

$$\Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \ , \tag{150}$$

with hypercharge $Y = 1/2$. Both ϕ_1 and ϕ_2 are complex scalars.

That's all the particle content. The SM is constructed with gauge symmetry as the guiding principle. It contains no pure gauge singlet.

With the above particles, we can write down the Lagrangian of SM

$$\begin{aligned} \mathcal{L}_{\rm SM} = & -\frac{1}{4} G^{\alpha}_{\mu\nu} G^{\alpha\mu\nu} - \frac{1}{4} W^{a}_{\mu\nu} W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\ & + \bar{Q}_{Lmi} i \not{D} Q_{Lmi} + \bar{L}_{Lm} i \not{D} L_{Lm} + \bar{u}_{Rmi} i \not{D} u_{Rmi} + \bar{d}_{Rmi} i \not{D} d_{Rmi} + \bar{e}_{Rm} i \not{D} e_{Rm} \\ & + (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi^\dagger \Phi) \\ & - \left[(Y_u)_{mn} \bar{Q}_{Lmi} \tilde{\Phi} u_{Rmi} + (Y_d)_{mn} \bar{Q}_{Lmi} \Phi d_{Rmi} + (Y_e)_{mn} \bar{L}_{Lm} \Phi e_{Rm} + \text{h.c.} \right] \ . \end{aligned} \tag{151}$$

To write down the Yukawa coupling for generating the up-type quark mass, we introduced

$$\tilde{\Phi} = i \sigma_2 \Phi^* \ , \tag{152}$$

which transforms the same way as Φ under $SU(2)_L$ (can be shown using the identity $\sigma_2 \vec{\sigma}^* = -\vec{\sigma} \sigma_2$), but has a hypercharge $-1/2$.

The covariant derivative acting on various fields take the form

$$\begin{aligned} D_\mu Q_{Lmi} &= \partial_\mu Q_{Lmi} - i \left(g_3 G^{\alpha}_{\mu} \frac{\lambda^{\alpha}_{ij}}{2} + g_2 W^{a}_{\mu} \frac{\sigma^a}{2} \delta_{ij} + g_1 B_\mu \frac{1}{6} \delta_{ij} \right) Q_{Lmj} \ , \\ D_\mu u_{Rmi} &= \partial_\mu u_{Rmi} - i \left(g_3 G^{\alpha}_{\mu} \frac{\lambda^{\alpha}_{ij}}{2} + \not{0} + g_1 B_\mu \frac{2}{3} \delta_{ij} \right) u_{Rmj} \ , \\ D_\mu d_{Rmi} &= \partial_\mu d_{Rmi} - i \left(g_3 G^{\alpha}_{\mu} \frac{\lambda^{\alpha}_{ij}}{2} + \not{0} + g_1 B_\mu \left(-\frac{1}{3} \right) \delta_{ij} \right) d_{Rmj} \ , \\ D_\mu L_{Lm} &= \partial_\mu L_{Lm} - i \left(\not{0} + g_2 W^{a}_{\mu} \frac{\sigma^a}{2} + g_1 B_\mu \left(-\frac{1}{2} \right) \right) L_{Lm} \ , \\ D_\mu e_{Rm} &= \partial_\mu e_{Rm} - i \left(\not{0} + \not{0} + g_1 B_\mu \left(-1 \right) \right) e_{Rm} \ , \\ D_\mu \Phi &= \partial_\mu \Phi - i \left(\not{0} + g_2 W^{a}_{\mu} \frac{\sigma^a}{2} + g_1 B_\mu \left(\frac{1}{2} \right) \right) \Phi \ . \end{aligned} \tag{153}$$

2.2 Electroweak symmetry breaking

The electroweak symmetries $SU(2)_L \times U(1)_Y$ are broken because the scalar doublet Φ develops a VEV. The scalar potential in Eq. (151) takes the form

$$V(\Phi^\dagger\Phi) = \lambda\left(\Phi^\dagger\Phi - v^2/2\right)^2 \ , \tag{154}$$

where v is a real parameter. Such a potential is minimized with

$$\langle\Phi^\dagger\Phi\rangle = |\phi_1|^2 + |\phi_2|^2 = v^2/2 \ . \tag{155}$$

The freedom to assign the VEV to either real or imaginary parts of ϕ_1 and ϕ_2 corresponds to the original $SU(2)_L$ symmetry. Hereafter, we make a choice where

$$\langle\Phi\rangle = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \ , \tag{156}$$

i.e., only the real part of ϕ_2 gets a VEV. As discussed earlier, this VEV is the classical field value of ϕ_2 that fills throughout the spacetime.

Next, we talk about excitations. In the nonlinear realization, the scalar doublet can be organized as

$$\Phi = \exp\left(iG_a\sigma_a/v\right)\begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \ , \tag{157}$$

where $h(x)$ and $G_a(x)$ are the Higgs boson and three would-be Goldstone bosons, respectively. Making analogy to the $U(1)$ example discussed in section 1.4.3, we can perform an $SU(2)_L \times U(1)_Y$ rotation and remove the would-be goldstone bosons with a redefinition of the gauge bosons. This leads us to the unitary gauge, where

$$\Phi = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \ . \tag{158}$$

Plugging Eq. (158) into the kinetic term of Φ in Eq. (151), we find

$$\begin{aligned} D_\mu\Phi &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \partial_\mu h \end{pmatrix} - \frac{i}{2\sqrt{2}}\begin{pmatrix} g_2W^3\mu + g_1B_\mu & \sqrt{2}W^+_\mu \\ \sqrt{2}W^-_\mu & -g_2W^3\mu + g_1B_\mu \end{pmatrix}\begin{pmatrix} 0 \\ v+h \end{pmatrix} \ , \\ (D_\mu\Phi)^\dagger(D^\mu\Phi) &= \frac{1}{2}\partial_\mu h\partial^\mu h \\ &\quad + \frac{1}{8}\left[2g_2^2W^-_\mu W^{+\mu} + (-g_2W^3_\mu + g_1B_\mu)(-g_2W^{3\mu} + g_1B^\mu)\right](v+h)^2 \ , \end{aligned} \tag{159}$$

where we have defined the charged W-boson fields

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \ . \tag{160}$$

Reading from Eq. (159), we find some gauge bosons become massive. First, the charged W-boson mass is

$$M_W = \frac{1}{2} g_2 v \ . \tag{161}$$

Nest, it is useful to define the neutral Z-boson

$$Z_\mu = \frac{1}{\sqrt{g_1^2 + g_2^2}} (g_2 W_\mu^3 - g_1 B_\mu) = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \ , \tag{162}$$

where we introduce the weak mixing angle θ_W,

$$\tan\theta_W \equiv \frac{g_1}{g_2} \ . \tag{163}$$

The mass of the Z-boson is

$$M_Z = \frac{1}{2}\sqrt{g_1^2 + g_2^2}\, v = \frac{M_W}{\cos\theta_W} \ . \tag{164}$$

The last line is a tree-level prediction of the SM. It originates from a remaining custodial symmetry $SO(4) \to SO(3)$ of the scalar potential. Experimentally, $M_{W,Z}$ and θ_W can be measured independently, leading to a test of the SM.

There is one massless gauge boson after the Higgs mechanism, which is orthogonal to the Z boson,

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \ . \tag{165}$$

It will play the role of the photon, for electromagnetism.

Plugging Eq. (158) into the scalar potential Eq. (154), we get

$$V(h) = \frac{\lambda}{4}(h^2 + 2vh)^2 \ . \tag{166}$$

Picking the quadratic term, we find the Higgs boson mass is

$$m_h^2 = 2\lambda v^2 \ . \tag{167}$$

The rest are Higgs boson self-interaction terms. These interactions are yet to be directly measured experimentally.

A nonzero electroweak breaking VEV also generate mass for quarks and charged leptons, through the Yukawa couplings in Eq. (151). Using $\langle\tilde{\Phi}\rangle = (v/\sqrt{2}, 0)^T$, we get the fermion mass terms

$$\mathcal{L}_m = -\left[\bar{u}_L M_u u_R + \bar{d}_L M_d d_R + \bar{e}_L M_e e_R + \text{h.c.}\right] \ , \tag{168}$$

where

$$M_u = Y_u \frac{v}{\sqrt{2}} \ , \quad M_d = Y_d \frac{v}{\sqrt{2}} \ , \quad M_e = Y_e \frac{v}{\sqrt{2}} \ , \tag{169}$$

are 3×3 mass matrices in the generation space. They are complex matrices. In general, each can be diagonalized with two unitary matrices

$$\begin{aligned} M_u &= U_u^\dagger \begin{pmatrix} m_u & & \\ & m_c & \\ & & m_t \end{pmatrix} V_u \equiv U_u^\dagger \hat{M}_u V_u \ , \\ M_d &= U_d^\dagger \begin{pmatrix} m_d & & \\ & m_s & \\ & & m_b \end{pmatrix} V_d \equiv U_d^\dagger \hat{M}_d V_d \ , \\ M_e &= U_e^\dagger \begin{pmatrix} m_e & & \\ & m_\mu & \\ & & m_\tau \end{pmatrix} V_e \equiv U_e^\dagger \hat{M}_e V_e \ , \end{aligned} \tag{170}$$

where the mass eigenvalues in the diagonalized matrices can be made real by properly redefining the phase of the fermion fields.

If we define,

$$\begin{aligned} \begin{pmatrix} u_{1L} \\ u_{2L} \\ u_{3L} \end{pmatrix} &= U_u \begin{pmatrix} u_L \\ c_L \\ t_L \end{pmatrix} \ , \quad \begin{pmatrix} u_{1R} \\ u_{2R} \\ u_{3R} \end{pmatrix} = V_u \begin{pmatrix} u_R \\ c_R \\ t_R \end{pmatrix} \ , \\ \begin{pmatrix} d_{1L} \\ d_{2L} \\ d_{3L} \end{pmatrix} &= U_d \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} \ , \quad \begin{pmatrix} d_{1R} \\ d_{2R} \\ d_{3R} \end{pmatrix} = V_d \begin{pmatrix} d_R \\ s_R \\ b_R \end{pmatrix} \ , \\ \begin{pmatrix} e_{1L} \\ e_{2L} \\ e_{3L} \end{pmatrix} &= U_e \begin{pmatrix} e_L \\ \mu_L \\ \tau_L \end{pmatrix} \ , \quad \begin{pmatrix} e_{1R} \\ e_{2R} \\ e_{3R} \end{pmatrix} = V_e \begin{pmatrix} e_R \\ \mu_R \\ \tau_R \end{pmatrix} \ , \end{aligned} \tag{171}$$

the mass term Lagrangian can be written as

$$\begin{aligned} \mathcal{L}_m = &- \begin{pmatrix} \bar{u}_L & \bar{c}_L & \bar{t}_L \end{pmatrix} \begin{pmatrix} m_u & & \\ & m_c & \\ & & m_t \end{pmatrix} \begin{pmatrix} u_R \\ c_R \\ t_R \end{pmatrix} - \begin{pmatrix} \bar{d}_L & \bar{s}_L & \bar{b}_L \end{pmatrix} \begin{pmatrix} m_d & & \\ & m_s & \\ & & m_b \end{pmatrix} \begin{pmatrix} d_R \\ s_R \\ b_R \end{pmatrix} \\ &- \begin{pmatrix} \bar{e}_L & \bar{\mu}_L & \bar{\tau}_L \end{pmatrix} \begin{pmatrix} m_e & & \\ & m_\mu & \\ & & m_\tau \end{pmatrix} \begin{pmatrix} e_R \\ \mu_R \\ \tau_R \end{pmatrix} + \text{h.c.} \ . \end{aligned} \tag{172}$$

Now in this Lagrangian, all the fermions are in the mass basis. They are physical fields shown in the table at the beginning of this section.

In summary, spontaneous breaking of the electroweak symmetries generates masses for the W, Z bosons, quarks and charged leptons. Their masses were forbidden by gauge invariance in the first place. Now that the gauge symmetries are spontaneously broken by the scalar doublet VEV, mass terms become possible. No wonder all the nonzero masses are proportional to the VEV v.

After electroweak symmetry breaking, the particles are remain massless include the photon (due to a remaining good gauge symmetry) and neutrinos (due

to the absence of right-handed neutrinos). The $SU(3)_c$ for strong interaction is not spontaneously broken, but it generates a mass gap for gluons at low energy for a different reason. In reality, neutrino oscillation experiments have shown that neutrinos must be massive. This calls for explanation from beyond the SM physics.

2.3 Interactions

2.3.1 Higgs boson interactions

To derive the Higgs boson couplings to gauge bosons and fermions, here is a useful rule of thumb. The observation is that the Higgs field only arises from the scalar doublet Φ, along with the electroweak VEV. See Eq. (158). As a result, whenever a v occurs in the Lagrangian, it is legitimate make the replacement

$$v \to v\left(1+\frac{h}{v}\right) \ . \tag{173}$$

Doing this for the mass terms, it can correctly generate Higgs interactions with the corresponding massive particles. [2]

The interactions between the Higgs boson and W, Z bosons are derived from their mass terms, by making the above replacement and using $M^2_{W,Z} \propto v^2$,

$$\begin{aligned} \mathcal{L}_{hVV} &= M_W^2\left(1+\frac{h}{v}\right)^2 W^+_\mu W^{-\mu} + \frac{1}{2}M_Z^2\left(1+\frac{h}{v}\right)^2 Z_\mu Z^\mu \\ &\supset M_W^2\left(\frac{h^2}{v^2}+\frac{2h}{v}\right) W^+_\mu W^{-\mu} + \frac{M_Z^2}{2}\left(\frac{h^2}{v^2}+\frac{2h}{v}\right) Z_\mu Z^\mu \ . \end{aligned} \tag{174}$$

Because the photon and gluons remain massless after electroweak symmetry breaking, the Higgs boson does not couple to them at tree level. Their couplings are generated at loop level.

The interactions between the Higgs boson and massive fermion f are derived similarly, by noticing $m_f \propto v$,

$$\mathcal{L}_{hf\bar{f}} = -m_f\left(1+\frac{h}{v}\right)\bar{f}f \supset -\frac{m_f}{v}h\bar{f}f \ . \tag{175}$$

2.3.2 Charged-current interactions

The charged-current (CC) interactions describes how the $W^\pm$ bosons couple to fermions. Starting from the original Lagrangian Eq. (151), in the covariant derivative terms, the interaction terms for $W^{1,2}_\mu$ bosons are

$$\mathcal{L}_{CC} = \frac{g_2}{\sqrt{2}}\left(\bar{u}_{Lm}\gamma^\mu d_{Lm} + \bar{\nu}_{Lm}\gamma^\mu e_{Lm}\right) W^+_\mu + \text{h.c.} \tag{176}$$

[2] Note the above replacement rule breaks down for the scalar potential which contains the v^2 parameter which does not originate from the doublet VEV. Instead, it is the origin for the VEV. This only affects the Higgs self-interactions, and we should simply expand Eq. (166).

Here the u_{Lm}, d_{Lm}, e_{Lm} are still in the weak basis. Going to the mass basis using the relations in Eq. (171), we get

$$\begin{aligned}\mathcal{L}_{CC} &= \frac{g_2}{\sqrt{2}}\begin{pmatrix}\bar{u} & \bar{c} & \bar{t}\end{pmatrix} U_u^\dagger U_d \mathbb{P}_L \begin{pmatrix}d\\ s\\ b\end{pmatrix} W_\mu^+ + \frac{g_2}{\sqrt{2}}\begin{pmatrix}\bar{\nu}_{1L} & \bar{\nu}_{2L} & \bar{\nu}_{3L}\end{pmatrix} U_e \mathbb{P}_L \begin{pmatrix}e\\ \mu\\ \tau\end{pmatrix} W_\mu^+ + \text{h.c.} \\ &= \frac{g_2}{\sqrt{2}}\begin{pmatrix}\bar{u} & \bar{c} & \bar{t}\end{pmatrix} V_{\text{CKM}} \mathbb{P}_L \begin{pmatrix}d\\ s\\ b\end{pmatrix} W_\mu^+ + \frac{g_2}{\sqrt{2}}\begin{pmatrix}\bar{\nu}_e & \bar{\nu}_\mu & \bar{\nu}_\tau\end{pmatrix} U_e \mathbb{P}_L \begin{pmatrix}e\\ \mu\\ \tau\end{pmatrix} W_\mu^+ + \text{h.c.}\end{aligned} \tag{177}$$

In the second line, we introduce the Cabibbo-Kobayash-Maskawa matrix as the product of left-handed up- and down-type quark rotation matrices. Because neutrinos are massless in the SM, we are free to take arbitrary linear combinations of them. For convenience, we define neutrino flavor eigenstates ν_e, ν_μ, ν_τ so that the CC interactions in the lepton sector are flavor diagonal.

The Wolfenstein parametrization of the CKM matrix takes the form

$$V_{\text{SKM}} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} , \tag{178}$$

where $\lambda = 0.22$ is a small parameter. The other parameters A, ρ, η are all order one. The CKM matrix is very close to a unit matrix, but with small flavor off-diagonal elements. This is the only way in the SM where fermions from different generations can couple to each other, which plays important role in heavy quark decays. The CKM matrix element is also complex in the (13) and (31) elements leading to CP violation effects.

Establishing the CKM structure as the origin of quark flavor and CP violation effects in the SM is an extremely exciting chapter that lasted until the first decade of this century. For more details, see `https://pdg.lbl.gov/2022/reviews/rpp2022-rev-ckm-matrix.pdf`.

2.3.3 Neutral-current interactions

The neutral-current (NC) interactions describes how the Z boson and photon couple to fermions. To derive them we reverse the relations in Eqs. (162) and (165), and resort to the gauge coupling terms of W_μ^3 and B_μ in the original Lagrangian Eq. (151). Without loss of generality, we write down their couplings with a fermion f, where f can be any fermion in the SM,

$$\mathcal{L}_{NC} = \bar{f}_L \left(g_2 \gamma^\mu W_\mu^3 T_{3L} + g_1 \gamma^\mu B_\mu Y_L\right) f_L + \bar{f}_R g_1 \gamma^\mu B_\mu Y_R f_R \ , \tag{179}$$

where T_{3L} is called the weak isospin quantum number, $T_{3L} = \frac{1}{2}$ for up-type quarks and neutrinos, $T_{3L} = -\frac{1}{2}$ for down-type quarks and charged leptons. For convenience we can define $T_{3L} = 0$ for all right-handed fermions. $Y_{L,R}$ are the hypercharges for f_L, f_R, respectively.

In terms of the Z_μ and A_μ fields, the NC interactions read

$$\begin{aligned}\mathcal{L}_{NC} = & e\bar{f}\gamma^\mu\left[(T_{3L}+Y_L)\mathbb{P}_L + Y_R\mathbb{P}_R\right]fA_\mu \\ & + \frac{e}{\sin\theta_W\cos\theta_W}\bar{f}\gamma^\mu\left[T_{3L}\mathbb{P}_L - Q\sin^2\theta_W\right]fZ_\mu \ .\end{aligned} \tag{180}$$

The first line corresponds to the regular QED interactions. It is a vector current interaction because for all SM fermions, we have $T_{3L} + Y_L = Y_R$. This allows us to define the electric charge of a fermion as

$$Q = T_{3L} + Y \ . \tag{181}$$

The second line in Eq. (180) is the neutral-current interaction mediated by the massive Z-boson. It is often useful to rewrite it as

$$\mathcal{L}^Z_{NC} = g_Z\bar{f}\gamma^\mu\left[g_V + g_A\gamma_5\right]fZ_\mu \ , \tag{182}$$

where

$$g_Z = \frac{e}{\sin\theta_W\cos\theta_W} \ , \quad g_V = \frac{1}{2}T_{3L} - Q\sin^2\theta_W \ , \quad g_A = -\frac{1}{2}T_{3L} \ . \tag{183}$$

Importantly, the Z coupling is universal for all three generations, thus we can directly consider the above f as the physical mass eigenstates.

2.3.4 Gluon interactions

The gluon quark interactions are also universal all three generations, thus we can directly work in the mass basis and write down

$$\mathcal{L}_{\mathrm{gf\bar{f}}} = \sum_{q=u,d,s,c,b,t}\frac{g_3}{2}\bar{q}_i\gamma^\mu\lambda^\alpha_{ij}q_jG^\alpha_\mu \ . \tag{184}$$

2.3.5 Gauge boson self-interactions

The non-abelian gauge bosons also have self-interactions, generated from their kinetic term. For example, gluons have three- and four-point self-interactions.

Because the $W^\pm$ bosons carry electric charge, they can form a current and couple to Z and γ.

Because the $U(1)$ hypercharge gauge boson does not self-interact, there are no interactions involving Z-bosons and photons.

The reader can find all these interactions in the appendix of textbook by Cheng and Li.

2.4 Gauge anomaly cancellation

2.4.1 What goes wrong with the ABJ anomaly?

In this note we discuss the problem with gauge anomalies in a toy model with gauged $U(1)_L \times U(1)_{e.m.}$ symmetries. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}X_{\mu\nu}X^{\mu\nu} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\gamma^\mu\partial_\mu\psi + e\bar{\psi}\gamma^\mu\psi A_\mu + g\bar{\psi}_L\gamma^\mu\psi_L X_\mu \ , \tag{185}$$

where A_μ is the photon and $F_{\mu\nu}$ is the corresponding field strength. X_μ is the gauge boson of the new $U(1)_L$ symmetry. Only the left-handed fermion ψ_L carries the $U(1)_L$ charge.

Under the $U(1)_L$ gauge transformation,

$$\psi_L \to e^{i\omega(x)}\psi_L \ , \quad \psi_R \to \psi_R \ , \quad X_\mu \to X_\mu + \partial_\mu \omega(x) \ , \tag{186}$$

The Lagrangian is invariant. So it looks like a good symmetry.

Due to the chiral nature of this theory, we are not allowed to write down a mass term for the fermion ψ.

We could further introduce a scalar charged under the $U(1)_L$ and write down its Yukawa coupling to $\bar{\psi}_L \psi_R$. It could generate a fermion mass if ψ gets a VEV. This corresponds to the spontaneous $U(1)_L$ symmetry breaking, which we are not interested here.

The problem occurs at quantum level. Through the triangle diagram with XAA as external states (details omitted), we can find the left-handed current $J^\mu_L = \bar{\psi}_L \gamma^\mu \psi_L$ is not conserved, because it contains an axial current part,

$$\partial_\mu J^\mu_L = -\frac{1}{2}\partial_\mu(\bar{\psi}\gamma^\mu\gamma_5\psi) = -\frac{1}{2} \times (-\frac{e^2}{16\pi^2})\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = \frac{e^2}{32\pi^2}\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} \ . \tag{187}$$

So the current is not conserved!

This is in contradiction with the general expectation that before any symmetry breaking, a gauge boson must couple to a conserved current. This is the origin of the Ward identity. In the above simple model, the Ward identity can be derived using the Euler-Lagrange equation for X_μ, which takes the form

$$\Box A^\mu - \partial^\mu \partial_\nu A^\nu + g J^\mu_L = 0 \ . \tag{188}$$

Apply another ∂_μ to every term, we get

$$\partial_\mu J^\mu_L = 0 \ , \tag{189}$$

which implies the current J^μ_L is conserved.

In the presence of the quantum level anomaly, to make it consistent with the equation of motion, we need to modify the original Lagrangian by adding an extra term (called the Wess-Zumino term)

$$\Delta\mathcal{L} = -g\frac{e^2}{16\pi^2}\varepsilon^{\mu\nu\rho\sigma}X_\mu A_\nu F_{\rho\sigma} \ . \tag{190}$$

This will add another term to the right-hand side of the above equation of motion Eq. (188), which now takes the form

$$\Box A^\mu - \partial^\mu \partial_\nu A^\nu + g J^\mu_L - g\frac{e^2}{16\pi^2}\varepsilon^{\mu\nu\rho\sigma}A_\nu F_{\rho\sigma} = 0 \ . \tag{191}$$

Like before, we apply another ∂_μ to every term which leads to

$$g\partial_\mu J^\mu_L - g\frac{e^2}{16\pi^2}\varepsilon^{\mu\nu\rho\sigma}\partial_\mu(A_\nu F_{\rho\sigma}) = 0 \ . \tag{192}$$

Clearly, in the last term, the derivative ∂_μ must hit A_ν for to be nonzero. We can rewrite it as (remove the common factor g)

$$\partial_\mu J_L^\mu - \frac{e^2}{32\pi^2}\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = 0 \ , \tag{193}$$

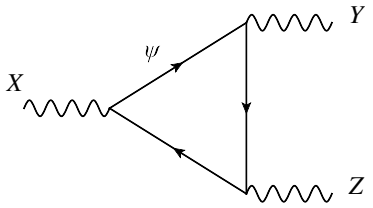
which is the same equation as Eq. (187).

However, the remaining problem here is that the Wess-Zumino term, Eq. (190), is not gauge invariant. In other word, the ABJ anomaly would make a gauge theory not gauge invariant. It must not exist or be canceled away.

2.4.2 How the SM works

Because the SM is chiral, it could suffer from the same problem of anomaly-induced loss of gauge invariance as the above toy model. Because the ABJ anomaly is independent of fermion masses, this is already an issue before electroweak symmetry breaking, as soon as we write down Eq. (151). One cannot argue it away by saying that the electroweak symmetry is broken thus gauge invariance is gone.

SM avoids this problem by cancellation.



Let's first consider a concrete case of the above triangle diagram where the X is W_μ^3 gauge boson, and Y, Z are both gluons. As the gauge boson from $SU(2)_L$, W^3 only couples to left-handed quarks and leptons. Among them only quarks can couple to gluons (Y, Z). Therefore, the relevant current that couples to W_μ^3 here is

$$J_L^{3\mu} = \frac{1}{2}\bar{Q}_L\gamma^\mu\sigma_3 Q_L = \frac{1}{2}\left(\bar{u}_L\gamma^\mu u_L - \bar{d}_L\gamma^\mu d_L\right) \ . \tag{194}$$

Therefore, although each term in this current suffers from the ABJ anomaly, the difference between $\bar{u}_L\gamma^\mu u_L$ and $\bar{d}_L\gamma^\mu d_L$ is anomaly free. This can be summarized as the traceless condition of σ_3. In other words, the triangle diagram here is proportional to

$$\mathrm{Tr}(\sigma_3)\mathrm{Tr}(\lambda_\alpha\lambda_\beta) = 0 \ . \tag{195}$$

With this experience, we can move on to enumerate all the combinations of X, Y, Z from the SM gauge groups.

- (3, 3, 3). f all the external gauge bosons are gluons. Because all gluons couple to vector currents which treat left- and right-handed fermions equally. There is no issue to ABJ anomaly.

- (3, 3, 2). We just worked out this example above and showed that it is anomaly free because $\mathrm{Tr}(\sigma_a) = 0$.

- (3, 3, 1). The hypercharge current is also chiral, thus the anomaly is potentially an issue. We have to put all quarks in the loop. Note that left-handed current and right-handed currents differ by a minus sign in their anomalies, due to the opposite signs in front of γ_5 (axial current part). The sum over all quark hypercharges from (Q_L, u_R, d_R) per generation is

$$3_{\text{color}} \times \left[2_{\text{doublet}} \times \frac{1}{6} - \frac{2}{3} - \left(-\frac{1}{3} \right) \right] = 0 \ . \tag{196}$$

- (3, 2, 2), (3, 2, 1), (3, 1, 1). They all vanish because they are proportional to a factor of $\mathrm{Tr}(\lambda^\alpha) = 0$.

- (2, 2, 2). This one is somewhat non-trivial. For each $SU(2)_L$ doublet fermion running in the loop, the triangle diagram is proportional to (has to be totally symmetric with respect to a, b, c)

$$\mathrm{Tr}(\sigma_a \{\sigma_b, \sigma_c\}) = 2\delta_{bc}\mathrm{Tr}(\sigma_a) = 0 \ . \tag{197}$$

  We are lucky because $SU(2)$ is special. Were $SU(3)$ chiral, the counterpart $\left\{\lambda^\alpha, \lambda^\beta\right\}$ is not proportional to a unit matrix and we would not get zero.

- (2, 2, 1). We need to sum up contributions from quark and lepton doublets. For each generation, the triangle diagram has a factor

$$3_{\text{color}} \times \frac{1}{6} + \left(-\frac{1}{2} \right) = 0 \ . \tag{198}$$

- (2, 1, 1). This contribution vanishes because $\mathrm{Tr}(\sigma_a) = 0$.

- (1, 1, 1). This contribution involves everybody. For each generation, it is proportional to

$$\begin{aligned} &3_{\text{color}} \times 2_{\text{doublet}} \times \left(\frac{1}{6} \right)^3 + 2_{\text{doublet}} \times \left(-\frac{1}{2} \right)^3 \\ &\quad - 3_{\text{color}} \times \left(\frac{2}{3} \right)^3 - 3_{\text{color}} \times \left(-\frac{1}{3} \right)^3 - (-1)^3 = 0 \ . \end{aligned} \tag{199}$$

- (1, gravity, gravity). This is called gravitational anomaly of $U(1)_Y$, which vanishes because

$$\begin{aligned} &3_{\text{color}} \times 2_{\text{doublet}} \times \left(\frac{1}{6} \right) + 2_{\text{doublet}} \times \left(-\frac{1}{2} \right) \\ &\quad - 3_{\text{color}} \times \left(\frac{2}{3} \right) - 3_{\text{color}} \times \left(-\frac{1}{3} \right) - (-1) = 0 \ . \end{aligned} \tag{200}$$

That is all the possibilities. We have verified that the SM is indeed free from all gauge anomalies.

As an additional remark, with the SM fermion fields, we can construct other currents such as the baryon or lepton number currents. They are also conserved currents at tree level and correspond to accidental global symmetries of the SM. However, after including the ABJ anomalies, they are no longer conserved. The above cancellations do not work for these global currents.

A theory with anomalous global currents is still healthy. They could even do good things for the nature, such as baryogenesis.

2.5 Θ terms

At renormalizable level, one can also add the following Θ terms to the SM Lagrangian,

$$\mathcal{L}_\Theta = \Theta_3 G^\alpha_{\mu\nu}\widetilde{G}^{\alpha\mu\nu} + \Theta_2 W^a_{\mu\nu}\widetilde{W}^{a\mu\nu} + \Theta_1 B_{\mu\nu}\widetilde{B}^{\mu\nu} \ , \tag{201}$$

where $\widetilde{G}^{\alpha\mu\nu} \equiv \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}G_{\rho\sigma}$, and similarly for W and B.

It is easiest to show that the last term is equivalent to a total derivative

$$B_{\mu\nu}\widetilde{B}^{\mu\nu} = 2(\partial_\mu B_\nu)\widetilde{B}^{\mu\nu} = 2\partial_\mu\left(B_\nu\widetilde{B}^{\mu\nu}\right) - 2B_\nu\partial_\mu\widetilde{B}^{\mu\nu} = \partial_\mu\left(\varepsilon^{\mu\nu\rho\sigma}B_\nu B_{\rho\sigma}\right) \ , \tag{202}$$

where we used $\partial_\mu\widetilde{B}^{\mu\nu}$ which is obvious.

For the non-abelian terms, there is an extra term in the "current". We work out the gluon case explicitly,

$$G^\alpha_{\mu\nu}\widetilde{G}^{\alpha\mu\nu} = \partial_\mu\left[\varepsilon^{\mu\nu\rho\sigma}\left(G^\alpha_\nu G^\alpha_{\rho\sigma} - \frac{1}{3}g_3 f^{\alpha\beta\gamma}G^\alpha_\nu G^\beta_\rho G^\gamma_\sigma\right)\right] \ . \tag{203}$$

It is straightforward to expand both sides the check the equality holds. One useful thing to note is the vanishment of the G^4 term

$$\begin{aligned}
&\varepsilon^{\mu\nu\rho\sigma}\left(f^{\alpha\beta\gamma}G^\beta_\mu G^\gamma_\nu\right)\left(f^{\alpha\beta'\gamma'}G^{\beta'}_\rho G^{\gamma'}_\sigma\right) \\
=&-2\varepsilon^{\mu\nu\rho\sigma}G^\beta_\mu G^\gamma_\nu G^{\beta'}_\rho G^{\gamma'}_\sigma \mathrm{Tr}\left([T^\beta,T^\gamma][T^{\beta'},T^{\gamma'}]\right) \\
=&-8\varepsilon^{\mu\nu\rho\sigma}G^\beta_\mu G^\gamma_\nu G^{\beta'}_\rho G^{\gamma'}_\sigma \mathrm{Tr}\left(T^\beta T^\gamma T^{\beta'}T^{\gamma'}\right) \\
=&+8\varepsilon^{\nu\rho\sigma\mu}G^\gamma_\nu G^{\beta'}_\rho G^{\gamma'}_\sigma G^\beta_\mu \mathrm{Tr}\left(T^\gamma T^{\beta'}T^{\gamma'}T^\beta\right) \ .
\end{aligned} \tag{204}$$

The last two lines are obviously opposite to each other, after renaming the labels correspondingly.

So all the Θ terms can be written as total derivatively. Naively, one would expect them to have no physical effect. This is indeed the case for the $SU(2)_L$ and $U(1)_Y$ Θ-terms. However, $SU(3)_c$ is different, because of its rich vacuum structure, and its coupling goes strong at low energies, leading to non-perturbative instanton processes that allow transitions to occur among various degenerate

minima carrying different topological numbers. These effects make the Θ_3 term physical. See `https://inspirehep.net/literature/3673` for more details.

Because the Θ terms is odd under parity and time-reversal transformations, a nonzero Θ_3 contributes to the neutron electric dipole moment. Currently experimental bound requires $\Theta_3 \lesssim 10^{10}$, in the basis where all quark masses are real.

2.6 Tools for calculation

2.6.1 Decay rate and cross section

The decay rate for $1 \to 2 + 3 + \cdots$ process is

$$\Gamma = \frac{1}{2M} \int d\Pi_f (2\pi)^4 \delta^4(p_1 - \sum_f p_f) \overline{|\mathcal{M}|^2} \ , \tag{205}$$

where M is mass of the decay particle 1.

The cross section for $1 + 2 \to 3 + 4 + \cdots$ process is

$$\sigma = \frac{1}{4f} \int d\Pi_f (2\pi)^4 \delta^4(p_1 + p_2 - \sum_f p_f) \overline{|\mathcal{M}|^2} \ , \tag{206}$$

where the prefactor can be rewritten as

$$\begin{aligned} 4f &= 4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2} = 4\sqrt{(E_1 E_2 - \vec{p}_1 \cdot \vec{p}_2)^2 - m_1^2 m_2^2} \\ &= 4E_1 E_2 \sqrt{(1 - \vec{v}_1 \cdot \vec{v}_2)^2 - (1 - |\vec{v}_1|^2)(1 - |\vec{v}_2|^2)} \\ &= 4E_1 E_2 \sqrt{(1 - 2\vec{v}_1 \cdot \vec{v}_2 + (\vec{v}_1 \cdot \vec{v}_2)^2) - (1 - |\vec{v}_1|^2 - |\vec{v}_2|^2 + |\vec{v}_1|^2 |\vec{v}_2|^2)} \\ &= 4E_1 E_2 \sqrt{(\vec{v}_1 - \vec{v}_2)^2 + (\vec{v}_1 \cdot \vec{v}_2)^2 - |\vec{v}_1|^2 |\vec{v}_2|^2} \\ &= 4E_1 E_2 \sqrt{(\vec{v}_1 - \vec{v}_2)^2 - (\vec{v}_1 \times \vec{v}_2)^2} \\ &\equiv 4E_1 E_2 v_{\text{Møller}} \ . \end{aligned} \tag{207}$$

In the last line we have introduced the definition of the Møller velocity,

$$v_{\text{Møller}} \equiv \sqrt{(\vec{v}_1 - \vec{v}_2)^2 - (\vec{v}_1 \times \vec{v}_2)^2}. \tag{208}$$

It is a Lorentz invariant quantity.

Special cases:

- For face-to-face scattering, such as LHC, Tevatron, LEP, we have $\vec{v}_1 /\!\!/ \vec{v}_2$. In this case $v_{\text{Møller}} = |\vec{v}_1 - \vec{v}_2|$ is equal to the relative velocity.
- For fixed-target experiments, we have $\vec{v}_2 = 0$. In this case $v_{\text{Møller}} = |\vec{v}_1 - \vec{v}_2| = |\vec{v}_1|$ is equal to the relative velocity.

- For non-relativistic particle scattering, such as WIMP dark matter freeze out in early universe, we have $|\vec{v}_1|, |\vec{v}_2| \ll c$. In this case $v_{\text{Møller}} \simeq |\vec{v}_1 - \vec{v}_2|$ is "approximately" the relative velocity.
- For ultra-relativistic particle interactions in a thermal plasma, such as very early universe, we have $|\vec{v}_1| = |\vec{v}_2| = c$. In general, the relative angle can take any value. In this case, $v_{\text{Møller}} = c\sqrt{2 - 2\cos\theta + \sin^2\theta} = (1 - \cos\theta)c$.

See the "kinematics" section of the Review of Particle Physics for further details, `https://pdg.lbl.gov/2022/reviews/rpp2022-rev-kinematics.pdf`, in particular, simplified two- and three-body final state phase space integrals for decay and scattering (for unpolarized processes).

2.6.2 SM at various energy scales

The Standard Model is a good theory and can successfully describe elementary particle physics processes from very low to very high energy scales.

The Higgs VEV is $v = 246\,\text{GeV}$, which is called the electroweak scale. All elementary particle masses in the SM are proportional to (below) this scale. If we build experiments that collide SM particles at very high center-of-mass energies, $E_{\text{CM}} \gg v$, the cross sections which is a function of E_{CM} and masses can be Taylor expanded in series of small m/E_{CM}

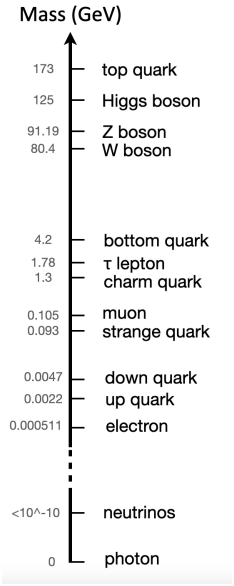
$$\sigma\left(E_{\text{CM}}, \frac{m}{E_{\text{CM}}}\right) = \sigma\left(E_{\text{CM}}, 0\right) + \cdots \ . \tag{209}$$

For $E_{\text{CM}} \gg m \sim v$, the zeroth order terms dominates. The scattering occurs approximately as if all particles are massless and electroweak symmetry is "restored". In this case, we simply have $\sigma \sim 1/E_{\text{CM}}^2$ from dimension analysis, which is a boring featureless power law.

For collisions with E_{CM} around or below the electroweak scale, the E_{CM} dependence in σ will be interesting, featured with various Breit-Wigner peaks when E_{CM} crosses particle mass thresholds (see Section 4).

At very low energies $E_{\text{CM}} \lesssim$ GeV scale, interesting phenomenon occurs for the strong interaction $SU(3)_c$, where quarks and gluons confine into hadrons (proton, neutron, pion, etc). To describe strong interaction at low energies, we need effective theories that describe the new degrees of freedom (see Section 5).

3 Particle Decays



3.1 Z and W boson decays

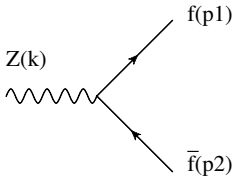
3.1.1 Unpolarized Z decay

In the SM, the Z-boson couples to all the fermions. Because of its large mass, the Z-boson can decay into all fermion-anti-fermion pairs, except for $t\bar{t}$. They comprise the leading decay modes. The Z-f-$\bar{f}$ interaction terms in the SM Lagrangian are

$$\mathcal{L}_{\rm int} = g_Z \bar{f}\gamma^\mu (g_V + g_A\gamma_5) f Z_\mu \ , \tag{210}$$

where $g_Z = e/(\sin\theta_W \cos\theta_W)$, $g_V = \frac{1}{2}T_{3L} - Q\sin^2\theta_W$, and $g_A = -\frac{1}{2}T_{3L}$.

The decay process $Z \to f\bar{f}$ is shown by the following Feynman diagram (time flows from left to right).



The corresponding decay amplitude is

$$i\mathcal{M} = g_Z \bar{u}(p_1, s_1)\gamma^\mu (g_V + g_A\gamma_5) v(p_2, s_2)\varepsilon_\mu(k, \lambda) \ . \tag{211}$$

For simplicity, we drop the arrow over the momenta, with the understanding that all particles are on-shell.

For unpolarized Z decay, we square the amplitude, sum over final state spins and colors, and average over the initial spin degrees of freedom for the Z-boson,

$$\begin{aligned}\overline{|\mathcal{M}|^2} &= \frac{1}{3}\sum_{s_1,s_2,\lambda} g_Z^2[\bar{u}(p_1,s_1)\gamma^\mu(g_V+g_A\gamma_5)v(p_2,s_2)]\varepsilon_\mu(k,\lambda) \\ &\qquad \times[\bar{v}(p_2,s_2)\gamma^\nu(g_V+g_A\gamma_5)u(p_1,s_1)]\varepsilon_{\nu *}(k,\lambda) \\ &= \frac{1}{3}g_Z^2\,\mathrm{Tr}\left[(\not{p}_1+m_f)\gamma^\mu(g_V+g_A\gamma_5)(\not{p}_2-m_f)\gamma^\nu(g_V+g_A\gamma_5)\right]\left(-g_{\mu\nu}+\frac{k_\mu k_\nu}{M_Z^2}\right)\ ,\end{aligned} \tag{212}$$

where in the second step, we used Eqs. (53) and (76).

To do the fermion trace, we need to know that $\mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma_5) = -4i\varepsilon^{\mu\nu\rho\sigma}$. If fewer γ matrices are taken trace with a γ_5, the result is zero. We also need to know the kinematics well, where $k = p_1 + p_2$, $k^2 = M_Z^2$, $p_1^2 = p_2^2 = m_f^2$, $p_1\cdot p_2 = \frac{1}{2}(M_Z^2 - 2m_f^2)$, $p_1\cdot k = p_2\cdot k = \frac{1}{2}M_Z^2$. The rest is just patience. We will eventually get

$$\overline{|\mathcal{M}|^2} = \frac{4N_c}{3}g_Z^2\left[(g_V^2+g_A^2)M_Z^2 + 2(g_V^2-2g_A^2)m_f^2\right]\ . \tag{213}$$

A color factor $N_c = 3$ is added in case Z decays into quarks. $N_c = 1$ for charged lepton and neutrino final states.

The unpolarized partial decay rate is then

$$\Gamma_{Z\to f\bar{f}} = \frac{1}{8\pi}\frac{|\vec{p}_{1\mathrm{cm}}|}{M_Z^2}\overline{|\mathcal{M}|^2} = \frac{N_c g_Z^2 M_Z}{12\pi}\left[(g_V^2+g_A^2)+2(g_V^2-2g_A^2)\frac{m_f^2}{M_Z^2}\right]\sqrt{1-\frac{4m_f^2}{M_Z^2}}\ , \tag{214}$$

where we used the decay momentum in center-of-mass frame,

$$|\vec{p}_{1\mathrm{cm}}| = \frac{M_Z}{2}\sqrt{1-\frac{4m_f^2}{M_Z^2}}\ . \tag{215}$$

In reality, the Z boson is much heavier than the heaviest fermion it can decay into, the bottom quark. Thus we can suppress the terms proportional to fermion mass in Γ, wen reach a simpler expression for the partial decay rate

$$\Gamma \simeq \frac{N_c(g_V^2+g_A^2)\sqrt{2}G_F M_Z^3}{3\pi}\ . \tag{216}$$

In the last step, we introduced the Fermi constant

$$G_F = \frac{\sqrt{2}g_2^2}{8M_W^2} = 1.16\times 10^{-5}\,\mathrm{GeV}^{-2}\ , \tag{217}$$

and used the tree-level relation $M_W/M_Z = \cos\theta_W$.

Let's go over each possible final state of Z decay:

- $f = u, c$. In this case, $T_{3L} = \frac{1}{2}$, $Q = \frac{2}{3}$. The value of $\sin^2\theta_W = 0.23$ is close to a quarter. This leads to

$$g_V^2 + g_A^2 \approx \frac{10}{144} \ . \tag{218}$$

- $f = d, s, b$. In this case, $T_{3L} = -\frac{1}{2}$, $Q = -\frac{1}{3}$. This leads to

$$g_V^2 + g_A^2 \approx \frac{13}{144} \ . \tag{219}$$

- $f = e, \mu, \tau$. In this case, $T_{3L} = -\frac{1}{2}$, $Q = -1$. This leads to

$$g_V^2 + g_A^2 \approx \frac{9}{144} \ . \tag{220}$$

- $f = \nu_e, \nu_\mu, \nu_\tau$. In this case, $T_{3L} = \frac{1}{2}$, $Q = 0$. This leads to

$$g_V^2 + g_A^2 \approx \frac{18}{144} \ . \tag{221}$$

The total decay width of the Z boson is

$$\Gamma_Z^{\rm total} \approx \frac{\sqrt{2} G_F M_Z^3}{3\pi} \left[2 \times 3 \times \frac{10}{144} + 3 \times 3 \times \frac{13}{144} + 3 \times \frac{9}{144} + 3 \times \frac{18}{144} \right] \approx 2.4\,{\rm GeV} \ . \tag{222}$$

This is a fairly large width. All the Z bosons will decay promptly after they are produced in particle collider experiments.

Experimentally, at e^+e^- colliders such as LEP, the Z-boson total decay width can be measured by studying the process $e^+e^- \to Z \to f\bar{f}$, and the invariant mass distribution of the final state $f, \bar{f}$. It works the best if the final states are e^+e^- or $\mu^+\mu^-$, whose four momenta can be most precisely measured. We will discuss the Breit-Wigner resonance in Section 4.

Another important concept when studying particle decays is the decay branching ratio into a particular type of final states. For the decay of Z, we can measure

- Hadronic decay mode: $Z \to q\bar{q}$, where q includes all types of quarks. The branching ratio is

$${\rm Br}_{\rm had} = \frac{\sum_q \Gamma_{Z\to q\bar{q}}}{\Gamma_Z^{\rm total}} \approx \frac{177}{258} \approx 70\% \ . \tag{223}$$

- Leptonic decay mode: $Z \to \ell^+\ell^-$, where $\ell = e, \mu, \tau$. The branching ratio is

$${\rm Br}_{\rm lep} = \frac{\sum_\ell \Gamma_{Z\to \ell^+\ell^-}}{\Gamma_Z^{\rm total}} \approx \frac{27}{258} \approx 10\% \ . \tag{224}$$

  Each charged lepton flavor has about 3% branching ratio.

- Invisible decay mode: $Z \to \nu\bar{\nu}$. Neutrinos are not visible to the detectors of collider experiments. The branching ratio is

$$\mathrm{Br}_{\mathrm{inv}} = \frac{\sum_\ell \Gamma_{Z\to\nu\bar{\nu}}}{\Gamma_Z^{\mathrm{total}}} \approx \frac{54}{258} \approx 20\% \ . \tag{225}$$

  Experimentally, we can tell the invisible decay occurs by comparing the expected number of produced Z bosons with those that have decayed visibly.

3.1.2 Polarized Z decay

For polarized Z decay, we can repeat the previous calculation up to Eq. (212), but in the last step, we do not sum over the index λ. For given λ, we work out the explicit form of the polarization vectors. Using Eq. (75), in the rest frame of the Z boson, we have

$$\begin{aligned} \varepsilon^{\mu,\lambda=\pm1} &= \frac{1}{\sqrt{2}}(0,1,\pm i,0) \ , \\ \varepsilon^{\mu,\lambda=0} &= (0,0,0,1) \ . \end{aligned} \tag{226}$$

The amplitude square for a given polarization is (neglecting the fermion masses for simplicity)

$$\begin{aligned} |\mathcal{M}_\lambda|^2 &= N_c g_Z^2 \mathrm{Tr}\left[(\not{p}_1 + m_f)\not{\varepsilon}^\lambda (g_V + g_A\gamma_5)(\not{p}_2 - m_f)\not{\varepsilon}^{\lambda *}(g_V + g_A\gamma_5)\right] \\ &\simeq N_c g_Z^2 \mathrm{Tr}\left[\not{p}_1\not{\varepsilon}^\lambda (g_V + g_A\gamma_5)\not{p}_2\not{\varepsilon}^{\lambda *}(g_V + g_A\gamma_5)\right] \\ &= N_c g_Z^2 (g_V^2 + g_A^2)\mathrm{Tr}\left[\not{p}_1\not{\varepsilon}^\lambda \not{p}_2\not{\varepsilon}^{\lambda *}\right] \\ &= 4N_c g_Z^2 (g_V^2 + g_A^2)\left[(p_1\cdot\varepsilon^\lambda)(p_2\cdot\varepsilon^{\lambda *}) - (p_1\cdot p_2)(\varepsilon^\lambda\cdot\varepsilon^{\lambda *}) + (p_1\cdot\varepsilon^{\lambda *})(p_2\cdot\varepsilon^\lambda)\right] \\ &= 4N_c g_Z^2 (g_V^2 + g_A^2)\left[(p_1\cdot\varepsilon^\lambda)(p_2\cdot\varepsilon^{\lambda *}) + \frac{1}{2}M_Z^2 + (p_1\cdot\varepsilon^{\lambda *})(p_2\cdot\varepsilon^\lambda)\right] \ . \end{aligned} \tag{227}$$

Next, we write down the four momenta $p_{1,2}$ explicitly as well. Assuming f travels in the (θ,ϕ) direction, and $\bar{f}$ in the opposite direction in the center-of-mass frame,

$$\begin{aligned} p_1^\mu &= \frac{M_Z}{2}(1, \sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \ , \\ p_2^\mu &= \frac{M_Z}{2}(1, -\sin\theta\cos\phi, -\sin\theta\sin\phi, -\cos\theta) \ . \end{aligned} \tag{228}$$

Together with Eq. (226), we find

$$|\mathcal{M}_\lambda|^2 = 2N_c g_Z^2 (g_V^2 + g_A^2) M_Z^2 \times \begin{cases} \frac{1}{2}(1+\cos^2\theta) \ , & \lambda = \pm 1 \ , \\ 1 - \cos^2\theta \ , & \lambda = 0 \ . \end{cases} \tag{229}$$

This leads to several important findings. First, the average of $|\mathcal{M}_{0,\pm1}|^2$ has no θ dependence, and reproduces the unpolarized result. Eq. (213).

Second, after integrating over the final state phase space, we obtain the same decay rate for all polarization cases,

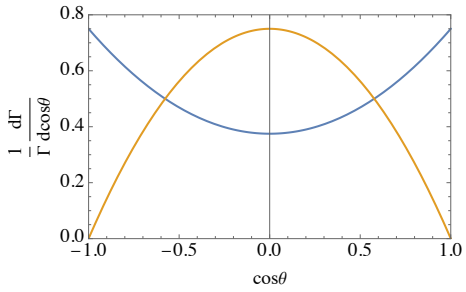
$$\Gamma_{\lambda=0} = \Gamma_{\lambda=\pm1} = \frac{N_c g_Z^2 (g_V^2 + g_A^2) M_Z}{12\pi} \ , \tag{230}$$

the same as Eq. (214) in the massless fermion limit. This is expected, because they are the same particle Z.

Third, the differential decay rate with respect to $\cos\theta$ is a measurable quantity experimentally. For the three cases, we have

$$\frac{1}{\Gamma_\lambda}\frac{d\Gamma_\lambda}{d\cos\theta} = \frac{3}{4} \times \begin{cases} \frac{1}{2}(1+\cos^2\theta) \ , & \lambda = \pm 1 \ , \\ 1 - \cos^2\theta \ , & \lambda = 0 \ . \end{cases} \tag{231}$$

These angular distribution functions are depicted in the following figure.



3.1.3 W decay

Similar to the Z boson, the decays of charged W boson is also dominated by fermion-anti-fermion final states. Using Eq. (177), the relevant interactions can be written as

$$\mathcal{L}_{CC} = \frac{g_2}{\sqrt{2}} W_\mu^+ \left(V_{ij}^{\rm CKM} \bar{u}_i \gamma^\mu \mathbb{P}_L d_j + \bar{\nu}_i \gamma^\mu \mathbb{P}_L d_j \ell_i \right) + \text{h.c.} \tag{232}$$

where the fermions are already in the mass basis. The W boson cannot decay into top quark because the latter is too heavy. The rest fermions are all must lighter than the W. By neglecting those masses, we can repeat the same calculation as for Z decays, but with the replacement

$$g_Z \to \frac{g_2}{\sqrt{2}} V_{ij}^{\rm CKM} \ , \quad g_V \to \frac{1}{2} \ , \quad g_A \to -\frac{1}{2} \ , \tag{233}$$

for decaying to quarks. For leptonic decays, the CKM matrix element is not there. Applying this replacement to Eq. (214), we obtain

$$\Gamma_{W^+ \to u_i \bar{d}_j} \simeq \frac{|V_{ij}^{\rm CKM}|^2 N_c g_2^2 M_Z}{48\pi} \ , \quad \Gamma_{W^+ \to \nu_i \bar{\ell}_i} \simeq \frac{g_2^2 M_Z}{48\pi} \ . \tag{234}$$

Summing over the final state multiplicity, we obtain

$$\Gamma_W^{\rm total} = \left(3 \sum_{i=1}^{2} \sum_{j=1}^{3} |V_{ij}^{\rm CKM}|^2 + 3 \right) \frac{g_2^2 M_Z}{48\pi} = \frac{3\sqrt{2} G_F M_W^3}{4\pi} = 2.04\,{\rm GeV}. \quad (235)$$

This is comparable to $\Gamma_Z^{\rm total}$. Both W and Z decay promptly at colliders.

The hadronic branching ratio of W decay is 67%. The leptonic branching ratio is 11% for every flavor (e, μ, τ).

3.2 Higgs boson decay (and Higgs physics in brief)

The Higgs boson has several ways to decay in the SM that contribute to its total decay width.

First, it can decay into fermion-anti-fermion pair, except for the top quark (too heavy) and neutrinos (massless). Importantly, Eq. (175) tells that the Higgs-fermion coupling is proportional the fermion mass. It implies that Higgs prefers to decay into heavier fermions than lighter ones. Without details, this decay rate is

$$\Gamma_{h \to f\bar{f}} = \frac{\sqrt{2} N_c G_F m_f^2 m_h}{8\pi} \left(1 - \frac{4m_f^2}{m_h^2} \right)^{3/2} . \quad (236)$$

Here the factor $(\dots)^{3/2}$ indicates this decay occurs in the P-wave. Higgs boson is a parity even scalar and a fermion-anti-fermion pair has parity $-(-1)^L$. Due to parity conservation of the Higgs-fermion interaction, the orbital angular momentum L must be an odd integer.

The heaviest fermion the Higgs boson can decay into is the bottom quark, which is much lighter than the Higgs. This means we can always approximate the $(\dots)^{3/2}$ factor as 1. As a result $\Gamma_{h \to f\bar{f}} \propto m_f^2$. The contribution from $h \to b\bar{b}$ decay to the Higgs total width is

$$\Gamma_{h \to b\bar{b}} \simeq \frac{3\sqrt{2} G_F m_b^2 m_h}{8\pi} \simeq 4.3\,{\rm MeV} , \quad (237)$$

where we used $v^2 = (\sqrt{2} G_F)^{-1}$. The corresponding branching ratio is 58%. This partial decay rate is thousand times lower than the Z, W decay width, because the decay rate involves a small Yukawa coupling parameter, $y_b = \sqrt{2} m_b / v = 0.024$. For Z, W decay, the weak interaction gauge coupling $g_2 \simeq 0.65$ is order one. The bottom quarks from Higgs boson decay is very hard to detect at the LHC, mainly due to the huge background of $b\bar{b}$ production from strong interaction. Jet substructure techniques have recently been proposed to search for boosted $h \to b\bar{b}$ decay (see `https://arxiv.org/pdf/2006.13251.pdf`), but the significance has not reached 5σ yet.

Another important fermionic decay of the Higgs is to $\mu^+\mu^-$. Because the muon is light, this branching ratio is much lower than $b\bar{b}$, by a extra factor of $(m_\mu/m_b)^2 \approx 6 \times 10^{-4}$. One of the task of the high-luminosity of LHC is to discover the Higgs boson through this channel and test the SM prediction.

Other Higgs boson decay modes include $h \to VV^* \to V\bar{f}f$ ($V = Z, W$) which is a three-body decay, and $h \to gg$ which occurs at loop level (top quark in the loop). They contribute to the other 42% of branching ratio.

The 2012 Higgs boson discovery announcement was based on two decay channels, $h \to \gamma$ and $h \to ZZ^* \to 4\mu$. Both have very small branching ratio but are very clean signals.

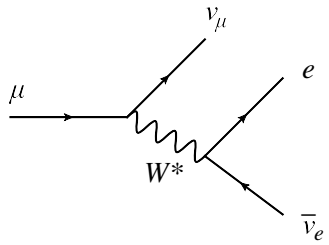
At hadron colliders (e.g. the LHC), the Higgs is mainly produced in the gluon fusion channel $gg \to h$. AT e^+e^- colliders (e.g. future Higgs factory), the dominant production is associate production, $q\bar{q}$ (or e^+e^-) $\to V^* \to Vh$. A great textbook to study the Higgs physics in depth is the *Higgs Hunter's Guide*, `https://catalogue.library.cern/literature/r8p7d-xgg53`.

3.3 Muon decay (and heavy quark decays)

In the SM, the forces carriers that the muon interacts with are γ, Z, h, W. Among them the γ, Z, h couplings always involve two muons and thus cannot make muon decay. The only way for muon to decay is

$$\mu^- \to \nu_\mu e^- \bar{\nu}_e \ , \tag{238}$$

which occurs through charged current interaction with an off-shell W boson exchange, as shown by the Feynman diagram below. The relevant interacting Lagrangian is the second term in Eq. (232).



Here we work out the calculation details of unpolarized muon decay rate. The momentum assignments we will use are $\mu(p_0) \to \nu_\mu(p_1) + e(p_2) + \bar{\nu}_e(p_3)$. The decay amplitude is

$$\begin{aligned} i\mathcal{M} = \left(\frac{ig_2}{\sqrt{2}}\right)^2 \Big[\bar{u}(p_1, s_1)\gamma^\mu \mathbb{P}_L u(p_0, s_0)\Big] \Big[\bar{u}(p_2, s_2)\gamma^\nu \mathbb{P}_L v(p_3, s_3)\Big] \\ \times \frac{i}{(p_2+p_3)^2 - M_W^2}\left(-g_{\mu\nu} + \frac{(p_0-p_1)_\mu (p_2+p_3)_\nu}{M_W^2}\right) \ , \end{aligned} \tag{239}$$

where in the last bracket, we used $p_0 - p_1 = p_2 + p_3$.

The first thing we want to argue is that in the last bracket, the second term inversely proportional to M_W^2 is negligible. This can be seen by contracting $(p_0 - p_1)_\mu$ and $(p_2 + p_3)_\nu$ with the fermion currents and apply their equations

of motion. Because neutrinos are massless, the only surviving terms are proportional to m_μ^2/M_W^2 or m_e^2/M_W^2. Both are much smaller than one. For this reason, we will only keep the contribution from the $g_{\mu\nu}$ term in below.

Next, we square the amplitude, and sum over final state fermions spins, and average over the initial state muon spins (assuming the muon is unpolarized). This leads to

$$\overline{|\mathcal{M}|^2} = \frac{1}{2}\frac{g_2^4}{4[(p_2+p_3)^2-M_W^2]^2}\mathrm{Tr}\left[\not{p}_1\gamma^\mu\mathbb{P}_L(\not{p}_0+m_\nu)\gamma^\nu\mathbb{P}_L\right]\mathrm{Tr}\left[(\not{p}_2+m_e)\gamma_\mu\mathbb{P}_L\not{p}_3\gamma_\nu\mathbb{P}_L\right] \ . \tag{240}$$

To complete the traces, we use the identities

$$\begin{aligned}\mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) &= 4(g^{\mu\nu}g^{\rho\sigma}-g^{\mu\rho}g^{\nu\sigma}+g^{\mu\sigma}g^{\nu\rho}) \ , \\ \mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma_5) &= -4i\varepsilon^{\mu\nu\rho\sigma} \ ,\end{aligned} \tag{241}$$

to obtain

$$\begin{aligned}\mathrm{Tr}\left[\not{p}_1\gamma^\mu\mathbb{P}_L(\not{p}_0+m_\nu)\gamma^\nu\mathbb{P}_L\right] &= 2(p_1^\mu p_0^\nu+p_1^\nu p_0^\mu-p_1\cdot p_0 g^{\mu\nu})+2i\varepsilon^{\alpha\mu\beta\nu}p_{1\alpha}p_{0\beta} \ , \\ \mathrm{Tr}\left[(\not{p}_2+m_e)\gamma_\mu\mathbb{P}_L\not{p}_3\gamma_\nu\mathbb{P}_L\right] &= 2(p_{2\mu}p_{3\nu}+p_{2\nu}p_{3\mu}-p_2\cdot p_3 g_{\mu\nu})+2i\varepsilon_{\alpha'\mu\beta'\nu}p_2^{\alpha'}p_3^{\beta'} \ .\end{aligned} \tag{242}$$

By further using the identity $\varepsilon^{\alpha\mu\beta\nu}\varepsilon_{\alpha'\mu\beta'\nu} = -2(g^\alpha_{\alpha'}g^\beta_{\beta'}-g^\alpha_{\beta'}g^\beta_{\alpha'})$, we obtain

$$\overline{|\mathcal{M}|^2} = \frac{2g_2^4}{[(p_2+p_3)^2-M_W^2]^2}(p_1\cdot p_2)(p_0\cdot p_3) \ . \tag{243}$$

For the above three-body decay, we can rewrite the momentum products as (neglecting the electron mass hereafter)

$$\begin{aligned}p_1\cdot p_2 &= \frac{(p_1+p_2)^2-p_1^2-p_2^2}{2} = \frac{m_{12}^2}{2} \ , \\ p_0\cdot p_3 &= \frac{p_0^2+p_3^2-(p_0-p_3)^2}{2} = \frac{p_0^2+p_3^2-(p_1+p_2)^2}{2} = \frac{m_\mu^2-m_{12}^2}{2} \ ,\end{aligned} \tag{244}$$

where $m_{12}^2 \equiv (p_1+p_2)^2$ is the invariant mass square of ν_μ and e.

We could also introduce two other invariant masses, $m_{23}^2 \equiv (p_2+p_3)^2$ for the invariant mass square of e and $\bar{\nu}_e$, and $m_{13}^2 = (p_1+p_3)^2$ for the invariant mass of ν_μ and $\bar{\nu}_e$. These three kinetic variables are not independent, they satisfy

$$m_{12}^2+m_{23}^2+m_{13}^2 = p_0^2+p_1^2+p_2^2+p_3^2 = m_\mu^2 \ . \tag{245}$$

Putting them together, we get

$$\overline{|\mathcal{M}|^2} = \frac{g_2^4}{2[m_{23}^2-M_W^2]^2}m_{12}^2(m_\mu^2-m_{12}^2) \simeq \frac{g_2^4}{2M_W^4}m_{12}^2(m_\mu^2-m_{12}^2) = 16G_F^2m_{12}^2(m_\mu^2-m_{12}^2) \ . \tag{246}$$

In the second step, we make use of the mass hierarchies $m_{23}^2 < m_\mu^2 \ll M_W^2$, and keep the leading term (the large M_W expansion).

To calculate the three-body decay rate, we use the simplified final state phase space integral provided in the PDG review,

$$\Gamma = \frac{1}{(2\pi)^3}\frac{1}{32m_\mu^3}\int dm_{12}^2 \int dm_{23}^3 \overline{|\mathcal{M}|^2} \ . \tag{247}$$

The integral limits are

$$(m_1+m_2)^2 \le m_{12}^2 \le (m_\mu - m_3)^2 \ , \tag{248}$$

where the lower bound corresponds to particles 1,2 travel in parallel and recoil against particle 3, and upper bound corresponds to 1,2 recoil against each other whereas 3 is at rest, and

$$\begin{aligned}(m_{23}^2)_{\rm min}^{\rm max} &= (E_2^* + E_3^*)^2 - \left(\sqrt{E_2^{*2}-m_2^2} \mp \sqrt{E_3^{*2}-m_3^2}\right)^2 \ , \\ E_2^* &= \frac{m_{12}^2 - m_1^2 + m_2^2}{2m_{12}} \ , \quad E_3^* = \frac{m_\mu^2 - m_{12}^2 - m_3^2}{2m_{12}} \ .\end{aligned} \tag{249}$$

For massless final state particles $1,2,3$ considered here,

$$0 \le m_{12}^2 \le m_\mu^2 \ , \quad 0 \le m_{23}^2 \le 4E_2^*E_3^* = m_\mu^2 - m_{12}^2 \ . \tag{250}$$

The decay rate can then be written as

$$\begin{aligned}\Gamma &= \frac{1}{256\pi^3 m_\mu^3}\int_0^{m_\mu^2} dm_{12}^2 \int_0^{m_\mu^2 - m_{12}^2} dm_{23}^3 \overline{|\mathcal{M}|^2} \\ &= \frac{G_F^2 m_\mu^5}{16\pi^3}\int_0^1 dx \int_0^{1-x} dy x(1-x) \ .\end{aligned} \tag{251}$$

The dimensionless integrals over x and y gives $\frac{1}{12}$. Thus the leading order muon decay rate is

$$\Gamma = \frac{G_F^2 m_\mu^5}{192\pi^3} \simeq 2.8 \times 10^{-19}\,{\rm GeV} \ . \tag{252}$$

The corresponding lifetime is

$$\tau = \frac{0.658 \times 10^{-24}\,{\rm GeV\,sec}}{\Gamma} \simeq 2 \times 10^{-6}\,{\rm sec} \ . \tag{253}$$

If the muon is produced relativistically in a experiment, it can travel at least a distance of $\sim c\tau = 600$ meters. This is a very large distance compared to the size of a typical detector. Therefore, in collider experiments, muons are treated as stable final state particles. (It is electrically charged, so not difficult to detect.)

Another useful picture to have in mind is the production of high energy particles due to cosmic ray bombarding on the earth's atmosphere. With a

boost factor of 100 (corresponding to 10 GeV in energy), the muon can travel a distance of $\gamma c\tau \sim 60\,\text{km}$, that is roughly the thickness of the atmosphere. This explains why we can see high energy particles by building a bubble chamber in the ground.

It is not as easy to stop energetic muons in matter as electrons. Each collision of muon with an atomic electron only loses a small fraction of its energy, because $m_\mu \gg m_e$. If we really need to stop the muons (such as in beam dump experiments), we need to put a very thick piece of lead (or earth), along the direction the muons travel.

I will not do polarized muon decay. It is a good exercise for you.

Before proceeding, we generalize the CC induced muon decay to other particles, including the τ lepton and heavy quarks (b, c, s). If the SM has no CC weak interactions, all these particles would be stable. In reality, they all decay via an off-shell W boson exchange (similar Feynman diagrams to muon decay). The only exception is the top quark, which is heavier than the W boson. The top quark simply decays as $t \to b + W^+$, with a large decay width (1.32 GeV).

3.4 Integrating out the W boson

In the $M_W \to \infty$ limit, the muon decay amplitude Eq. (239) can be written as

$$i\mathcal{M} \simeq -i2\sqrt{2}G_F \left[\bar{u}(p_1, s_1)\gamma^\mu \mathbb{P}_L u(p_0, s_0)\right] \left[\bar{u}(p_2, s_2)\gamma_\mu \mathbb{P}_L v(p_3, s_3)\right] \ . \quad (254)$$

This corresponds to an effective four-fermion interacting Lagrangian

$$\mathcal{L}_{\text{eff}} \simeq -2\sqrt{2}G_F \left[\bar{\nu}_\mu \gamma^\mu \mathbb{P}_L \mu\right] \left[\bar{e}\gamma_\mu \mathbb{P}_L \nu_e\right] \ , \quad (255)$$

where we used the identity $2\sqrt{2}G_F = g_2^2/(2M_W^2)$. Such a Lagrangian can be formally obtained through the procedure of integrating out the W boson. The relevant pieces from the original Standard Model Lagrangian are

$$\mathcal{L} = -\frac{1}{4}W^{+\mu\nu}W^-_{\mu\nu} + M_W^2 W^+_\mu W^{-\mu} + \frac{g}{\sqrt{2}}W^+_\mu \left(\bar{u}\gamma^\mu d + \bar{\nu}\gamma^\mu e\right) + \text{h.c.} \ , \quad (256)$$

where we keep only one fermion generation for simplicity. When the W boson is heavy, we are can drop its kinetic term. As a result, the W field becomes a Lagrangian multiplier. Its equation of motion becomes

$$W^{-\mu} = \frac{g}{\sqrt{2}M_W^2}\left(\bar{u}\gamma^\mu d + \bar{\nu}\gamma^\mu e\right) \ , \quad W^{+\mu} = \frac{g}{\sqrt{2}M_W^2}\left(\bar{d}\gamma^\mu u + \bar{e}\gamma^\mu \nu\right) \ . \quad (257)$$

Plugging this back to the Lagrangian (with the W kinetic term already removed), we get

$$\mathcal{L}_{\text{eff}} \simeq -2\sqrt{2}G_F \left[\bar{\nu}\gamma^\mu \mathbb{P}_L e + \bar{u}\gamma^\mu \mathbb{P}_L d\right] \left[\bar{e}\gamma_\mu \mathbb{P}_L \nu_e + \bar{d}\gamma^\mu \mathbb{P}_L u\right] \ , \quad (258)$$

which includes the effective Lagrangian Eq. (255). As a bit more details, a cancellation will occur between two of the terms in the Lagrangian, but the third term (the h.c. part) survives. Eq. (258) also contains a term that can induce the beta decay of neutron.

3.5 Polarized neutron decay

As the last exercise of this section, we explore polarized neutron decay $n \to pe\bar{\nu}_e$ through weak interaction, and try to appreciate how parity violation manifests experimentally.

The effective Lagrangian for neutron decay is

$$\mathcal{L}_{\text{eff}} = -2\sqrt{2}G_F \left[\bar{p}\gamma^\mu \mathbb{P}_L n\right] \left[\bar{e}\gamma_\mu \mathbb{P}_L \nu_e\right] \ . \tag{259}$$

This is an approximation, which we will use for discussions in this subsection. In reality, the $1-\gamma_5$ structure on the hadron side is given by $g_V - g_A\gamma_5$. At low energies, $g_V \simeq 1$, $g_A \simeq 1.2$.

Consider a neutron at rest with in a fixed spin state spin. Using Eq. (50), the helicity spinor for such a neutron is

$$u_n = \sqrt{m_n}\begin{pmatrix}\chi_s\\ \chi_s\end{pmatrix} \ , \quad \text{where } \chi_s = a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}1\\0\end{pmatrix} \ , \tag{260}$$

where $|a^2| + |b|^2 = 1$. The matrix element square for neutron decay is

$$\begin{aligned}
|\mathcal{M}|^2 &= 8G_F^2 \left[\bar{u}_n\gamma^\mu \mathbb{P}_L(\not{p}_p + m_p)\gamma^\nu \mathbb{P}_L u_n\right] \mathrm{Tr}(\not{p}_e\gamma_\mu \mathbb{P}_L \not{p}_{\bar{\nu}}\gamma_\nu \mathbb{P}_L) \\
&\simeq 8G_F^2 m_p \left[\bar{u}_n\gamma^\mu\gamma^0\gamma^\nu \mathbb{P}_L u_n\right] \mathrm{Tr}(\not{p}_e\gamma_\mu \not{p}_{\bar{\nu}}\gamma_\nu \mathbb{P}_L) \\
&= 8G_F^2 m_p \left[\chi_s^\dagger(\sigma^\mu + \bar{\sigma}^\mu)\bar{\sigma}^\nu \chi_s\right] \mathrm{Tr}(\not{p}_e\gamma_\mu \not{p}_{\bar{\nu}}\gamma_\nu \mathbb{P}_L) \\
&= 16G_F^2 m_p \delta_0^\mu \left[\chi_s^\dagger \bar{\sigma}^\nu \chi_s\right] \mathrm{Tr}(\not{p}_e\gamma_\mu \not{p}_{\bar{\nu}}\gamma_\nu \mathbb{P}_L) \ .
\end{aligned} \tag{261}$$

We have summed over the spins of final state fermions, but not for the initial state neutron which is polarized. We make the approximation that the final state proton is still non-relativistic, thus $\not{p}_p \simeq m_p\gamma^0$ in the second step. We also introduce $\sigma^\mu \equiv (1, \vec{\sigma})$, $\bar{\sigma}^\mu \equiv (1, -\vec{\sigma})$. The factor $(\sigma^\mu + \bar{\sigma}^\mu)$ forces $\mu = 0$ which explains the last step.

To account for the information of neutron polarization, we will pick the 3-vector component out of the $\bar{\sigma}^\nu$, i.e., letting $\nu = j$.

The trace can be simply calculated (with $\mu = 0$ and $\nu = j$),

$$\mathrm{Tr}(\not{p}_e\gamma_0 \not{p}_{\bar{\nu}}\gamma_j \mathbb{P}_L) = 2E_e(-p_{\bar{\nu}}^j) + 2E_{\bar{\nu}}(-p_e^j) - 2p_e \cdot p_{\bar{\nu}} g_{0j} + \cdots \ . \tag{262}$$

The $\cdots$ represents the contribution with a γ_5 in the trace which produces a term proportional to $(\vec{p}_e \times \vec{p}_{\bar{\nu}})^j$. It is relevant for CP violation, but not the parity violation we are interested here.

Keeping the terms in the amplitude square that knows the neutron spin, we have

$$|\mathcal{M}|^2 \simeq 32G_F^2 m_p \chi_s^\dagger \left[E_e\vec{\sigma}\cdot\vec{p}_{\bar{\nu}} + E_{\bar{\nu}}\vec{\sigma}\cdot\vec{p}_e\right]\chi_s + (\text{spin independent terms}) \ . \tag{263}$$

Clearly, the $\vec{\sigma}\cdot\vec{p}_e$ term makes a different for electron traveling along or opposite to the neutron spin direction. Recall that $|\mathcal{M}|^2$ contributes to the differential

decay rate with respect to the angular variables, e.g., $d\Gamma/d\cos\theta(\vec{p}_e, \vec{s}_n) \propto |\mathcal{M}|^2$. The presence of $\vec{\sigma}\cdot\vec{p}$ terms clearly indicate parity violation, because under parity transformation 3-momentum $\vec{p}$ is odd whereas spin $\vec{s}_n = \langle\vec{\sigma}/2\rangle$ stays invariant.

Looking back, had the weak interaction preserved parity, and the above calculation are repeated without $\mathbb{P}_L$, we would get

$$\bar{u}_n\gamma^\mu\gamma^0\gamma^\nu u_n = \chi_s^\dagger(\sigma^\mu + \bar{\sigma}^\mu)(\sigma^\nu + \bar{\sigma}^\nu)\chi_s = 4\chi_s^\dagger\chi_s\delta_0^\mu\delta_0^\nu \ . \tag{264}$$

No information about the neutron spin will be left over.

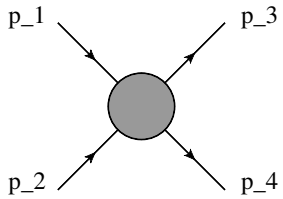
Evidence for parity violation in weak interactions was first discovered experimentally using the beta decay of cobalt-60 in the Wu experiment, where it was found that the most of the emitted electrons favored a direction opposite to that of the nuclear spin. `https://en.wikipedia.org/wiki/Wu_experiment`

A similar idea has also been applied to measure the polarization of the top quark experimentally, using the charged leptons in its decay final states.

4 Particle Collisions

4.1 Mandelstam variables

In this subsection and next, we focus on $2 \to 2$ scattering processes, represented by the diagram below.



The squared scattering amplitude of any process $|\mathcal{M}|^2(p_1, p_2, p_3, p_4)$ depends on the initial and final state momenta but is a Lorentz scalar, thus it must depend on the scalar products on these momenta. Using energy-momentum conservation, $p_4 = p_1 + p_2 - p_3$. With the three independent momenta, we can simply square each of them which gives particle masses. Nontrivial scalar products occur between different momenta, $p_1 \cdot p_2$, $p_1 \cdot p_3$, $p_2 \cdot p_3$. They can be reorganized to define the Mandelstam kinematic variables

$$\begin{aligned} s &\equiv (p_1 + p_2)^2 = m_1^2 + m_2^2 + 2p_1 \cdot p_2 \ , \\ t &\equiv (p_1 - p_3)^2 = m_1^2 + m_3^2 - 2p_1 \cdot p_3 \ , \\ u &\equiv (p_2 - p_3)^2 = m_2^2 + m_3^2 - 2p_2 \cdot p_3 \ . \end{aligned} \tag{265}$$

By noting that

$$\begin{aligned} &s + t + u = 2m_1^2 + 2m_2^2 + 2m_3^2 + 2(p_1 \cdot p_2 - p_1 \cdot p_3 - p_2 \cdot p_3) \ , \\ &p_4^2 = m_4^2 = (p_1 + p_2 - p_3)^2 = m_1^2 + m_2^2 + m_3^2 + 2(p_1 \cdot p_2 - p_1 \cdot p_3 - p_2 \cdot p_3) \ , \\ &\qquad \Rightarrow s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2 \ , \end{aligned} \tag{266}$$

we conclude that only two are independent among the s, t, u variables. The amplitude square is often written as a function of s and t.

In practice, in calculations we often work in the center-of-mass (CM) frame. Recall the cross section defined in Eq. (206) is Lorentz invariant. In this case, we have $\vec{p}_{1\text{cm}} = -\vec{p}_{2\text{cm}}$, and $s = (E_{1\text{cm}} + E_{2\text{cm}})^2$. We can also derive

$$\begin{aligned} &|\vec{p}_{1\text{cm}}| = |\vec{p}_{2\text{cm}}| = \frac{1}{2\sqrt{s}} \sqrt{(s - (m_1 + m_2)^2)\,(s - (m_1 - m_2)^2)} \ , \\ &E_{1\text{cm}} = \frac{s + m_1^2 - m_2^2}{2\sqrt{s}} \ , \quad E_{2\text{cm}} = \frac{s + m_2^2 - m_1^2}{2\sqrt{s}} \ . \end{aligned} \tag{267}$$

Similarly, we can find $|\vec{p}_{3\text{cm}}|, |\vec{p}_{4\text{cm}}|, E_{3\text{cm}}, E_{4\text{cm}}$ with the replacement $1, 2 \to 3, 4$ in the above equation.

In the CM frame, the t variable has a one-to-one correspondance to the scattering angle

$$t = m_1^2 + m_3^2 - 2E_{1\mathrm{cm}}E_{3\mathrm{cm}} + 2|\vec{p}_{1\mathrm{cm}}||\vec{p}_{3\mathrm{cm}}|\cos\theta \ . \tag{268}$$

The allowed range of t can then be worked out for $0 \leq \theta \leq \pi$. In high energy collision processes, we often deals with ultra-relativistic initial and final state particles. In the limit where all masses are negligible, it is straightforward to check that

$$-s \leq t \leq 0. \tag{269}$$

t is always negative, and so is u.

Finally, the total cross section can be obtained by integrating over the final state phase space, including the scattering angle, or t. As a result, it is only a function of s,

$$\sigma(s) \ . \tag{270}$$

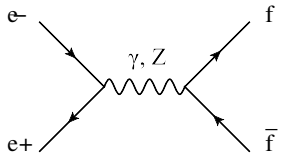
In general, the interaction cross section is different from the geometric cross section. For a fundamental particle, the latter is related to its de Broglie or Compton wavelength.

All discussions in this subsection are model independent.

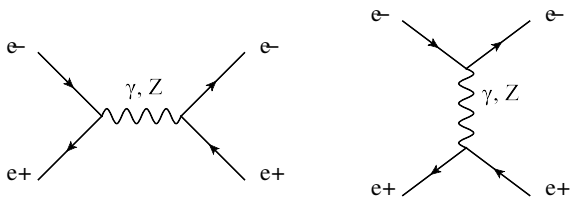
4.2 Lepton colliders

We first consider lepton collider physics, where e^+e^- annihilate onto fermion-anti-fermion pair, denoted by $f\bar{f}$. Depending on the identity of f, there are three classes of possible Feynman diagrams.

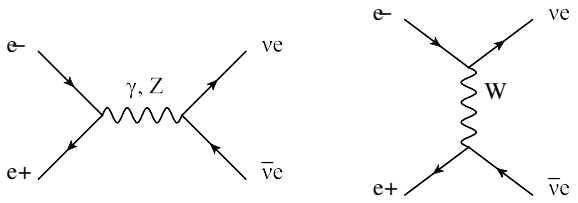
1) For $f \neq e$ or ν_e, the annihilation occurs only via s-channel photon and Z-boson exchange. (My time always goes from left to right.)



2) For $f = e$, the annihilation can also occur via t-channel photon and Z-boson exchange, in addition to the s-channel.



3) For $f = \nu_e$, the annihilation can also occur via t-channel W-boson exchange, in addition to the s-channel photon and Z-boson exchange.



In general, all diagrams contribute coherently at the scattering amplitude level, and they can interfere with each other.

For the rest of this subsection, we will consider f is the muon or quarks. In this case, only the s-channel photon and Z diagrams contribute. The scattering amplitude is

$$\begin{aligned}
i\mathcal{M} &= (ie)^2 Q_f [\bar{v}(p_2,s_2)\gamma^\mu u(p_1,s_1)] \frac{-ig_{\mu\nu}}{s} [\bar{u}(p_3,s_3)\gamma^\nu v(p_4,s_4)] \\
&\quad + (ig_Z)^2 [\bar{v}(p_2,s_2)\gamma^\mu (g_{eL}\mathbb{P}_L + g_{eR}\mathbb{P}_R) u(p_1,s_1)] \frac{i\left(-g_{\mu\nu} + \frac{(p_1+p_2)_\mu (p_1+p_2)_\nu}{M_Z^2}\right)}{s - M_Z^2} \\
&\qquad\qquad \times [\bar{u}(p_3,s_3)\gamma^\nu (g_{fL}\mathbb{P}_L + g_{fR}\mathbb{P}_R) v(p_4,s_4)] \\
&= \frac{ie^2 Q_f}{s} [\bar{v}(p_2,s_2)\gamma^\mu u(p_1,s_1)][\bar{u}(p_3,s_3)\gamma_\mu v(p_4,s_4)] \\
&\quad + \frac{ig_Z^2}{s - M_Z^2} [\bar{v}(p_2,s_2)\gamma^\mu (g_{eL}\mathbb{P}_L + g_{eR}\mathbb{P}_R) u(p_1,s_1)][\bar{u}(p_3,s_3)\gamma_\mu (g_{fL}\mathbb{P}_L + g_{fR}\mathbb{P}_R) v(p_4,s_4)] \ ,
\end{aligned} \tag{271}$$

where in the second step, we applied the equations of motion for the electron and muon spinors and neglected terms suppressed by $m_e m_\nu / M_Z^2$ (similar to what we did in the muon decay exercise). Using Eq. (180), the Z couplings are $g_Z = g_2/\cos\theta_W$, $g_L = T_{3L} - Q\sin^2\sin^2\theta_W$, and $g_R = -Q\sin^2\sin^2\theta_W$.

Next, we will consider special energy ranges of the scattering where only one of the two amplitudes dominates the scattering process.

4.2.1 e^+e^- annihilation at low energies

We first consider the CM energy of scattering in the regime $m_e, m_\mu, m_q \ll s \ll M_Z^2$. This allows us to drop all light fermion masses in the initial and final states, and drop the terms inversely proportional to the Z-boson mass. Indeed, comparing the prefactors in the two terms of Eq. (271), we find

$$\frac{e^2 Q_f}{s} \gg \frac{g_Z^2}{M_Z^2} \ . \tag{272}$$

This suggests that the photon exchange contribution dominates over that of Z exchange. The leading amplitude can be approximated as

$$i\mathcal{M} = \frac{ie^2 Q_f}{s} [\bar{v}(p_2,s_2)\gamma^\mu u(p_1,s_1)][\bar{u}(p_3,s_3)\gamma_\mu v(p_4,s_4)] \ . \tag{273}$$

The corresponding squared amplitude for unpolarized electron and position beams is

$$\begin{aligned}\overline{|\mathcal{M}|^2} &= \frac{1}{4}\frac{e^4Q_f^2}{s^2}\mathrm{Tr}[\not{p}_2\gamma^\mu\not{p}_1\gamma^\nu]\mathrm{Tr}[\not{p}_3\gamma^\mu\not{p}_4\gamma^\nu] \\ &= \frac{8e^4Q_f^2}{s^2}\left[(p_2\cdot p_3)(p_1\cdot p_4)+(p_1\cdot p_3)(p_2\cdot p_4)\right] \\ &= \frac{2e^4Q_f^2}{s^2}(u^2+t^2) \\ &\simeq \frac{2e^4Q_f^2}{s^2}\left[(s+t)^2+t^2\right] \ .\end{aligned} \tag{274}$$

In the last step, we used the relation $s+t+u\approx 0$ here.

The differential cross section is

$$\frac{d\sigma}{dt} = \frac{1}{64\pi s}\frac{1}{|\vec{p}_{1\mathrm{cm}}|^2}\overline{|\mathcal{M}|^2} \ , \tag{275}$$

where $|\vec{p}_{1\mathrm{cm}}|$ is given by Eq. (267). and the general range of t is

$$t_{\mathrm{min}}^{\mathrm{max}} = \left(\frac{m_1^2-m_2^2-m_3^2+m_4^2}{2\sqrt{s}}\right)^2 - \left(|\vec{p}_{1\mathrm{cm}}|\mp|\vec{p}_{3\mathrm{cm}}|\right)^2 \ . \tag{276}$$

Here for massless fermions, we have $|\vec{p}_{1\mathrm{cm}}| = \sqrt{s}/2$, and $-s\le t\le 0$. The integral can be readily done,

$$\sigma^{(\gamma\,\mathrm{exchange})}_{e^+e^-\to f\bar{f}}(s) = \frac{4\pi\alpha^2Q_f^2N_c}{3s} \ , \tag{277}$$

where the color factor $N_c=1$ for $f=\mu$, and $N_c=3$ for $f=$ quarks.

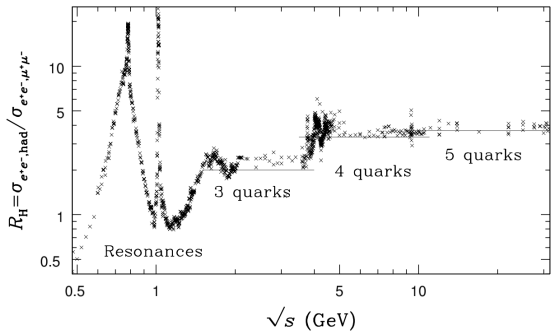
For e^+e^- collider running at CM energy $\sqrt{s}$, the annihilation will produce all light fermions with mass $m_f<\sqrt{s}/2$. It is useful to define an observable called the R factor, as the ratio between two cross sections. The hadronic R factor is defined as

$$R_H(s) = \sum_{\text{light } q}\frac{\sigma^{(\gamma\,\mathrm{exchange})}_{e^+e^-\to q\bar{q}}(s)}{\sigma^{(\gamma\,\mathrm{exchange})}_{e^+e^-\to\mu^+\mu^-}(s)} = \sum_{\text{light } q}3Q_f^2 \ , \tag{278}$$

For B factory experiments (BaBar, Belle, Belle-II) that runs above twice of the bottom quark mass, light quark includes u,d,s,c,b, and $R_H=3$. For τ-charm factory experiments (BES) that runs above twice of the charm quark mass, light quark includes u,d,s,c, and $R_H=8/3$. For experiments running at GeV $\lesssim\sqrt{s}<2m_c$, we have $R_H=2$. For experiments with $\sqrt{s}\lesssim$ GeV, the above rule breaks down because we are entering the mesonland of low energy QCD. In this case, resonance are expected when $\sqrt{s}$ equals mass of a meson.

The following plot is stolen from the textbook by Burgess and Moore, which displays the experimentally measured values of R_H as a function of $\sqrt{s}$.

Verifying the above predicted value of R_H also provides a measurement of the number of colors in QCD, $N_c = 3$.



4.2.2 e^+e^- annihilation near the Z-pole

Next, we consider another special energy regime of e^+e^- collision with $\sqrt{s} \sim M_Z$, i.e., the lepton collider runs at energies that scans the Z-pole. Again, we comparing the prefactors in the two terms of Eq. (271) and find

$$\frac{e^2 Q_f}{s} \gg \left| \frac{g_Z^2}{s - M_Z^2} \right| . \tag{279}$$

for $|\sqrt{s} - M_Z| \ll M_Z$. This suggests that the Z-boson exchange contribution dominates over that of photon exchange.

The way to regularize the scattering amplitude when $\sqrt{s} = M_Z$ is to include the decay with of Z boson in the propagator,

$$\frac{1}{s - M_Z^2} \to \frac{1}{s - M_Z^2 + i M_Z \Gamma_Z} , \tag{280}$$

where $\Gamma_Z = 2.4\,\mathrm{GeV}$ is the total decay width of the Z boson. Because $\Gamma_Z \ll M_Z$, Eq. (279) still holds after the above replacement.

The squared scattering amplitude near Z pole is

$$\begin{aligned}
\overline{|\mathcal{M}|^2} &= \frac{g_Z^4}{4} \frac{1}{(s - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \mathrm{Tr}[\not{p}_2 \gamma^\mu (g_{eL} \mathbb{P}_L + g_{eR} \mathbb{P}_R) \not{p}_1 \gamma^\nu (g_{eL} \mathbb{P}_L + g_{eR} \mathbb{P}_R)] \\
&\qquad \times \mathrm{Tr}[\not{p}_3 \gamma_\mu (g_{fL} \mathbb{P}_L + g_{fR} \mathbb{P}_R) \not{p}_4 \gamma_\nu (g_{fL} \mathbb{P}_L + g_{fR} \mathbb{P}_R)] \\
&= \frac{g_Z^4}{4} \frac{1}{(s - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \left\{ g_{eL}^2 \mathrm{Tr}[\not{p}_2 \gamma^\mu \not{p}_1 \gamma^\nu \mathbb{P}_L] + g_{eR}^2 \mathrm{Tr}[\not{p}_2 \gamma^\mu \not{p}_1 \gamma^\nu \mathbb{P}_R] \right\} \\
&\qquad \times \left\{ g_{fL}^2 \mathrm{Tr}[\not{p}_3 \gamma_\mu \not{p}_4 \gamma_\nu \mathbb{P}_L] + g_{fR}^2 \mathrm{Tr}[\not{p}_3 \gamma_\mu \not{p}_4 \gamma_\nu \mathbb{P}_R] \right\} \\
&= \frac{g_Z^4}{2} \frac{1}{(s - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \left[(g_{eL}^2 + g_{eR}^2)(g_{fL}^2 + g_{fR}^2)(u^2 + t^2) + (g_{eL}^2 - g_{eR}^2)(g_{fL}^2 - g_{fR}^2)(u^2 - t^2) \right] .
\end{aligned} \tag{281}$$

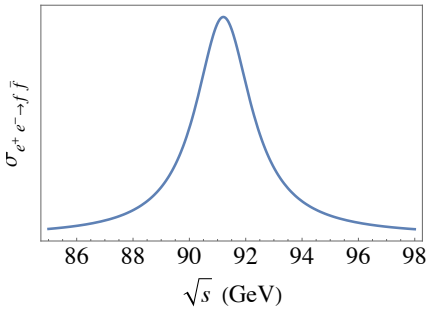
In the last step, we have expressed the momentum scalar products as the Mandelstam variables.

The unpolarized cross section is

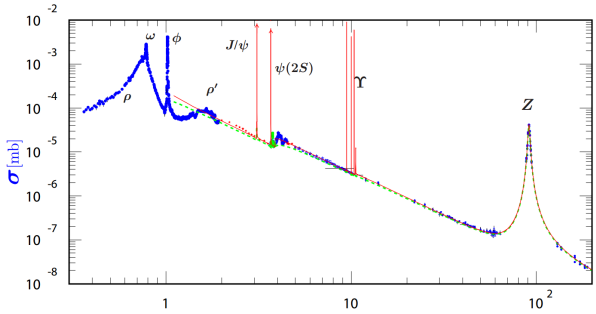
$$\begin{aligned}\sigma_{e^+e^-\to f\bar{f}}(s) &= \frac{1}{64\pi s}\frac{1}{|\vec{p}_{1\mathrm{cm}}|^2}\int_{-s}^{0} dt\overline{|\mathcal{M}|^2} \\ &= \frac{N_c g_Z^4 (g_{eL}^2+g_{eR}^2)(g_{fL}^2+g_{fR}^2)}{48\pi}\frac{s}{(s-M_Z^2)^2+M_Z^2\Gamma_Z^2}\ ,\end{aligned} \tag{282}$$

where we used $g_Z^4 = 32G_F^2M_Z^4$, and $\int_{-s}^{0} dt(t^2+u^2) = \frac{2}{3}s^3$, $\int_{-s}^{0} dt(t^2-u^2) = 0$.

In the picture below, we plot the cross section as a function of the CM energy $\sqrt{s}$. The peak around M_Z is called the Breit-Wigner resonance.



By scanning the e^+e^- collider energy over a wide range, a number of resonances have been discovered in the Standard Model, including Z, Υ, J/ψ, ρ, ω, etc. Some of them are very narrow (small decay width).



4.2.3 Narrow width approximation

A very useful approximation often made in collider physics analysis is the narrow width approximation. It works well when the total decay width of a particle is much smaller compared to its mass.

The following quantity has the asymptotic form as the Dirac-δ function,

$$\lim_{\Gamma_Z \to 0} \frac{1}{(s-M_Z^2)^2 + M_Z^2 \Gamma_Z^2} = \frac{\pi}{M_Z \Gamma_Z} \delta(s - M_Z^2) \ . \tag{283}$$

In the narrow limit, the $e^+e^- \to f\bar{f}$ cross section can be written as

$$\sigma_{e^+e^- \to f\bar{f}} = \frac{N_c g_Z^4 (g_{eL}^2 + g_{eR}^2)(g_{fL}^2 + g_{fR}^2) M_Z}{48 \Gamma_Z} \delta(s - M_Z^2) \ . \tag{284}$$

Let us recall the unpolarized $Z \to f\bar{f}$ partial decay width calculated in Eq. (216). In the $m_f = 0$ limit,

$$\Gamma_{Z \to f\bar{f}} \simeq \frac{N_c g_Z^2 M_Z}{24\pi} (g_L^2 + g_R^2) \ , \tag{285}$$

where we used the relation $g_{L,R} = g_V \mp g_A$, thus $g_{fL}^2 + g_{fR}^2 = 2(g_{fV}^2 + g_{fA}^2)$. This allows us to write

$$\sigma_{e^+e^- \to f\bar{f}} = \frac{12\pi^2 \Gamma_{Z \to e^+e^-} \Gamma_{Z \to f\bar{f}}}{M_Z \Gamma_Z} \delta(s - M_Z^2) \ . \tag{286}$$

Recall that the ratio $\Gamma_{Z \to f\bar{f}} / \Gamma_Z$ is nothing but the branching ratio for Z decaying into $f\bar{f}$. Thus, the above cross section $\sigma_{e^+e^- \to f\bar{f}}$ factorizes into two parts, the production cross section of $e^+e^- \to Z$, and the branching ratio for Z to decay into a particular $f\bar{f}$ final state, where

$$\sigma_{e^+e^- \to Z} = \frac{12\pi^2 \Gamma_{Z \to e^+e^-}}{M_Z} \delta(s - M_Z^2) \ . \tag{287}$$

The prefactor we got here is consistent with the famous lecture notes by Tao Han, `https://arxiv.org/pdf/hep-ph/0508097.pdf`. This cross section can also be verified by directly calculating the $2 \to 1$ annihilation cross section. Crossing symmetry of the S matrix tells that

$$|\mathcal{M}|^2_{e^+e^- \to Z} = |\mathcal{M}|^2_{Z \to e^+e^-} \ , \tag{288}$$

which holds without initial state spin average. Using Eq. (213), we get

$$|\mathcal{M}|^2_{e^+e^- \to Z} = 2 g_Z^2 (g_{eL}^2 + g_{eR}^2) M_Z^2 \ . \tag{289}$$

Averaging over the initial state spin and feed this to the general cross section formula, Eq. (206),

$$\begin{aligned} \sigma_{e^+e^- \to Z} &= \frac{1}{4f} \int d\Pi_f (2\pi)^4 \delta^4(p_1 + p_2 - \sum_f p_f) |\mathcal{M}|^2_{e^+e^- \to Z} \\ &= \frac{1}{2s} \frac{1}{(2\pi)^3 2M_Z} (2\pi)^4 \delta(\sqrt{s} - M_Z) \frac{1}{4} 2 g_Z^2 (g_{eL}^2 + g_{eR}^2) M_Z^2 \\ &= \frac{\pi g_Z^2 (g_{eL}^2 + g_{eR}^2)}{4 M_Z} \delta(\sqrt{s} - M_Z) \\ &= \frac{\pi g_Z^2 (g_{eL}^2 + g_{eR}^2)}{2} \delta(s - M_Z^2) = \frac{12\pi^2 \Gamma_{Z \to e^+e^-}}{M_Z} \delta(s - M_Z^2) \ . \end{aligned} \tag{290}$$

In the last step we used Eq. (285).

The "parton" level cross section of the form Eq. (287) will be extremely useful with evaluating the total cross sections by including the parton distribution functions.

4.2.4 Z-pole observables

The most obvious Z-pole observables are the mass and width of the Z boson, which can be obtained by studying $\sigma_{e^+e^-\to f\bar{f}}$ as a function of $\sqrt{s}$ (see the figure two pages ago). In addition, the overall Z production rate and its branching ratio are also sensitive to combinations of g_Z, $\sin^2\theta_W$, as well as the $SU(2)_L \times U(1)_Y$ quantum numbers of various fermions.

In this subsection, we point out Z-pole observables that allow us to measure $\sin^2\theta_W$ alone. They are the asymmetries of final state distributions or asymmetries due to different choices of polarized e^+e^- beams.

Forward-backward asymmetry

Let's consider a concrete process $e^+e^- \to \mu^+mu^-$ that occurs around the Z pole. From Eq. (269), it has been understood that for scattering angle θ taking values between 0 and π, the Mandelstam variable t ranges from $-s$ to 0. This time, we do the final state phase integral Eq. (282) more carefully, but dividing the range of scattering angle into two regions $0 \le \theta \le \pi/2$, and $\pi/2 \le \theta \le \pi$. They define the forward and backward directions for the final state μ^- to travel with respect to the initial e^- beam. The corresponding ranges of t are $-s/2 \le t \le 0$ and $-s \le t \le -s/2$, respectively. With these, we define and forward and backward parts of the cross sections as

$$\begin{aligned}\sigma_+ &= \frac{1}{64\pi s}\frac{1}{|\vec{p}_{1\mathrm{cm}}|^2}\int_{-s/2}^{0} dt\overline{|\mathcal{M}|^2}\ ,\\ \sigma_- &= \frac{1}{64\pi s}\frac{1}{|\vec{p}_{1\mathrm{cm}}|^2}\int_{-s}^{-s/2} dt\overline{|\mathcal{M}|^2}\ ,\end{aligned} \tag{291}$$

where $\overline{|\mathcal{M}|^2}$ is given by Eq. (281). With them, we can define the forward-backward asymmetry for $e^+e^- \to f\bar{f}$ at the Z-pole

$$A_{\mathrm{FB}}^f = \frac{\sigma_+ - \sigma_-}{\sigma_+ + \sigma_-}\ . \tag{292}$$

In practice, the measurement can be best done for $\mu^+\mu^-$ final states, where the electron charge of muons can be easily measured.

The relevant integrals can be done by using

$$\begin{aligned}&\int_{-s/2}^{0} dt(u^2+t^2) = \frac{1}{3}s^3\ , \quad \int_{-s/2}^{0} dt(u^2-t^2) = \frac{1}{4}s^3\ ,\\ &\int_{-s}^{-s/2} dt(u^2+t^2) = \frac{1}{3}s^3\ , \quad \int_{-s}^{-s/2} dt(u^2-t^2) = -\frac{1}{4}s^3\ .\end{aligned} \tag{293}$$

They lead to

$$A_{\rm FB}^f = \frac{3}{4}\frac{(g_{eL}^2 - g_{eR}^2)(g_{fL}^2 - g_{fR}^2)}{(g_{eL}^2 + g_{eR}^2)(g_{fL}^2 + g_{fR}^2)} \ . \tag{294}$$

It is remarkable to notice that $A_{\rm FB}^f$ only depends on the weak mixing angle θ_W. Indeed, taking $e^+e^- \to \mu^+\mu^-$ as example, we have

$$g_{eL} = g_{\mu L} = -\frac{1}{2} + \sin^2\theta_W \ , \quad g_{eR} = g_{\mu R} = \sin^2\theta_W \ . \tag{295}$$

The measurement of the forward-backward asymmetry provides a direct probe of the weak mixing angle as a fundamental parameter of the Standard Model. The value of $\sin^2\theta_W$ has been measured at the Z-pole to be 0.23.

Left-right asymmetry

Next, we consider the case of polarized electron-positron beams. Introduced polarized cross sections near the Z-pole,

$$\begin{aligned}
\sigma_{e_L^- e_R^+ \to \mu^+\mu^-} &= \frac{N_c g_Z^4 g_{eL}^2 (g_{\mu L}^2 + g_{\mu R}^2)}{48\pi} \frac{s}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \ , \\
\sigma_{e_R^- e_L^+ \to \mu^+\mu^-} &= \frac{N_c g_Z^4 g_{eR}^2 (g_{\mu L}^2 + g_{\mu R}^2)}{48\pi} \frac{s}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \ .
\end{aligned} \tag{296}$$

Here we are able to write down the cross sections directly without redoing the calculation with polarized electron helicity states, because we can read the contributions from $e_L^- e_R^+$ and $e_R^- e_L^+$ in the total cross section Eq. (282). The cross sections for other beam combinations $e_L^- e_L^+$, $e_R^- e_R^+$ must occur through the insertion of the small electron mass and are suppressed.

The left-right asymmetry is defined as

$$A_{\rm LR}^f = \frac{\sigma_{e_L^- e_R^+ \to \mu^+\mu^-} - \sigma_{e_R^- e_L^+ \to \mu^+\mu^-}}{\sigma_{e_L^- e_R^+ \to \mu^+\mu^-} + \sigma_{e_R^- e_L^+ \to \mu^+\mu^-}} = \frac{g_{eL}^2 - g_{eR}^2}{g_{eL}^2 + g_{eR}^2} \ . \tag{297}$$

Clearly, it is also a function of $\sin^2\theta_W$. Polarized electron-positron beams provides an independent measurement of $\sin^2\theta_W$ using $A_{\rm LR}^f$, which can be cross checked against the result from $A_{\rm FB}^f$.

4.2.5 Testing the electroweak theory

Here we briefly tests of the electroweak theory. The input parameters of the Standard Model related to electroweak symmetry breaking and the Higgs mechanism are

$$g_1 \ , \quad g_2 \ , \quad v \ . \tag{298}$$

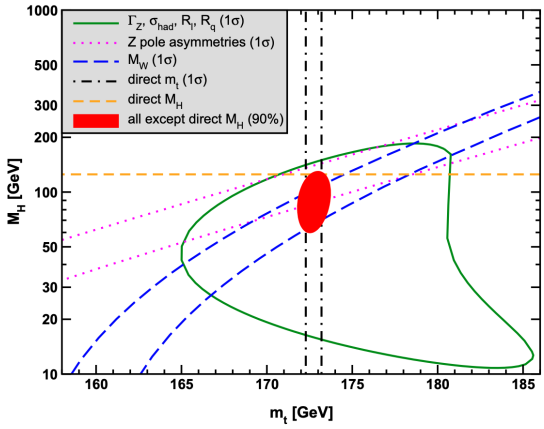
They can be reorganized into 3 independent parameters that are more widely used parameters

$$\alpha = \frac{1}{4\pi}\frac{(g_1 g_2)^2}{g_1^2 + g_2^2} \ , \quad G_F = \frac{\sqrt{2}g_2^2}{8M_W^2} = \frac{1}{\sqrt{2}v^2} \ , \quad \sin^2\theta_W = \frac{g_1^2}{g_1^2 + g_2^2} \ . \tag{299}$$

The fine-structure constant α is most precisely measured in $g-2$ and atomic physics experiments. The Fermi constant is most precisely measured in the weak decay rate of the muon. As just discussed in the previous subsection, $\sin^2\theta_W$ can be measured using asymmetry observables at the Z pole.

In addition, there are many other observables that are sensitive to combination of the above parameters, such as the Z, W boson masses and decay widths, production cross sections in e^+e^- collisions, and the R factor discussed earlier. Clearly, there are more observables than the number of input parameters. The problem is over constrained, which in turn provide consistency tests of Standard Model predictions. Remarkably, the Standard Model has passed all these tests so far. The $SU(2) \times U(1)$ gauge theory that unifies the electromagnetic and weak interactions is extremely successful.

The above arguments are based on tree level considerations. In fact, the LEP experiment has tested these observables to per mille level, sensitive even to one-loop radiative corrections. At loop level, additional input parameters are involved, including the masses of the top quark and Higgs boson. Precision tests of the Standard Model can also provide an indirect measurement of these parameters. Below is a nice plot illustrating the interplay among various measurements, taken from `https://arxiv.org/pdf/1908.07327.pdf`.

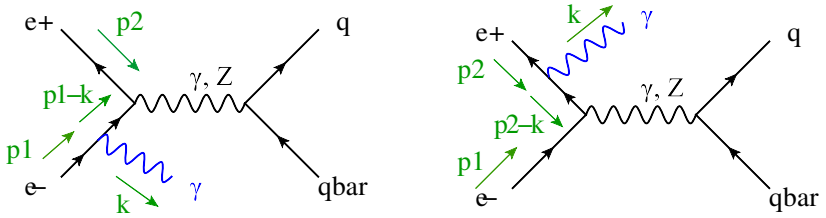


4.3 Sudakov

Up to now in this section, we have only discussed $2 \to 2$ processes. In this subsection, we take one step further and consider $e^+e^- \to q\bar{q} + \gamma$. Naively, one would expect that radiating an extra photon involves an extra factor of e at amplitude level, and in turn an extra factor of $\alpha/\pi \sim 1/400$ at cross section level, which makes it very suppressed compared to $e^+e^- \to q\bar{q}$. As we will see, such a naive estimate is not always true.

In particular, we consider the photon to be radiated from the initial state, and moreover, the photon is soft and collinear with respect to one of the beams.

With the momentum assignment in the Feynman diagrams below, this corresponds to $k \ll p_{1,2}$, and $\vec{k} /\!/ \vec{p}_1$ or $\vec{p}_2$.



For the left diagram, let's only write down the e^- part of the diagram that connects to the $s$channel γ, Z which includes the propagator

$$\begin{aligned}
&\frac{i}{\not{p}_1 - \not{k} - m_e}(-ie\gamma^\mu)u(p_1, s_1)\varepsilon_\mu^{\lambda *}(k) \\
&= e\varepsilon_\mu^{\lambda *}(k)\frac{\not{p}_1 - \not{k} + m_e}{(p_1 - k_1)^2 - m_e^2}\gamma^\mu u(p_1, s_1) \\
&= e\varepsilon_\mu^{\lambda *}(k)\frac{2p_1^\mu - \gamma^\mu\not{p}_1 - \not{k}\gamma^\mu + \gamma^\mu m_e}{(p_1 - k_1)^2 - m_e^2}u(p_1, s_1) \\
&= e\varepsilon_\mu^{\lambda *}(k)\frac{2p_1^\mu - \not{k}\gamma^\mu}{-2p_1 \cdot k}u(p_1, s_1) \ ,
\end{aligned} \tag{300}$$

where in the second step we interchange the position of $\not{p}_1$ and γ^μ, and in the last step, we used equation of motion $(\not{p}_1 - m_e)u(p_1, s_1) = 0$ and on-shell conditions $p_1^2 = m_e^2$, $k^2 = 0$.

Next, we apply the soft photon assumption and drop the $\not{k}\gamma^\mu$ terms in the numerator. This leads to

$$-\frac{ep_1 \cdot \varepsilon^{\lambda *}(k)}{p_1 \cdot k}u(p_1, s_1) \ . \tag{301}$$

We can similarly work out the right diagram with the soft photon radiated from the e^+ beam. This amounts to multiply the corresponding spinor with a similar factor

$$\frac{ep_2 \cdot \varepsilon^{\lambda *}(k)}{p_2 \cdot k}\bar{v}(p_2, s_2) \ . \tag{302}$$

Because the factors in front of $u(p_1, s_1)$ or $v(p_2, s_2)$ contain no γ matrices, they are simplify a multiplication factors. We can merge the two spinors with the rest part of the scattering amplitudes, and discover the following

$$\mathcal{M}_{e^+e^- \to q\bar{q}\gamma} \simeq \mathcal{M}_{e^+e^- \to q\bar{q}} \times e\left(\frac{p_2 \cdot \varepsilon^{\lambda *}(k)}{p_2 \cdot k} - \frac{p_1 \cdot \varepsilon^{\lambda *}(k)}{p_1 \cdot k}\right) \ , \tag{303}$$

which holds for $k \ll p_{1,2}$. The terms in the bracket are famously known as the soft photon factorization.

To calculate the total cross section, we square the amplitude and average (sum) over initial (final) state spins and colours, and then integrate over the 3-body final state phase space.

$$\begin{aligned}
\sigma_{e^+e^-\to q\bar{q}\gamma} &\simeq \frac{1}{4f}\int \frac{d^3\vec{k}}{(2\pi)^3 2E_k}\frac{d^3\vec{p}_3}{(2\pi)^3 2E_3}\frac{d^3\vec{p}_4}{(2\pi)^3 2E_4}(2\pi)^4\delta^4(p_1+p_2-p_3-p_4-k) \\
&\qquad \times \overline{|\mathcal{M}_{e^+e^-\to q\bar{q}|}|^2} \times e^2 \sum_\lambda \left|\frac{p_2\cdot\varepsilon^{\lambda *}}{p_2\cdot k}-\frac{p_1\cdot\varepsilon^{\lambda *}}{p_1\cdot k}\right|^2 \\
&\simeq \frac{1}{4f}\int \frac{d^3\vec{p}_3}{(2\pi)^3 2E_3}\frac{d^3\vec{p}_4}{(2\pi)^3 2E_4}(2\pi)^4\delta^4(p_1+p_2-p_3-p_4)\overline{|\mathcal{M}_{e^+e^-\to q\bar{q}}|^2} \\
&\qquad \times e^2 \int \frac{d^3\vec{k}}{(2\pi)^3 2E_k} \sum_\lambda \left|\frac{p_2\cdot\varepsilon^{\lambda *}}{p_2\cdot k}-\frac{p_1\cdot\varepsilon^{\lambda *}}{p_1\cdot k}\right|^2 \\
&= \sigma^{\rm LO}_{e^+e^-\to q\bar{q}} \times e^2 \int_{\rm soft} \frac{d^3\vec{k}}{(2\pi)^3 2E_k} \sum_\lambda \left|\frac{p_2\cdot\varepsilon^{\lambda *}}{p_2\cdot k}-\frac{p_1\cdot\varepsilon^{\lambda *}}{p_1\cdot k}\right|^2 \ .
\end{aligned} \tag{304}$$

In the second step, we eliminate k from the δ functions, and restrict the k integral to the soft regime (to be quantified). As a result the first line asymptotes to the leading order $e^+e^-\to q\bar{q}$ cross section. To evaluate the cross section for $2\to 3$ process with very soft photon radiated from the initial states, we simply only need to know the $2\to 2$ cross section and complete the extra multiplicative factor in the last line of Eq. (304).

Employing the photon polarization sum formula presented in Eq. (84), the multiplicative factor can be written as

$$\begin{aligned}
X_1 &= e^2 \int_{\rm soft} \frac{d^3\vec{k}}{(2\pi)^3 2E_k} \sum_\lambda \left|\frac{p_2\cdot\varepsilon^{\lambda *}}{p_2\cdot k}-\frac{p_1\cdot\varepsilon^{\lambda *}}{p_1\cdot k}\right|^2 \\
&= -e^2 \int_{\rm soft} \frac{d^3\vec{k}}{(2\pi)^3 2E_k} \left(\frac{p_2^\mu}{p_2\cdot k}-\frac{p_1^\mu}{p_1\cdot k}\right)\left(\frac{p_{2\mu}}{p_2\cdot k}-\frac{p_{1\mu}}{p_1\cdot k}\right) \\
&\simeq 2e^2 \int_{\rm soft} \frac{d^3\vec{k}}{(2\pi)^3 2E_k} \frac{p_1\cdot p_2}{(p_1\cdot k)(p_2\cdot k)} \\
&= \frac{e^2}{4\pi^2}\int_{\rm soft} E_k dE_k \int_{-1}^{1} d\cos\theta \frac{p_1\cdot p_2}{(p_1\cdot k)(p_2\cdot k)} \\
&= \frac{e^2 s}{8\pi^2}\int_{\rm soft} E_k dE_k \int_{-1}^{1} d\cos\theta \frac{1}{(p_1\cdot k)(p_2\cdot k)} \ ,
\end{aligned} \tag{305}$$

where we neglected terms proportional to m_e^2 in the third step. In the next to last step, we have chosen $\vec{p}_1$ to be along the $\hat{z}$ axis, and θ is the polar angle of radiated photon direction. Working in the CM frame, we have used $\vec{p}_1=-\vec{p}_2$ and $p_1\cdot p_2 = s/2$. Clearly, the above integrand goes to infinity in the limit $E_k\to 0$ (thus $k^\mu=(E_k,\vec{k})\to 0$ in every component for on-shell photon). This is the soft divergence.

Next, we use the other feature of the radiation photon – its collinearity with respect to either of the beam. We first make the argument in the massless electron limit. The above integrand has another way to be divergent with particular choices of the θ. If $\theta = 0$, we have $k^\mu = z p_1^\mu$ where z is a small number, thus $k \cdot p_1 = 0$. If $\theta = \pi$, we have $k^\mu = z p_2^\mu$ where z is a small number, thus $k \cdot p_2 = 0$. These are called collinear divergences.

In practice, the collinear divergences can be regularized with a non-zero electron mass. Near $\theta = 0$, we have

$$\begin{aligned} &p_1 \cdot k = E_1 E_k - \vec{p}_1 \cdot \vec{k} = E_1 E_k - \sqrt{E_1^2 - m_e^2} E_k \cos\theta \simeq E_1 E_k \left(\frac{m_e^2}{2E_1^2} + 1 - \cos\theta \right) \ , \\ &p_2 \cdot k \simeq 2 E_2 E_k \ . \end{aligned} \tag{306}$$

Near $\theta = \pi$, we have

$$\begin{aligned} &p_1 \cdot k = 2 E_1 E_k \ , \\ &p_2 \cdot k = E_2 E_k - \vec{p}_2 \cdot \vec{k} = E_2 E_k - \sqrt{E_2^2 - m_e^2} E_k \cos(\pi - \theta) \simeq E_2 E_k \left(\frac{m_e^2}{2E_2^2} + 1 + \cos\theta \right) \ . \end{aligned} \tag{307}$$

With this knowledge, we restrict the evaluation of X_1 factor defined in Eq. (305) in the regime where photon is both soft and collinear,

$$\begin{aligned} X_1 &= e^2 \int_{\text{soft\&collinear}} \frac{d^3\vec{k}}{(2\pi)^3 2E_k} \sum_\lambda \left| \frac{p_2 \cdot \varepsilon^{\lambda *}}{p_2 \cdot k} - \frac{p_1 \cdot \varepsilon^{\lambda *}}{p_1 \cdot k} \right|^2 \\ &= \frac{e^2 s}{8\pi^2} \int_{\text{soft}} E_k dE_k \left[\int_{\theta \simeq 0} d\cos\theta \frac{1}{(p_1 \cdot k)(p_2 \cdot k)} + \int_{\theta \simeq \pi} d\cos\theta \frac{1}{(p_1 \cdot k)(p_2 \cdot k)} \right] \\ &= \frac{e^2}{4\pi^2} \int_{\text{soft}} \frac{dE_k}{E_k} \left[\int_{\theta \simeq 0} d\cos\theta \frac{1}{\frac{m_e^2}{2E_1^2} + 1 - \cos\theta} + \int_{\theta \simeq \pi} d\cos\theta \frac{1}{\frac{m_e^2}{2E_2^2} + 1 + \cos\theta} \right] \\ &= \frac{e^2}{2\pi^2} \log \frac{E_k^{\max}}{\mu} \left[\log \frac{s}{m_e^2} + \mathcal{O}(1) \right] \ , \end{aligned} \tag{308}$$

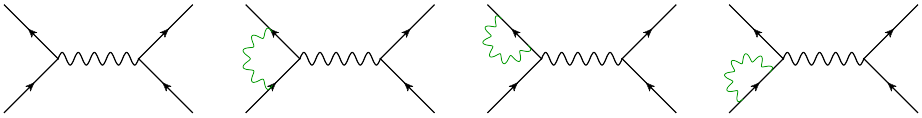
where we used $E_1 = E_2 \simeq \sqrt{s}/2$ for $\sqrt{s} \gg m_e$. In the last line, the first logarithmic factor comes from the E_k integral, where we introduce the fictitious photon mass μ to regularize it. The quantity $E_k^{\max}$ defines how soft a photon needs to be in order to be considered soft. Experimentally, it can be identified as the energy threshold for a detector, below which the photon cannot be identified and recorded to the tapes. From now on, we will call $E_k^{\max} = E_{\text{th}}$. The square bracket comes from the θ integral. The order 1 factor depends how far the θ is integrated away from 0 or π, and it is much smaller than the logarithmic factor.

To summarize, the soft-collinear photon radiation corresponds to a cross section

$$\sigma_{e^+e^- \to q\bar{q}\gamma} \simeq \frac{\alpha}{\pi} \left(\log \frac{E_{\text{th}}^2}{\mu^2} \right) \left(\log \frac{s}{m_e^2} \right) \sigma^{\text{LO}}_{e^+e^- \to q\bar{q}} \ . \tag{309}$$

Compared to the $2 \to 2$ cross section, the suppression factor is not simply α/π, but comes with two log enhancements. One of them is even divergent because photon is massless.

Experimentally, if the final state photon is not measured, we would consider $e^+e^- \to q\bar{q}\gamma$ to be part of the $e^+e^- \to q\bar{q}$ process and add the two cross sections up. But the new one is divergent. This is because we are still missing a class of important Feynman diagrams at one loop level, as shown below.



The first diagram is the leading order $2 \to 2$ diagram. The remaining three diagrams are loop corrections with virtual photon exchange among the initial states. Their amplitude carry extra factor of $e^2/(16\pi^2)$ compared to the leading diagram. The interference terms between tree-level and loop level diagrams, $\mathrm{Re}(\mathcal{M}_{\mathrm{tree}}\mathcal{M}^*_{\mathrm{loop}})$, contributes to the total cross section, and is of order α/π compared to the leading contribution from $|\mathcal{M}_{\mathrm{tree}}|^2$. Importantly, there are also logarithmic factors. Without detailed computation (see Peskin Chapter 6.4), the $2 \to 2$ cross section after including one-loop corrections is

$$\begin{aligned} \sigma^{\text{one-loop}}_{e^+e^-\to q\bar{q}} &= X_2 \sigma^{\mathrm{LO}}_{e^+e^-\to q\bar{q}} \\ X_2 &\simeq -\frac{\alpha}{\pi}\left(\log\frac{s}{\mu^2}\right)\left(\log\frac{s}{m_e^2}\right) \ . \end{aligned} \tag{310}$$

The "effective" next-to-leading order cross section for $e^+e^- \to q\bar{q}$ (which can be used to compare with experimental data collected a detector with threshold E_{th}) is

$$\begin{aligned} \sigma^{\mathrm{NLO}}_{e^+e^-\to q\bar{q}} &= \sigma^{\mathrm{LO}}_{e^+e^-\to q\bar{q}} + \sigma^{\text{one-loop}}_{e^+e^-\to q\bar{q}} + \sigma_{e^+e^-\to q\bar{q}\gamma} \\ &= (1 + X_1 + X_2)\sigma^{\mathrm{LO}}_{e^+e^-\to q\bar{q}} \\ &\simeq \left[1 - \frac{\alpha}{\pi}\left(\log\frac{s}{E_{\mathrm{th}}^2}\right)\left(\log\frac{s}{m_e^2}\right)\right]\sigma^{\mathrm{LO}}_{e^+e^-\to q\bar{q}} \ . \end{aligned} \tag{311}$$

Remarkable, adding X_1 and X_2 gives a finite result. The factor $X = X_1 + X_2$ is called the Sudakov double logarithm factor.

Even without talking about the detector threshold or any detector at all, we still need to add up the real and virtual photon processes. The theoretical reason is, when k is small, both the real photon and the loop photon are close to on-shell. Its wavelength is very so large that after a photon is radiated from the beam, one cannot tell whether it travels to infinity or gets absorbed back to the beam. In the spirit of path integral, one has to include all the possible fate of the photon. This is the fundamental reason for treating the real and virtual photon processes on the same footing.

In perturbative calculations, ressumation is a technique to sum up a class (not all) of Feynman diagrams. We could perform a ressumation of multiple soft-collinear photon radiations from the initial states, and include loop corrects to the same order of α. This leads to

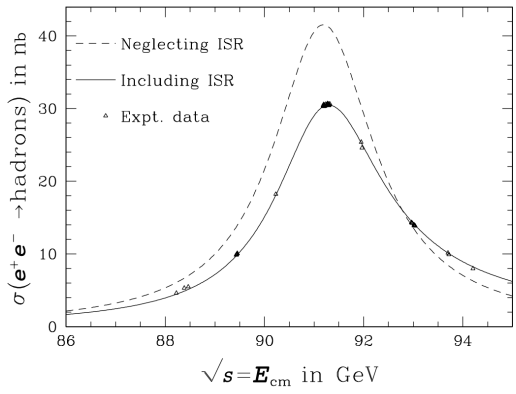
$$\begin{aligned}\sigma^{\text{resummed}}_{e^+e^-\to q\bar{q}} &= \sigma^{\text{LO}}_{e^+e^-\to q\bar{q}} \sum_{n=0}^{\infty} \frac{X^n}{n!} = e^X \sigma^{\text{LO}}_{e^+e^-\to q\bar{q}} \\ &= \exp\left[-\frac{\alpha}{\pi}\left(\log\frac{s}{E^2_{\text{th}}}\right)\left(\log\frac{s}{m_e^2}\right)\right]\sigma^{\text{LO}}_{e^+e^-\to q\bar{q}} \ ,\end{aligned} \tag{312}$$

where the factor $n!$ arises because all the radiated photons are identical particles. The exponential is the Sudakov factor.

You may have noticed that all the real and virtual photon processes considered above are from the initial states (beams). Other diagrams involve the final state quarks do exist. However, they do not introduce any logarithmic factors, and are simply suppressed by powers of α/π compared to leading order cross section. We have neglected them.

The following plot are stolen from Fig. 6.10 of Burgess and Moore textbook. It shows the comparison of leading order cross section for $e^+e^- \to$ hadronic final states (dashed curve) and the next-to-leading order one (solid curve). The center of mass energies are around the Z pole. Clearly, radiative correction (dominated by the Sudakov factor) makes a $\sim -30\%$ change from the leading order result, and it is confirmed by experimental data.

Another interesting thing to notice is that radiative correction makes the Breit-Wigner peak of Z boson asymmetric, due to the energy dependence in the Sudakov factor.



In the above discussions, we considered QED processes of real photon radiation. Similar results hold for QCD. If we work with hadron colliders with quark-anti-quark initial states, there could be real gluon radiations as well. In the latter case, the Sudakov factor is proportional to $\alpha_s = g_3^2/(4\pi)$ instead of α. The corresponding radiative correction can play an even more significant role.

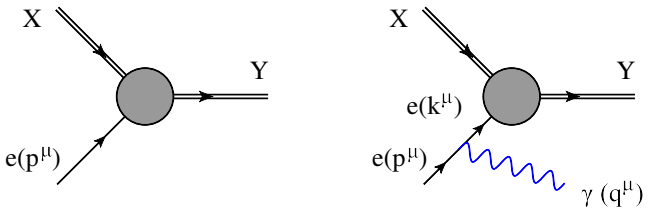
4.4 Parton distribution functions

In this section, we will continue to deal with singularities of real photons radiation off an electron beam. This time we consider collinear photons that are not necessarily soft. The photons are still lost to observation because they continue to travel along the beam direction. Such a discussion will reveal an important aspect of the electron beam, which is actually composite. There are nonzero probabilities to find the electron carrying only a fraction (any value between 0 and 1) of the total beam energy. The same applies to finding the photon in the beam as well.

To proceed, let us consider a general process where an electron beam is injected and the electron can participate in a process $e + X \to Y$, as shown by the Feynman diagram in the left of the figure below. Here X could be another beam or a fixed target, and Y represents the resulting set of final state particles. If you want a concrete example, we could consider elastic $e^-\mu^-$ collision with t-channel photon and Z boson exchanges. We denote the scattering amplitude as $\mathcal{M}_{eX\to Y}(E)$, where E is the energy of the electron coming into the vertex (gray blob). For convenience, we factorize out the electron's spinor,

$$\mathcal{M}_{eX\to Y}(E) = Au(p,s) \ , \tag{313}$$

where A accounts for the rest part of the amplitude, and $p^\mu = (E, \vec{p})$ is the four momentum of the electron.



Next, we consider a real photon radiation off the incoming electron, carrying momentum q^μ. See the right diagram above. This results in an electron propagator with momentum $k^\mu = p^\mu - q^\mu$. The amplitude for this $eX \to \gamma Y$ process can be written as

$$\begin{aligned} \mathcal{M}_{eX\to\gamma Y}(E) &= A\frac{i}{\not{k} - m_e}(-ie\gamma^\mu)u(p,s)\varepsilon^{\lambda *}_\mu(q) \\ &= eA\frac{\not{k} + m_e}{k^2 - m_e^2}\gamma^\mu u(p,s)\varepsilon^{\lambda *}_\mu(q) \ . \end{aligned} \tag{314}$$

Because we consider high energy electron scattering, the electron mass will be dropped until the point where it is needed for regularizing momentum integrals.

Next, we take into account of the fact that the photon is very collinear with the electron. At leading order

$$k^2 \simeq 0 \ . \tag{315}$$

This implies that the internal electron is nearly on shell. It allows us to approximately apply the spin sum relation for fermions

$$\not{k} + m_e \simeq \sum_{s'} u(k,s')\bar{u}(k,s') \ , \tag{316}$$

and rewrite the $eX \to \gamma Y$ as

$$\mathcal{M}_{eX\to\gamma Y}(p) \simeq eA\frac{\sum_{s'} u(k,s')\bar{u}(k,s')}{k^2}\gamma^\mu u(p,s)\varepsilon_\mu^{\lambda *}(q) \ . \tag{317}$$

We will see shortly why this is useful.

Consider a polarized electron beam where the electron is left-handed. Neglecting the electron mass, a chirality eigenstate is also a helicity eigenstate, thus chirality is preserved within the beam. Upon interactions, it is important to note that QED is vector current interaction, which also preserves the electron chirality before and after the photon radiation. This fixes $s' = s$, i.e., the electron remains left-handed until its hard collision with X. As a result, the $eX \to \gamma Y$ amplitude factorizes into two parts and rewrite the $eX \to \gamma Y$ as

$$\begin{aligned}\mathcal{M}_{e_L X\to\gamma Y}(E) &\simeq \Big[Au(k,s)\Big]\left[\frac{e}{k^2}\bar{u}(k,s)\gamma^\mu u(p,s)\varepsilon_\mu^{\lambda *}(q)\right] \\ &= \mathcal{M}_{e_L X\to Y}(E_k)\frac{1}{k^2}\mathcal{M}_{e_L\to e_L\gamma}(p,q) \ .\end{aligned} \tag{318}$$

The first factor is identified as the amplitude for $eX \to Y$, define in Eq. (313), but with incoming electron momentum equal to k^μ.

The leads to a nice observation that the radiation of a collinear photon simply amounts to factorization of the full amplitude into the hard scattering part ($eX \to Y$) and the photon radiation part ($e \to e\gamma$).

We still need to evaluate remaining part of Eq. (318). To do so, we take more care in writing down the four momenta, by allowing for a small transverse component of the photon with respect to the incoming electron beam (not exact collinear). While still considering the electron to be massless, we have

$$\begin{aligned}p^\mu &\simeq (E,0,0,E) \ , \\ q^\mu &\simeq \left(zE, p_\perp, 0, zE - \frac{p_\perp^2}{2zE}\right) \ .\end{aligned} \tag{319}$$

Here we introduce a small transverse momentum for the photon

$$m_e \ll p_\perp \ll E \ . \tag{320}$$

The photon carries away a faction $z \in (0,1)$ of the beam energy. The photon's four momentum square q^2 vanishes up to $p_\perp^4$ order.

Momentum conservation implies that

$$k^\mu = p^\mu - q^\mu \simeq \left((1-z)E, -p_\perp, 0, (1-z)E + \frac{p_\perp^2}{2zE}\right) \ . \tag{321}$$

As a result,

$$k^2 \simeq -\frac{1}{z}p_\perp^2 \ , \tag{322}$$

which fixes the denominator Eq. (318).

The remaining part to evaluate is the helicity amplitude $\mathcal{M}_{e_L\to e_L\gamma}(p,q)$, where we must make use of the spinor representations that are helicity eigenstates (see Eq. (58)). In particular, we have

$$\begin{aligned} u(p,+) &= \sqrt{2E}\begin{pmatrix}1\\0\\0\\0\end{pmatrix} \\ u(k,+) &= \sqrt{2(1-z)E}\begin{pmatrix}\cos\frac{\theta}{2}\\ \sin\frac{\theta}{2}e^{i\phi}\\0\\0\end{pmatrix} \simeq \sqrt{2(1-z)E}\begin{pmatrix}1\\ -\frac{p_\perp}{2(1-z)E}\\0\\0\end{pmatrix} \ , \end{aligned} \tag{323}$$

where we have used $\tan\theta \simeq -\frac{p_\perp}{(1-z)E} \ll 1$ and $\phi = 0$, based on the form of k^μ in Eq. (321). We neglected k_z which is much smaller than k_x.

Using the γ matrix forms used throughout the lecture notes, Eq. (37), we obtain

$$\mathcal{M}_{e_L\to e_L\gamma} = e\sqrt{2E}\sqrt{2(1-z)E}\chi'^\dagger\sigma^\mu\chi\varepsilon_\mu^{\lambda *}(q) \ , \tag{324}$$

where

$$\chi = \begin{pmatrix}1\\0\end{pmatrix} \ , \quad \chi' = \begin{pmatrix}1\\ -\frac{p_\perp}{2(1-z)E}\end{pmatrix} \ . \tag{325}$$

The two polarizations vectors pf photon are

$$\begin{aligned} \varepsilon_\mu^{+1*}(q) &\simeq \frac{1}{\sqrt{2}}\left(0,1,i,-\frac{p_\perp}{zE}\right) \ , \\ \varepsilon_\mu^{-1*}(q) &\simeq \frac{1}{\sqrt{2}}\left(0,1,-i,-\frac{p_\perp}{zE}\right) \ . \end{aligned} \tag{326}$$

Both satisfy $q\cdot\varepsilon_\mu^{\lambda *}(q) = 0$. Using $(\sigma_x + i\sigma_y)\chi = 0$, we obtain

$$\mathcal{M}_{e_L\to e_L\gamma_-} = -ep_\perp\frac{\sqrt{2(1-z)}}{z} \ . \tag{327}$$

The other one is slightly harder

$$\begin{aligned} \mathcal{M}_{e_L\to e_L\gamma_+} &= e\sqrt{2E}\sqrt{2(1-z)E}\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -\frac{p_\perp}{2(1-z)E}\end{pmatrix}\begin{pmatrix}-\frac{p_\perp}{zE} & 0\\ 2 & \frac{p_\perp}{zE}\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} \\ &= -ep_\perp\frac{\sqrt{2(1-z)}}{z(1-z)} + \mathcal{O}(p_\perp^2) \ . \end{aligned} \tag{328}$$

These lead to

$$\sum_{\lambda=\pm1}|\mathcal{M}_{e_L\to e_L\gamma_\lambda}|^2 = \frac{2e^2p_\perp^2}{z(1-z)}\frac{1+(1-z)^2}{z} \ . \tag{329}$$

Because QED respects parity, we can derive the same result with a right-handed polarized electron beam.

$$\sum_{\lambda=\pm 1} |\mathcal{M}_{e_R\to e_R\gamma_\lambda}|^2 = \frac{2e^2p_\perp^2}{z(1-z)}\frac{1+(1-z)^2}{z} \ . \tag{330}$$

We are now ready to write down the total cross section for $eX\to\gamma Y$ and organize it in an elegant way. The unpolarized cross section is

$$\begin{aligned}
\sigma_{eX\to\gamma Y} &= \frac{1}{4f}\int \Pi_{\text{final}} \sum_\lambda \overline{|\mathcal{M}_{eX\to\gamma_\lambda Y}|^2}(2\pi)^4\delta^4(p+p_X-q-p_Y) \\
&= \frac{1}{4f}\int \Pi_{\text{final}} \sum_\lambda \frac{1}{2}\left[|\mathcal{M}_{e_LX\to\gamma_\lambda Y}|^2+|\mathcal{M}_{e_RX\to\gamma_\lambda Y}|^2\right](2\pi)^4\delta^4(p+p_X-q-p_Y) \\
&= \frac{1}{4f}\int \Pi_{\text{final}} \frac{1}{2k^4}\left[|\mathcal{M}_{e_LX\to Y}(E_k)|^2\sum_\lambda|\mathcal{M}_{e_L\to e_L\gamma_\lambda}|^2+|\mathcal{M}_{e_RX\to Y}(E_k)|^2\sum_\lambda|\mathcal{M}_{e_R\to e_R\gamma_\lambda}|^2\right] \\
&\qquad\qquad \times(2\pi)^4\delta^4(p+p_X-q-p_Y) \\
&= \frac{1}{4f}\int \Pi_{\text{final}} \frac{1}{2}\left[|\mathcal{M}_{e_LX\to Y}(E_k)|^2+|\mathcal{M}_{e_RX\to Y}(E_k)|^2\right](2\pi)^4\delta^4(k+p_X-p_Y) \\
&\qquad\qquad \times\int\frac{d^3\vec{q}}{(2\pi)^3 2E_q}\frac{1}{k^4}\sum_\lambda|\mathcal{M}_{e_{L,R}\to e_{L,R}\gamma_\lambda}|^2 \ ,
\end{aligned} \tag{331}$$

where the prefactor is $f=\sqrt{(p\cdot p_X)^2-m_e^2m_X^2}\simeq p\cdot p_X=\frac{1}{1-z}k\cdot p_X$. Let's hold the four momentum of the other beam p_X fixed. In the last step of Eq. (331), the first line is equal to

$$(1-z)\sigma_{eX\to Y}\left(E_k\right) \ , \quad E_k=(1-z)E \ . \tag{332}$$

In the second line, the $d^3\vec{q}$ integral can be evaluated with using Eq. (319), where $q_z=zE$ and $q_x^2+q_y^2=p_\perp^2$. Thus $\int dq_z = E\int dz$, $\int dq_xdq_y=\int_0^{2\pi}d\phi p_\perp dp_\perp = \frac{1}{2}\int_0^{2\pi}d\phi dp_\perp^2$. And because $q_z\gg q_x,q_y$, we have $E_q\simeq q_z=zE$. Further using Eqs. (329) and (330), we obtain

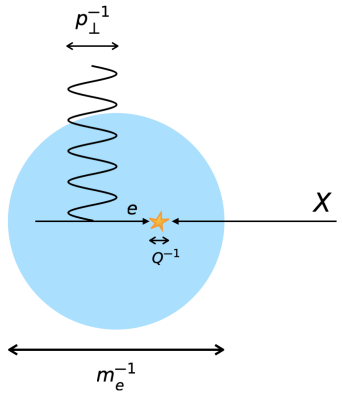
$$\begin{aligned}
&\int\frac{d^3\vec{q}}{(2\pi)^3 2E_q}\frac{1}{k^4}\sum_\lambda|\mathcal{M}_{e_{L,R}\to e_{L,R}\gamma_\lambda}|^2 \\
&= \frac{1}{16\pi^2}\int_0^1\frac{dz}{z}\int dp_\perp^2\frac{1}{k^4}\sum_\lambda|\mathcal{M}_{e_{L,R}\to e_{L,R}\gamma_\lambda}|^2 \\
&= \frac{1}{16\pi^2}\int_0^1\frac{dz}{z}\int dp_\perp^2\frac{z^2}{p_\perp^4}\sum_\lambda|\mathcal{M}_{e_{L,R}\to e_{L,R}\gamma_\lambda}|^2 \\
&= \frac{e^2}{8\pi^2}\int_0^1 dz\frac{1+(1-z)^2}{z(1-z)}\int\frac{dp_\perp^2}{p_\perp^2} \ .
\end{aligned} \tag{333}$$

Putting Eqs. (332) and (333) together, we get

$$\sigma_{eX\to\gamma Y}(E) \simeq \frac{e^2}{8\pi^2}\int_0^1 dz \frac{1+(1-z)^2}{z}\int \frac{dp_\perp^2}{p_\perp^2}\sigma_{eX\to Y}\left(E_k\right) \ . \tag{334}$$

The $p_\perp$ integral is log divergent. It can be regularized by the nonzero electron mass. We choose the upper limit to be $Q = \sqrt{-q^2} > 0$, referred to as the resolution scale associated with the $eX \to Y$ hard scattering. Physically, Q is the momentum transfer in the hard scattering, or transverse momenta p_T carried by the resulting final state particles measured by the detector. For a given high-energy collider running at CM energy $\sqrt{s}$, the typical value of Q is of same order as $\sqrt{s}$. That is why we build such a collider for. In contrast, the initial state photon radiation considered above must be sufficiently collinear and does not affect the observation of the hard scattering process. This sets an upper bound on its transverse momentum $p_\perp$, which is Q. Or maybe $Q/10$, but the fudge factor of 10 is not important because it eventually appears in a log. [3] Otherwise, if the photon carries $p_\perp$ higher than Q, we would consider it as part of the hard scattering process, rather than the beam being composite.

It is useful to think about the above process from the point of view of the electron (in its rest frame), where one can appreciate the hierarchy among the scales (see Eq. (320)). In this frame, the photon still has transverse momentum $p_\perp$, and the corresponding wavelength is $p_\perp^{-1}$. It can probe where the electron is within its Compton wavelength given by m_e^{-1}, with a resolution given by $p_\perp^{-1}$. In contrast, the eX hard scattering occurs at much shorter distances.



With the above picture, we complete the $p_\perp$ integral with Q as upper bound and use the electron mass to regularize the singularity near the lower bound

$$\int \frac{dp_\perp^2}{p_\perp^2} \simeq \log\frac{Q^2}{m_e^2} \ . \tag{335}$$

[3] More generally, one could keep the $p_\perp$ dependence, and the resulting parton distribution will depend on both x and $p_\perp$. That is called the transverse momentum dependent PDF. For more details and references, see `https://en.wikipedia.org/wiki/Transverse_momentum_distributions`.

The log divergence near $p_\perp = 0$ is regularized by the electron mass. Note Q should always be much higher than m_e.

Defining $x = 1 - z$, we finally obtain

$$\sigma_{eX\to\gamma Y} \simeq \frac{\alpha}{2\pi} \log\frac{Q^2}{m_e^2} \int_0^1 dx \frac{1+x^2}{1-x} \sigma_{eX\to Y} \ . \tag{336}$$

The physical meaning of this formula is clear. With an electron beam with energy E, there are probabilities for the electron out of the beam to participate in hard scatterings with energy xE, with $0 \le x \le 1$. The rest of beam energy is carried away by a photon. The suggests that the beam is composite, and electron is a parton of the beam.

The probability density for finding electron to carry energy fraction x from the beam can be read from Eq. (336),

$$f_e(x, Q^2) = \frac{\alpha}{2\pi} \left(\log\frac{Q^2}{m_e^2}\right) \frac{1+x^2}{1-x} \ . \tag{337}$$

However, there is still a subtlety in this expression when $x = 1$, we need to fix the singularity. Moreover, the above $f_e(x, Q^2)$ is proportional to α, indicating a photon must be radiated. It does not yet account for the possibility where the beam electron does not radiate any photon at all, which is nothing but the leading order $eX \to Y$ process. After including this contribution, Eq. (336) is modified to

$$f_e(x, Q^2) = \delta(1-x) + \frac{\alpha}{2\pi} \left(\log\frac{Q^2}{m_e^2}\right) \left[\frac{1+x^2}{(1-x)_+} + \frac{3}{2}\delta(1-x)\right] \ , \tag{338}$$

where the $1/(1-x)_+$ is the same as $1/(1-x)$ for all values of $0 \le x < 1$, but regularizes the singularity when $x \simeq 1$. It takes the form

$$\frac{1}{(1-x)_+} = \lim_{\epsilon\to 0} \left[\frac{1}{1-x}\Theta(1-x-\epsilon) - \delta(1-x)\int_0^1 dx' \frac{1}{1-x'}\right] \ , \tag{339}$$

where Θ is the step function. This implies

$$\int_0^1 dx \frac{g(x)}{(1-x)_+} = \int_0^1 dx \frac{g(x) - g(1)}{1-x} \ , \tag{340}$$

for any real function $g(x)$. For the first term in the square bracket, we have

$$\int_0^1 dx \frac{1+x^2}{(1-x)_+} = \int_0^1 dx \frac{(1+x^2)-2}{1-x} = -\frac{3}{2} \ , \tag{341}$$

As a result, the integral over x vanishes for the second term in Eq. (338). Eq. (338) is the electron parton distribution function (PDF) in an electron beam, valid up to the first of order in α. It satisfies the unitarity condition

$$\int_0^1 dx f_e(x, Q^2) = 1 \ . \tag{342}$$

The total probability of finding an electron as a parton of an electron beam (minus that of finding a positron) where the parton carries all possible fractions of beam energy must be 1.

To simplify the notation, we introduce the splitting function

$$P_{e\leftarrow e}(x) = \frac{1+x^2}{(1-x)_+} + \frac{3}{2}\delta(1-x) \ . \tag{343}$$

With this, the electron PDF can be written as

$$f_e(x,Q^2) = \delta(1-x) + \frac{\alpha}{2\pi}\left(\log\frac{Q^2}{m_e^2}\right) P_{e\leftarrow e}(x) \ . \tag{344}$$

With the electron PDF, we can define the inclusive $eX \to Y$ reaction cross section in the beam electron experiment

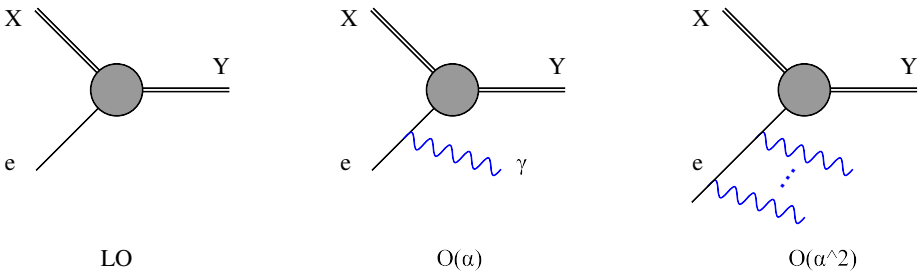
$$\sigma_{eX\to Y+(0,1,\,\ldots\,\gamma)} = \int_0^1 dx f_e(x,Q^2)\sigma_{eX\to Y} \ . \tag{345}$$

To include more than one photon in the final state, the PDF needs to be calculated beyond the first order in α. It is worth noting that both $\sigma_{eX\to Y+(0\,\mathrm{or}\,1\,\gamma)}$ and $\sigma_{eX\to Y}$ share the same argument Q^2, the resolution scale that the hard scattering vertex is probing.

Plugging Eq. (344) into Eq. (345), we obtain

$$\sigma_{eX\to Y+(0,1,\,\ldots\,\gamma)} = \sigma_{eX\to Y} + \frac{\alpha}{2\pi}\left(\log\frac{Q^2}{m_e^2}\right)\int_0^1 dx P_{e\leftarrow e}(x)\sigma_{eX\to Y} + \mathcal{O}(\alpha^2) \ . \tag{346}$$

Term by term, it corresponds to the following series of Feynman diagrams. Here, all radiated photons are collinear.



An electron beam also contains photon as a parton, and the corresponding parton distribution function is (can be read from Eq. (334))

$$f_\gamma(z,Q^2) = \frac{\alpha}{2\pi}\log\frac{Q^2}{m_e^2}P_{\gamma\leftarrow e} \ , \quad P_{\gamma\leftarrow e} = \frac{1+(1-z)^2}{z} \ . \tag{347}$$

where $P_{\gamma\leftarrow e}$ is another splitting function. For an electron beam, $f_\gamma(z,Q^2)$ must occur at least at $\mathcal{O}(\alpha)$ level. There is no need to regularize the singularity near

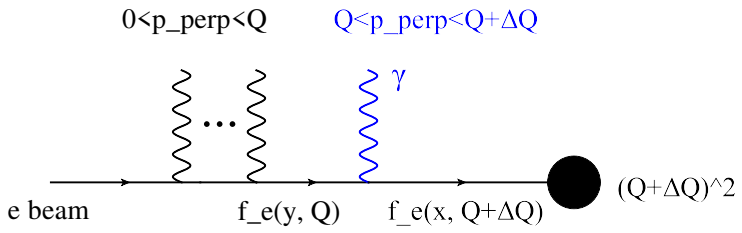
$z = 0$ because for the photon (as a parton) to participate in hard scattering processes, it must still carry a substantial amount of the original beam energy.

If we prepare a photon beam experimentally, it also contains electron and position as its parton. We will not present the derivation of the corresponding PDFs, but refer to Peskin Chapter 17.5 for further details.

4.4.1 Parton evolution

The parton distribution functions defined in Eqs. (338) and (347) depend on two arguments, the fraction of beam energy x, and the transverse momentum Q. In this subsection, we explore the Q dependence in more detail and derive a differential equation that can relate PDF at various Q energy scales.

In particular, we consider the following Feynman diagram that relates two electron PDFs corresponding to two different Q, up to order α.



Starting with an electron beam, after a number of photon radiations (all with $p_\perp < Q$), we reach an electron PDF denoted by $f_e(y, Q)$. On top of it, consider one more photon is radiated (the blue one), with $Q < p_\perp < Q + \Delta Q$, and the resulting electron PDF is $f_e(x, Q + \Delta Q)$. Here, x must be smaller than y due to extra photon radiation. Because the last photon has the highest energy, it occurs at a shorter distance than all the other photons. This explains why we put it the closest to the hard-scattering vertex with momentum transfer $Q + \Delta Q$.

After all the photon radiations, the resulting electron with energy carries energy fraction x out of the energy E from the original e beam (on the left). Effectively, we can also consider the electron before the last photon radiation as an "intermediate beam" with energy yE, and the electron after the last photon radiation as the parton of this intermediate beam with energy fraction x/y.

At leading order, the difference between $f_e(x, Q + \Delta Q)$ and $f_e(x, Q)$ is due to the radiation of the last photon with $Q < p_\perp < Q + \Delta Q$. We could do a

similar calculation similar to Eq. (334), and obtain

$$\begin{aligned} f_e(x, Q+\Delta Q) - f_e(x,Q) &= \int_x^1 dy f_e(y,Q) \left[\int dx' \frac{\alpha}{2\pi} \int_Q^{Q+\Delta Q} \frac{dp_\perp^2}{p_\perp^2} P_{e\leftarrow e}(x') \right] \delta(x - yx') \\ &= \frac{\alpha}{\pi} \log \frac{Q+\Delta Q}{Q} \int_x^1 \frac{dx'}{x'} P_{e\leftarrow e}(x') f_e\left(\frac{x}{x'}, Q\right) \\ &\xrightarrow{\Delta Q \to 0} \frac{\alpha}{\pi} \frac{\Delta Q}{Q} \int_x^1 \frac{dx'}{x'} P_{e\leftarrow e}(x') f_e\left(\frac{x}{x'}, Q\right) \ . \end{aligned} \tag{348}$$

In the first line, the square bracket is the probability of finding an electron carrying energy fraction $x' = x/y$ from the intermediate "beam" with energy yE. In the second line, we complete the y integral first using the δ function. The resulting range of x' integral is set such that the first argument in f_e does not exceed 1, i.e., $x/x' \leq 1$. The last step allows us to write down a differential equation for f_e in the $\Delta Q \to 0$ limit,

$$\frac{df_e(x,Q)}{d\log Q} = \frac{\alpha}{\pi} \int_x^1 \frac{dx'}{x'} P_{e\leftarrow e}(x') f_e\left(\frac{x}{x'}, Q\right) \ . \tag{349}$$

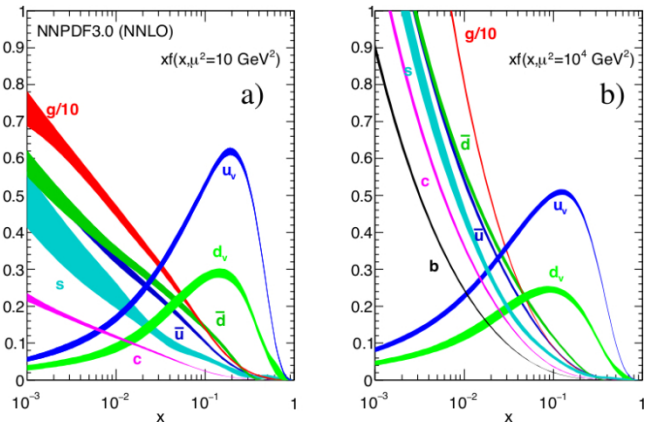
This represents one of the DEGLAP equations. It is an integro-differential equation.

For the case of QED, one might not feel very impressed about this. By inserting Eq. (338) to the both sides of the DEGLAP equation and keeping the leading term on the right (only the δ function from f_e in the integrand), we find it works trivially. This is indeed the case because QED is completely perturbative, and we can derive a PDF from the first principle.

The DEGLAP equation is a lot more useful for the case of QCD, when we consider quark/gluon PDFs out of a proton beam, which are no longer calculable due to its non-perturbative nature. Thus, we are not able to derive anything similar to Eq. (338) from first principle. Instead, the above DEGLAP equation still holds (with the replacement $\alpha \to \alpha_3$). What we could do is to first measure the PDF experimentally at certain energy scale Q, and then use the DEGLAP equation to extrapolate it another scale Q', as along as the gauge coupling α_3 is perturbative between the two scales.

This is exactly the state of art of high-energy collider physics. The PDFs of proton can be measured in deep-inelastic scattering processes with $Q \sim$ a few GeV (to be discussed in the next section). They are then extrapolated to high scales (hundreds of GeV to several TeV) for estimating the production cross sections of heavy particles at the LHC, using the DEGLAP equation.

The following picture illustrates the scale Q dependence in the proton PDFs. A useful place to download the PDF data set that has Mathematica interface is `https://nnpdf.hepforge.org/old/html/mathematica.html`.



4.4.2 Proton beam: LHC cross sections

We have gained sufficient understanding of PDF to directly write down general cross sections at LHC (proton-proton collider), where at the fundamental level two partons scatter

$$\sigma(s) = \sum_{\text{A,B}} \int dx_A dx_B f_A(x_A) f_B(x_B) \hat{\sigma}_{AB}(x_A x_B s) \ , \tag{350}$$

where A, B goes through quarks, antiquarks and gluons. $s = (14\,\text{TeV})^2$ is the CM energy square of two proton beams, and $\hat{s} = (p_A + p_B)^2 \simeq 2p_A \cdot p_B \simeq x_A x_B s$ is the center-of-mass energy of two partons. This is a good approximation for $\hat{s} \gg \text{GeV}^2$ where the masses of proton and partons are all negligible.

The proton PDFs satisfy the following sum rules

$$\begin{aligned}
&\int_0^1 dx \left[f_u(x, Q^2) - f_{\bar{u}}(x, Q^2) \right] = 2 \ , \\
&\int_0^1 dx \left[f_d(x, Q^2) - f_{\bar{d}}(x, Q^2) \right] = 1 \ , \\
&\int_0^1 dx \left[f_q(x, Q^2) - f_{\bar{q}}(x, Q^2) \right] = 0 \ , \quad (q = s, c, b) \\
&\sum_A \int_0^1 dx x f_A(x, Q^2) = 1 \ .
\end{aligned} \tag{351}$$

The above general formation can be simplified for resonance production. i.e., the $2 \to 1$ process discussed in sec. 4.2.3. Examples include gluon fusion production of the Higgs boson, or $q\bar{q} \to Z$ production. Let's work with the latter process. Making analogy to Eq. (287), for $q\bar{q} \to Z$, we have

$$\sigma_{q\bar{q}\to Z} = \frac{1}{9} \times \frac{12\pi^2 \Gamma_{Z\to q\bar{q}}}{M_Z} \delta(\hat{s} - M_Z^2) = \frac{4\pi^2 \Gamma_{Z\to q\bar{q}}}{3M_Z} \delta(\hat{s} - M_Z^2) \ , \tag{352}$$

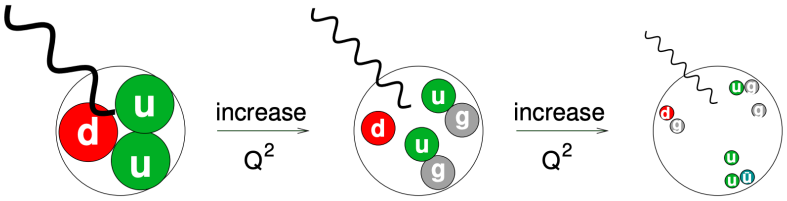
where the extra 1/9 comes from averaging the initial state quark, antiquark color degrees of freedom. Plugging this into Eq. (350), we get

$$\sigma_{pp\to Z}(s) = \int dx_A dx_B \left[f_q(x_A) f_{\bar{q}}(x_B) + f_{\bar{q}}(x_A) f_q(x_B)\right] \frac{4\pi^2 \Gamma_{Z\to q\bar{q}}}{3M_Z} \delta(x_A x_B s - M_Z^2)$$
$$= \frac{4\pi^2 \Gamma_{Z\to q\bar{q}}}{3M_Z} \frac{1}{s} \int_{M_Z^2/s}^1 \frac{dx_A}{x_A} \left[f_q(x_A) f_{\bar{q}}\left(\frac{M_Z^2/s}{x_A}\right) + f_{\bar{q}}(x_A) f_q\left(\frac{M_Z^2/s}{x_A}\right)\right] . \tag{353}$$

Here the scale factor in the PDF should be set to $Q \simeq M_Z$.

4.5 Deep inelastic scattering

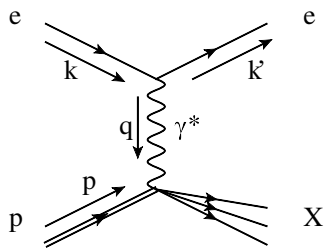
There are two ways of looking at the high energy collision processes discussed earlier with the proton beam. What we have described is in the lab frame where the proton sends out one of its partons to participate in a hard scattering with the second beam. Because both cross sections and PDFs are Lorentz invariant quantities, we can boost to be rest frame of the proton, and equivalently describe the above process as using the second beam to probe the internal structure of the proton, with momentum transfer Q. Using the uncertainty principle, Q^{-1} is the length scale that can be probed. Clearly, larger Q allows us to see the internal structure of the proton better (see the picture below). That explains why Q is called the resolution scale.



In this subsection, we consider fixed target experiments as the second way of probing the internal structure of proton. The experiment has an electron beam and at rest protons in the target material. We will discuss experimental evidences showing that the parton picture indeed works for high energy collisions. The PDFs can correctly describe the internal structure of proton in these circumstances. These experiments also provides a valuable way of measuring the PDFs.

The deep inelastic scattering process is illustrated by the following diagram, where proton at rest is struck by an energetic electron beam. This occurs through t-channel γ or Z exchange, where for the low-energy ($\sqrt{s} \sim$ few GeV) collision considered here, photon exchange dominates. After the collision, the electron gets deflected whereas proton breaks apart into a bunch of final state hadrons, denoted by X. The first experiment in this class is

the SLAC-MIT experiment, https://www.nobelprize.org/prizes/physics/1990/9594-the-slac-mit-experiment/.



4.5.1 Kinematics

First, let's work out the kinematics of the scattering, by introducing two variables

$$\begin{aligned} x &= \frac{Q^2}{2p\cdot q} \ , \\ \nu &= \frac{2p\cdot q}{m_p} \ , \end{aligned} \tag{354}$$

where $Q^2 = -q^2 > 0$. x is a dimensionless quantity. ν carries unit of energy. We work in the lab frame where the proton is at rest, $p^\mu = (m_p, 0, 0, 0)$. Using momentum conservation $q = k - k'$, we find ν is related to the energy transfer of the process,

$$\nu = 2(E_k - E_{k'}) \ . \tag{355}$$

On the other hand, the product of x and ν is related to the three-momentum transfer (scattering angle) of the process,

$$Q^2 = q^2 = (k-k')^2 \simeq 2k\cdot k' = 2E_k E_{k'}(1-\cos\theta) \ . \tag{356}$$

We could dig deeper on the meaning of x, by making connection with the parton model. The parton model states that for each high energy collision event, the electron only scatters with a parton inside the proton and does not care about the rest parts of proton. Here to interact with the photon, the parton must be a quark or antiquark. The parton can carry a fraction of the proton's four-momemtum, $\tilde{p}^\mu = \tilde{x}p^\mu = (\xi m_p, 0, 0, 0)$. The probability density for this to occur is given by the PDF $f_q(\xi, Q^2)$. Assuming the final state parton has four-momemtum $\tilde{p}'$, we have

$$\tilde{p}' = \tilde{p} + q \ . \tag{357}$$

Squaring both sides, we get

$$\tilde{p}'^2 = \tilde{p}^2 + q^2 + 2\tilde{p}\cdot q \ . \tag{358}$$

Working in the limit where $Q^2 = -q^2 \gg m_p^2$, we find

$$Q^2 \simeq 2\tilde{p} \cdot q \ , \quad \Rightarrow \quad \xi = \frac{Q^2}{2p \cdot q} \ . \tag{359}$$

Comparing with Eq. (354), we find

$$x = \xi \ . \tag{360}$$

In other words, if the parton model works, the kinematic variable x would tell the fraction of energy carried by the parton in the collision!

4.5.2 Testing the parton model

It is time to work out the differential cross section of the process. We first do it without assuming the parton model. The matrix element is

$$i\mathcal{M} = \bar{u}(k')(-ie\gamma^\mu)u(k)\frac{-ig_{\mu\nu}}{q^2}\langle X|(-ieJ^\mu)|p\rangle \ , \tag{361}$$

where J^μ is made of quark vector current that sees the photon. Without the possibility of knowing the detailed final state X, we cannot work out $\langle X|J^\mu|p\rangle$ but try to proceed with this symbolic notation.

The spin averaged matrix element square is

$$\begin{aligned} \overline{|\mathcal{M}|^2} &= \frac{1}{4}\frac{e^4}{q^4}\mathrm{Tr}\left[(\not{k}' + m_e)\gamma^\mu(\not{k} + m_e)\gamma^\nu\right]\langle p|J_\mu|X\rangle\langle X|J_\nu|p\rangle \\ &= \frac{e^4}{q^4} \times L^{\mu\nu} \times \frac{1}{2}\langle p|J_\mu|X\rangle\langle X|J_\nu|p\rangle \end{aligned} \tag{362}$$

where electron mass is neglected in the second step, and

$$L^{\mu\nu} = 2(k^\mu k'^\nu + k^\nu k'^\mu - k \cdot k' g^{\mu\nu}) \ . \tag{363}$$

The inclusive $ep \to eX$ cross section is

$$\begin{aligned} \sigma_{ep\to eX} &= \frac{1}{4f}\int \frac{d^3\vec{k}'}{(2\pi)^3 2E_{k'}}\sum_X \overline{|\mathcal{M}|^2}(2\pi)^4\delta^4(p + k - k' - p_X) \\ &= \frac{1}{4f}\int \frac{d^3\vec{k}'}{(2\pi)^3 2E_{k'}}\frac{e^4}{q^4}L^{\mu\nu}(4\pi)W_{\mu\nu} \ , \end{aligned} \tag{364}$$

where $\sum_X$ includes the sum over all possible X final states and contains the corresponding phase space integrals. In the last step we defined

$$W_{\mu\nu} = \frac{1}{4\pi}\sum_X \frac{1}{2}\langle p|J_\mu|X\rangle\langle X|J_\nu|p\rangle(2\pi)^4\delta^4(p + k - k' - p_X) \ . \tag{365}$$

There will be enough integrals to kill all the δ functions.

Before worrying about $W_{\mu\nu}$, we first simplify the electron final state phase space integral, by introducing another kinematic variable

$$y = \frac{q\cdot p}{k\cdot p} = \frac{q^0}{E_k} = \frac{E_k - E_{k'}}{E_k} \ . \tag{366}$$

Again, we used $p^\mu = (m_p, 0, 0, 0)$. As a result, the k' phase space integral can be rewritten as

$$\begin{aligned}
\int \frac{d^3\vec{k}'}{(2\pi)^3 2E_{k'}} &= \frac{1}{8\pi^3}\int d\phi \int d\cos\theta \int k'^2 dk' \frac{1}{2E_{k'}} \\
&\simeq \frac{1}{8\pi^2}\int d\cos\theta \int E_{k'} dE_{k'} \\
&= \frac{1}{8\pi^2}\frac{m_p}{E_k}\int dx \int (E_k - E_{k'}) d(E_k - E_{k'}) \\
&= \frac{1}{8\pi^2} m_p E_k \int_0^1 dx \int_0^1 y dy \ ,
\end{aligned} \tag{367}$$

where $k' = |\vec{k}'|$ and we work in the massless electron limit such that $E_{k'} = k'$. In the third step, we used

$$x = \frac{Q^2}{2p\cdot q} = \frac{E_k E_{k'}(1-\cos\theta)}{m_p(E_k - E_{k'})} \ , \quad \Rightarrow \quad d\cos\theta = -\frac{m_p(E_k - E_{k'})dx}{E_k E_{k'}} \ . \tag{368}$$

For the electron fixed target collision, we have $4f = 4E_k m_p$. As a result,

$$\frac{1}{4f}\int \frac{d^3\vec{k}'}{(2\pi)^3 2E_{k'}} = \frac{1}{32\pi^2}\int_0^1 dx \int_0^1 y dy \ , \tag{369}$$

and we can write the cross section as

$$\sigma_{ep\to eX} = 2\pi\alpha^2 \int_0^1 dx \int_0^1 y dy \frac{1}{q^4} L^{\mu\nu} W_{\mu\nu} \ . \tag{370}$$

We could express $W_{\mu\nu}$ in terms of form factors, keeping in mind that it is a function of p^μ and q^μ. The Ward identity states that in QED photon must always couple to a conserved current, thus $q^\mu W_{\mu\nu} = 0$. The general form of $W_{\mu\nu}$ is

$$W_{\mu\nu} = \left(-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2}\right) F_1(x, Q^2) - \left(\frac{p_\mu}{p\cdot q} - \frac{q_\mu}{q^2}\right)\left(\frac{p_\nu}{p\cdot q} - \frac{q_\nu}{q^2}\right)(p\cdot q) F_2(x, Q^2) \ , \tag{371}$$

where $F_{1,2}$ are two dimensionless form factors (structure functions) characterizing the internal structure of proton. They must depend on the scalar products made out of p^μ and q^μ. The non-trivial ones are $p\cdot q$ and q^2, which are replaced by $x = Q^2/(2p\cdot q)$ and $Q^2 = -q^2$.

The contraction between $L^{\mu\nu}$ and $W_{\mu\nu}$ can now be worked out in terms of the form factors. Finally, we get

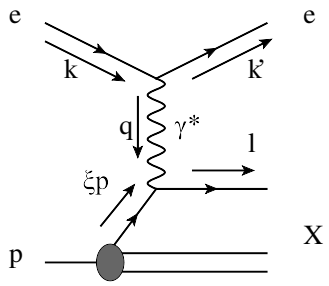
$$\frac{d\sigma_{ep\to eX}}{dxydy} = \frac{4\pi\alpha^2}{q^4}\left\{q^2 F_1(x,Q^2) - \left[(1-y)\frac{q^2}{xy^2} - xm_p^2\right]F_2(x,Q^2)\right\} \ . \quad (372)$$

We could measure the two structure functions experimentally by exploring the differential cross section.

Next, we work in the parton model and consider the scatterings of electron with the fundamental partons inside the proton, which are quarks (u,d) and anti-quarks $(\bar{u},\bar{d})$. We apply the master formula Eq. (353) (neglecting QED corrections, only the proton beam is composite),

$$\sigma_{ep\to eX} = \sum_q \int d\xi \left[f_q(\xi,Q^2)\sigma_{eq\to eq} + f_{\bar{q}}(\xi,Q^2)\sigma_{e\bar{q}\to e\bar{q}}\right] \ , \quad (373)$$

where the parton level cross sections $\sigma_{eq\to eq}$ and $\sigma_{e\bar{q}\to e\bar{q}}$ are calculated assuming the initial state parton carries four-momentum $\xi p^\mu = (\xi m_p, 0, 0, 0)$.



It is straightforward to compute the $2\to 2$ cross section for $e(k)+q(\xi p)\to e(k')+q(\ell)$ (as shown by the above Feynman diagram),

$$\begin{aligned}
\sigma_{eq\to eq} &= \frac{1}{4\xi E_k m_p}\int\frac{d^3\vec{k}'}{(2\pi)^3 2E_{k'}}\int\frac{d^3\vec{\ell}}{(2\pi)^3 2E_\ell}(2\pi)^4\delta^4(k+\xi p-k'-\ell)\frac{e^4Q_q^2}{q^4}L^{\mu\nu} \\
&\qquad\qquad \times 2\xi(p_\mu\ell_\nu + p_\nu\ell_\mu - p\cdot\ell g_{\mu\nu}) \\
&= \frac{1}{4E_k m_p}\int\frac{d^3\vec{k}'}{(2\pi)^3 2E_{k'}}\frac{1}{2E_\ell}(2\pi)\delta(E_k - E_{k'} + \xi m_p - E_\ell)\frac{e^4Q_q^2}{q^4}L^{\mu\nu} \\
&\qquad\qquad \times 2(p_\mu\ell_\nu + p_\nu\ell_\mu - p\cdot\ell g_{\mu\nu}) \\
&= \frac{1}{4E_k m_p}\int\frac{d^3\vec{k}'}{(2\pi)^3 2E_{k'}}(2\pi)\delta\left(E_\ell^2 - (E_k - E_{k'} + \xi m_p)^2\right)\frac{e^4Q_q^2}{q^4}L^{\mu\nu} \\
&\qquad\qquad \times 2(p_\mu\ell_\nu + p_\nu\ell_\mu - p\cdot\ell g_{\mu\nu}) \ .
\end{aligned} \quad (374)$$

Because the final state parton is also on-shell, we have

$$E_\ell^2 = |\vec{\ell}|^2 + m_q^2 \simeq |\vec{\ell}|^2 = |\vec{q}+\xi\vec{p}|^2 \ . \quad (375)$$

This allows us to rewrite the remaining δ function as

$$\begin{aligned}\delta\left(|\vec{q}+\xi\vec{p}|^2-(q^0+\xi m_p)^2\right) &= \delta\left((q+\xi p)^2\right)\\ &= \delta\left(q^2+2\xi p\cdot q+\xi^2 m_p^2\right)\\ &\simeq \delta\left(q^2+2\xi p\cdot q\right)\\ &= \frac{1}{2p\cdot q}\delta\left(\xi+\frac{q^2}{2p\cdot q}\right)\\ &= \frac{-x}{q^2}\delta\left(\xi-x\right)\ .\end{aligned} \tag{376}$$

Plugging this back to the above parton level cross section, we find

$$\sigma_{eq\to eq}=\frac{1}{4E_k m_p}\int\frac{d^3\vec{k}'}{(2\pi)^3 2E_{k'}}\frac{2\pi x}{-q^2}\delta(\xi-x)\frac{e^4Q_q^2}{q^4}L^{\mu\nu}2(p_\mu\ell_\nu+p_\nu\ell_\mu-p\cdot\ell g_{\mu\nu})\ . \tag{377}$$

It contributes to the proton level cross section as

$$\begin{aligned}\sigma_{ep\to eX} &\ni \int d\xi f_q(\xi,Q^2)\sigma_{eq\to eq}\\ &= \frac{1}{4E_k m_p}\int\frac{d^3\vec{k}'}{(2\pi)^3 2E_{k'}}\frac{e^4}{q^4}L^{\mu\nu}\frac{4\pi Q_q^2 x}{-q^2}(p_\mu\ell_\nu+p_\nu\ell_\mu-p\cdot\ell g_{\mu\nu})f_q(x,Q^2)\ .\end{aligned} \tag{378}$$

As expected, the δ-function ensures $\xi=x$. This can be used to compare with the last line of Eq. (364). Matching the two, we obtain

$$W_{\mu\nu}=\frac{Q_q^2 x}{-q^2}(p_\mu\ell_\nu+p_\nu\ell_\mu-p\cdot\ell g_{\mu\nu})f_q(x,Q^2)\ . \tag{379}$$

Recall that $\ell^\mu=xp^\mu+q^\mu$. We have

$$p\cdot\ell=xm_p^2+p\cdot q\simeq p\cdot q=-\frac{q^2}{2x}\ , \tag{380}$$

and

$$\begin{aligned}x(p_\mu\ell_\nu+p_\nu\ell_\mu-p\cdot\ell g_{\mu\nu}) &= x\left(xp_\mu p_\nu+p_\mu q_\nu+xp_\mu p_\nu+p_\nu q_\mu+\frac{q^2}{2x}g_{\mu\nu}\right)\\ &= x\left[\frac{q^2}{2x}\left(g_{\mu\nu}-\frac{q_\mu q_\nu}{q^2}\right)+\frac{1}{2x}q_\mu q_\nu+p_\mu q_\nu+p_\nu q_\mu+2xp_\mu p_\nu\right]\\ &= \frac{q^2}{2}\left(g_{\mu\nu}-\frac{q_\mu q_\nu}{q^2}\right)+x\left[-\frac{p\cdot q}{q^2}q_\mu q_\nu+p_\mu q_\nu+p_\nu q_\mu-\frac{q^2}{p\cdot q}p_\mu p_\nu\right]\\ &= \frac{q^2}{2}\left(g_{\mu\nu}-\frac{q_\mu q_\nu}{q^2}\right)-xq^2(p\cdot q)\left[\frac{q_\mu}{q^2}\frac{q_\nu}{q^2}-\frac{p_\mu}{p\cdot q}\frac{q_\nu}{q^2}-\frac{p_\nu}{p\cdot q}\frac{q_\mu}{q^2}+\frac{p_\mu}{p\cdot q}\frac{p_\nu}{p\cdot q}\right]\\ &= \frac{q^2}{2}\left(g_{\mu\nu}-\frac{q_\mu q_\nu}{q^2}\right)-xq^2(p\cdot q)\left(\frac{p_\mu}{p\cdot q}-\frac{q_\mu}{q^2}\right)\left(\frac{p_\nu}{p\cdot q}-\frac{q_\nu}{q^2}\right)\ .\end{aligned} \tag{381}$$

They lead to

$$W_{\mu\nu} = -\frac{Q_q^2}{2} f_q(x,Q^2) \left(g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) + x Q_q^2 f_q(x,Q^2)(p\cdot q) \left(\frac{p_\mu}{p\cdot q} - \frac{q_\mu}{q^2} \right) \left(\frac{p_\nu}{p\cdot q} - \frac{q_\nu}{q^2} \right) \ . \tag{382}$$

Comparing this with Eq. (371) where the form factors are introduced, we obtain

$$\begin{aligned} F_1(x,Q^2) &= \sum_q \frac{Q_q^2}{2} \left[f_q(x,Q^2) + f_{\bar{q}}(x,Q^2) \right] \ , \\ F_2(x,Q^2) &= x \sum_q Q_q^2 \left[f_q(x,Q^2) + f_{\bar{q}}(x,Q^2) \right] \ . \end{aligned} \tag{383}$$

The parton model implies a close relationship between the two form factors

$$F_2(x,Q^2) = 2x F_1(x,Q^2) \ , \tag{384}$$

which has been tested experimentally.

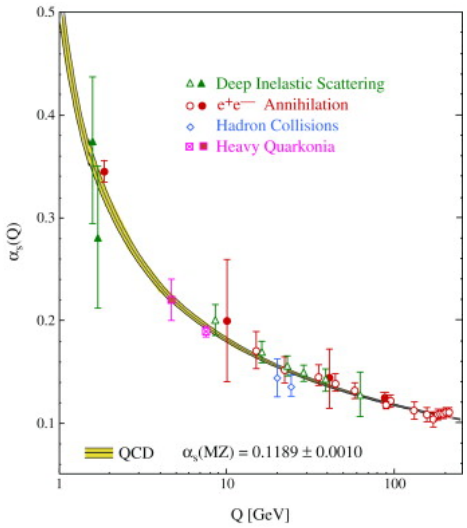
Another remarkable observation is the Q dependence in the PDFs are only logarithmic. It in turn implies that the form factors only mildly depend on Q. Neglecting their Q dependence, the form factors only depend on x. This feature is called Bjorken scaling. In general, there is no reason to expect such a feature to be there. It is a unique prediction of the parton model.

* * *

In this section, we only touched the basics of the parton model and discussed its application to high-energy hadron colliders. We have left many important aspects of perturbative QCD untouched, such as final state radiations and jet physics. You are referred to this very nice lecture note to explore further, `https://arxiv.org/pdf/0910.4182.pdf`.

5 QCD at Low Energies

A novel property of strong interaction is asymptotic free. The value strong coupling α_3 depends on the energy scale of the hard scattering process Q. It decreases as Q gets larger. For Q larger than several GeV, α_3 becomes small enough so that perturbation theory works, where we can treat QCD in a similar fashion as QED. We did this when discussing deep inelastic scattering and high-energy collider experiments. On the other hand, as Q goes below a few GeV, the strong coupling becomes very large, as shown by the plot below.



The running of α_3 is governed by the QCD β function,

$$\frac{d\alpha_3}{d\log Q} = -\left(11 - \frac{2n_f}{3}\right)\frac{\alpha_3^2}{2\pi} + \cdots \,, \tag{385}$$

where $\cdots$ are terms of higher power in α_3. They are not negligible when α_3 becomes order 1 or larger. The scale where α_3 becomes infinity is an intrinsic scale the QCD. It is called the QCD scale, $\Lambda_{\rm QCD} \simeq 200\,{\rm MeV}$. Around this scale, two things happen to the quarks and gluons which participate in the strong interaction.

One is confinement, where quarks, antiquarks and gluons are no longer the free particles for constructing the S matrix at low energies, or at length scales above a fm. Instead, they confine into colour singlet hadrons. The simples states include mesons, made of $(q^i_m \bar{q}^j_n)\delta_{ij}$, and baryons, made of $(q^i_m q^j_n q^k_l)\varepsilon_{ijk}$. Here we use the same convention as introduced in section 2.1, where i, j, k are color indices and m, n, l are flavor indices. The top quark is an exception. It decay so fast that there is no time for it to form any hadron states. Gluons will form massive glueballs.

In this section, we mainly deal with mesons and baryons made of light quarks, u, d, s. Heavy quark mesons or and baryons (containing c, b) do exist. Because the masses of c, b lie above GeV scale, the value of α_s is border line perturbative. As a result, heavy quarkonia such as $\Upsilon = (b\bar{b})$ and $J/\psi = (c\bar{c})$ can be described by non-relativistic effective theory of QCD, similar to the positronium described by NRQED. For mesons containing one heavy quark and one light quark, because the heavy quark mass is much higher than the QCD scale, we can picture them as a cloud of light antiquark surrounding a non-relativistic core of heavy quark. The size of the cloud is of order $\Lambda_{\rm QCD}^{-1}$. There exist heavy quark effective theories (HQET) to deal with such a system. In contrast, hadrons made of only light quarks are tightly bounded states with masses around or below GeV scale.

5.1 Chiral symmetries and the eightfold way

Our first goal is to find an effective Lagrangian for the lightest mesons made of light quarks and antiquarks, using symmetries (and their breaking) as the guiding principle. It will turn out to be highly nontrivial. Let's look at the strong interaction Lagrangian involving the three light quarks,

$$\mathcal{L}_{\rm QCD} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} + \sum_{q=u,d,s} (\bar{q} i \not{D} q - m_q \bar{q} q) \ . \tag{386}$$

In the limit $m_q \to 0$, the Lagrangian has a large global symmetry

$$SU(3)_L \times SU(3)_R \times U(1)_B \ . \tag{387}$$

The $SU(3)_L$ and $SU(3)_R$ are chiral symmetries describing unitary rotations among the three flavours of left-handed and right-handed quarks, respectively. The $U(1)_B$ symmetry describe making a common phase redefinition to all the quarks fields (both left and right). It is the baryon number symmetry!

Another $U(1)_A$ axial symmetry that makes a common phase redefinition to all left-handed quarks, and an opposite phase to all right-handed quarks can also leave the above Lagrangian invariant. However, at quantum level, it receives gauge anomaly with respect to $SU(3)_c^2$. Although it is OK for a global symmetry to be anomalous, the $U(1)_A$ is badly broken and not a suitable symmetry to start with. [4]

The non-zero light quark masse breaks the $SU(3)_L \times SU(3)_R$ symmetries explicitly, but only slightly. These symmetries are also not compatible with QED, due to different electric charges of u and d, s. Because α is small, these symmetries are only slightly broken.

The most significant symmetry breaking effect is actually spontaneous, due to the second thing that occurs for QCD below GeV scale. It is called chiral symmetry breaking, where colour-singlet quark-antiquark vacuum condensates form

$$\langle \bar{u} u \rangle = \langle \bar{d} d \rangle = \langle \bar{s} s \rangle \ . \tag{388}$$

[4] Without the $U(1)_A$ being anomalous, we would have the ninth pseudo-Goldstone boson η'. In reality, η' is heavy thanks to the $U(1)_A$ anomaly.

The values of these condensates are about $(300\,\mathrm{MeV})^3$. They are much higher than the light quark masses. Thus we first work in the limit of $m_q \to 0$ and $\alpha = 0$. In this case, the quark condensates spontaneously breaks $SU(3)_L \times SU(3)_R$. The baryon symmetry is still good because the quark condensate is neutral in $U(1)_B$.

To see what $SU(3)_L \times SU(3)_R$ breaks down to, we rewrite the quark bilinear fields in a matrix form

$$\begin{aligned} \mathcal{M} &= \begin{pmatrix} \bar{u}_R u_L & \bar{d}_R u_L & \bar{s}_R u_L \\ \bar{u}_R d_L & \bar{d}_R d_L & \bar{s}_R d_L \\ \bar{u}_R s_L & \bar{d}_R s_L & \bar{s}_R s_L \end{pmatrix} \\ &= \begin{pmatrix} u_L \\ d_L \\ s_L \end{pmatrix} \begin{pmatrix} \bar{u}_R & \bar{d}_R & \bar{s}_R \end{pmatrix} \ , \end{aligned} \tag{389}$$

where it understood that when q_L and $\bar{q}'_R$ are put together, they automatically switch position and form a Lorentz scalar.

Under the $SU(3)_L \times SU(3)_R$ chiral symmetry transformations,

$$\begin{pmatrix} u_L \\ d_L \\ s_L \end{pmatrix} \to V_L \begin{pmatrix} u_L \\ d_L \\ s_L \end{pmatrix} \ , \quad \begin{pmatrix} u_R \\ d_R \\ s_R \end{pmatrix} \to V_R \begin{pmatrix} u_R \\ d_R \\ s_R \end{pmatrix} \ , \tag{390}$$

where $V_{L,R}$ are unitary matrices, the matrix $\mathcal{M}$ transforms as

$$\mathcal{M} \to V_L \mathcal{M} V_R^\dagger \ . \tag{391}$$

The vacuum condensate of quarks corresponds to the diagonal elements of $\mathcal{M}$ plus its Hermitian conjugation, and

$$\langle \mathcal{M} \rangle = \begin{pmatrix} \langle \bar{u}_R u_L \rangle & & \\ & \langle \bar{d}_R d_L \rangle & \\ & & \langle \bar{s}_R s_L \rangle \end{pmatrix} = \langle \bar{q}_R q_L \rangle \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \ . \tag{392}$$

Clearly, $\langle \mathcal{M} \rangle$ is still invariant under joint chiral transformations with $V_L = V_R$. This shows that the quark condensates (equal for three flavours) spontaneously breaks $SU(3)_L \times SU(3)_R$ down to a diagonal one $SU(3)_V$. Because $SU(3)$ has 8 generators, the spontaneous (global) symmetry breaking leads to 8 massless Goldstone bosons. They are excitations on top of $\langle \mathcal{M} \rangle$. Each of them is a colour singlet bound state of a quark and antiquark.

Using nonlinear realization, we can write

$$\mathcal{M} = e^{i\Sigma/F_\pi} \langle \mathcal{M} \rangle = \langle \bar{q}_R q_L \rangle e^{i\Sigma/F_\pi} \ , \tag{393}$$

where all the Golstones are inside the 3×3 matrix Σ,

$$\Sigma = \begin{pmatrix} \pi^0 + \frac{1}{\sqrt{3}}\eta & \sqrt{2}\pi^+ & \sqrt{2}K^+ \\ \sqrt{2}\pi^- & -\pi^0 + \frac{1}{\sqrt{3}}\eta & \sqrt{2}K^0 \\ \sqrt{2}K^- & \sqrt{2}\bar{K}^0 & -\frac{2}{\sqrt{3}}\eta \end{pmatrix} \ . \tag{394}$$

A fermion-anti-fermion system has parity $P = (-1)^{L+1}$ where L is orbital angular momentum. The lowest energy state has $L = 0$. Therefore, the above Goldstone bosons are pseudo-scalar mesons. Their compositions are shown by the following ket states

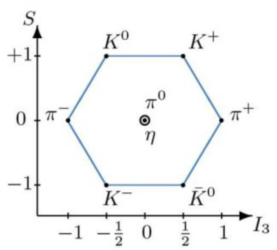
$$\begin{aligned}
\pi^0 &= \frac{|u\bar{u}\rangle - |d\bar{d}\rangle}{\sqrt{2}} \equiv \phi_3 \ , \\
\eta &= \frac{|u\bar{u}\rangle + |d\bar{d}\rangle - 2|s\bar{s}\rangle}{\sqrt{6}} \equiv \phi_8 \ , \\
\pi^+ &= |u\bar{d}\rangle \equiv \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2) \ , \\
\pi^- &= |d\bar{u}\rangle \equiv \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \ , \\
K^+ &= |u\bar{s}\rangle \equiv \frac{1}{\sqrt{2}}(\phi_4 - i\phi_5) \ , \\
K^- &= |s\bar{u}\rangle \equiv \frac{1}{\sqrt{2}}(\phi_4 + i\phi_5) \ , \\
K^0 &= |d\bar{s}\rangle \equiv \frac{1}{\sqrt{2}}(\phi_6 - i\phi_7) \ , \\
\overline{K}^0 &= |s\bar{d}\rangle \equiv \frac{1}{\sqrt{2}}(\phi_6 + i\phi_7) \ .
\end{aligned} \tag{395}$$

These particles are the lightest among all the hadrons. They are the building blocks of the low energy effective theory of QCD.

Using the ϕ_i fields, we can also write Σ in terms of the Gell-Mann matrices,

$$\Sigma = \sum_{i=1}^{8} \phi_i \lambda_i \ . \tag{396}$$

The eight ϕ_i fields form an adjoint representation of the broken $SU(3)$. They are often presented in the following eight-fold way.



There are two diagonal Gell-Man matrices, λ_3 and λ_8 (see Eq. (99)). Acting on the light quark fields, they define the isospin and strangeness quantum numbers, where $I_3(u) = 1/2$, $I_3(d) = -1/2$, $I_3(s) = 0$. Only s has strangeness which is -1. The quantum numbers for corresponding antiquarks are the opposite. These numbers are additive when quarks and antiquark form mesons.

5.2 Effective Lagrangian for pions

Next questions, what is the effective Lagrangian for the above light pseudo-scalar mesons, and how come they end up being massive? We make close analogy to the non-linear σ model. See first line of Eq. (129), but drop the radial mode which is heavy. We define

$$U = \exp(i\Sigma/F_\pi) \ , \tag{397}$$

where $F_\pi = 93\,\text{MeV}$ is called the pion decay constant. Comparing with Eq. (393), we know that

$$\mathcal{M} = \langle \bar{q}_R q_L \rangle U = \frac{\langle \bar{q} q \rangle}{2} U \ , \tag{398}$$

where we used $\langle \bar{q}_R q_L \rangle = \langle \bar{q}_L q_R \rangle = \frac{1}{2}\langle \bar{q} q \rangle$. QCD preserves parity.

The leading Lagrangian for the Goldstone bosons is

$$\mathcal{L}_{\text{eff,kinetic}} = \frac{F_\pi^2}{4} \text{Tr}\left(\partial_\mu U \partial^\mu U^\dagger\right) \ . \tag{399}$$

Taylor expanding each U up to linear order in Σ, and using Eq. (396), we get the canonical kinetic terms for ϕ_i's,

$$\mathcal{L}_{\text{eff,kinetic}} \supset \frac{1}{4}\text{Tr}\left(\partial_\mu \Sigma \partial^\mu \Sigma\right) = \frac{1}{4}\text{Tr}\left(\lambda^a \lambda^b\right) \partial_\mu \phi_a \partial^\mu \phi_b = \frac{1}{2} \partial_\mu \phi_a \partial^\mu \phi_a \ , \tag{400}$$

where we used $\text{Tr}\left(\lambda^a \lambda^b\right) = 2\delta^{ab}$ for Gell-Mann matrices.

Going beyond quadratic order, we find all terms from $\mathcal{L}_{\text{eff}}$ contain derivatives, thus none of them are responsible for the mass generation of the Goldstone bosons. The key for their masses is due to the quark mass terms in the QCD Lagrangian, Eq. (386). We rewrite them as

$$\mathcal{L}_{\text{quark masses}} = \text{Tr}\left[\begin{pmatrix} m_u & & \\ & m_d & \\ & & m_s \end{pmatrix} \begin{pmatrix} u_L \\ d_L \\ s_L \end{pmatrix} \begin{pmatrix} \bar{u}_R & \bar{d}_R & \bar{s}_R \end{pmatrix}\right] + \text{h.c.} = \text{Tr}\left[M_q \mathcal{M}\right] + \text{h.c.} \ , \tag{401}$$

where we used Eq. (389) and defined

$$M_q = \begin{pmatrix} m_u & & \\ & m_d & \\ & & m_s \end{pmatrix} \ . \tag{402}$$

Further using Eq. (398) we find the effective Lagrangian for the quark mass term,

$$\mathcal{L}_{\text{eff, mass}} = \frac{\langle \bar{q} q \rangle}{2} \text{Tr}\left[M_q U\right] + \text{h.c.} \ . \tag{403}$$

Expanding U to quadratic order in Σ, we get the mass term (note the term

linear in Σ vanishes due to the Hermitian conjugate),

$$\begin{aligned}\mathcal{L}_{\text{eff, mass}} \supset & -\frac{\langle\bar{q}q\rangle}{2F_\pi^2}\text{Tr}\left[M_q\Sigma\Sigma\right] \\ = & \frac{\langle\bar{q}q\rangle}{F_\pi^2}\left[\frac{1}{2}(m_u+m_d)(\pi^0)^2+(m_u+m_d)\pi^+\pi^- + (m_u+m_s)K^+K^-\right. \\ & \left. +(m_d+m_s)K^0\overline{K}^0+\frac{1}{6}(m_u+m_d+4m_s)\eta^2+\frac{m_u-m_d}{\sqrt{3}}\pi^0\eta\right] .\end{aligned} \tag{404}$$

Because the quark mass terms break the $SU(3)_L \times SU(3)_R$ chiral symmetry explicitly, the pseudoscalar mesons are pseudo-Goldstone bosons. Their masses are (neglecting the π^0-η mixing term, justified because $m_{u,d} \ll m_s$)

$$\begin{aligned} m_{\pi^0}^2 \simeq m_{\pi^\pm}^2 &= \frac{\langle\bar{q}q\rangle}{F_\pi^2}(m_u+m_d) , \\ m_{K^\pm}^2 &= \frac{\langle\bar{q}q\rangle}{F_\pi^2}(m_u+m_s) , \\ m_{K^0}^2 = m_{\overline{K}^0}^2 &= \frac{\langle\bar{q}q\rangle}{F_\pi^2}(m_d+m_s) , \\ m_\eta^2 &\simeq \frac{\langle\bar{q}q\rangle}{3F_\pi^2}(m_u+m_d+4m_s) . \end{aligned} \tag{405}$$

Defining the averaged kaon mass

$$m_K^2 = \frac{1}{2}(m_{K^\pm}^2+m_{K^0}^2) = \frac{\langle\bar{q}q\rangle}{F_\pi^2}\left(\frac{m_u+m_d}{2}+m_s\right) , \tag{406}$$

we find the Gell-Mann Okubo formula that relates different meson masses

$$m_\eta^2 \simeq \frac{1}{3}(4m_K^2-m_\pi^2) , \tag{407}$$

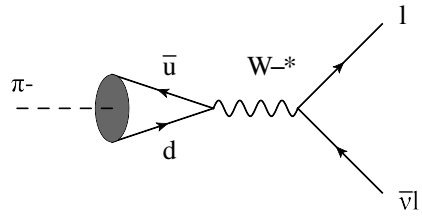
which works well experimentally.

The sum of Eqs. (399) and (403) are the leading terms in the effective chiral Lagrangian,

$$\mathcal{L}_{\text{eff}} = \frac{F_\pi^2}{4}\text{Tr}\left(\partial_\mu U\partial^\mu U^\dagger\right)+\frac{\langle\bar{q}q\rangle}{2}\text{Tr}\left[M_qU\right]+\text{h.c.} \tag{408}$$

5.3 Charged pion decay

$\pi^\pm$ are the lightest mesons states that carries electric charge. If the Standard Model only has QED+QCD gauge interactions but no weak interactions, $\pi^\pm$ would be stable. In the presence of weak interaction, charged pions can decay via the following Feynman diagram. It occurs through a virtual W boson exchange. The final state particles are $\mu^-\bar{\nu}_\mu$ or $e^-\bar{\nu}_e$.



Because the W boson is much heavier than the pion, we can integrate it out the work with the effective Lagrangian

$$\mathcal{L}_{\text{eff}} = -2\sqrt{2}G_F V_{ud}^{\text{CKM}} \left(\bar{\ell}\gamma^\mu \mathbb{P}_L \nu_\ell\right) \left(\bar{u}\gamma^\mu \mathbb{P}_L d\right) \ , \tag{409}$$

where V_{ud}^{CKM} is the CKM matrix element, and numerically very close to 1. The most important step here is to match the $d\bar{u}$ quark states to the pion. This is achieved using the partially conserved axial current (PCAC) relation. At matrix element level, it states

$$\langle 0|\bar{u}\gamma^\mu\gamma_5 d|\pi^-(q)\rangle = -i\sqrt{2}F_\pi q^\mu e^{-iq\cdot x} \ , \tag{410}$$

where $F_\pi = 93\,\text{MeV}$. Physically, we can treat an axial current made of fundamental quarks as a pion field and use it to annihilate a pion state. In coordinate space, the matching works as

$$\bar{u}\gamma^\mu\gamma_5 d \leftrightarrow \sqrt{2}F_\pi \partial^\mu \pi^-(x) \ . \tag{411}$$

The vector current counterpart $\langle 0|\bar{u}\gamma^\mu d|\pi^-(q)\rangle$ is zero because parity does not work.

The PCAC relation allows us to match Eq. (409) further to the low-energy effective Lagrangian in the language of pion

$$\mathcal{L}_{\text{eff}} = -\sqrt{2}G_F F_\pi V_{ud}^{\text{CKM}} \left(\bar{\ell}\gamma^\mu \mathbb{P}_L \nu_\ell\right) \partial^\mu \pi^- \ . \tag{412}$$

With this, we can readily write down the decay amplitude of charged pion, $\pi^-(P) \to \ell^-(p_1) + \bar{\nu}_\ell(p_2)$,

$$\begin{aligned} i\mathcal{M} &= -i\sqrt{2}G_F F_\pi \left[\bar{u}(p_1,s_1)\gamma^\mu \mathbb{P}_L v(p_2,s_2)\right](-iP_\mu) \\ &= -\sqrt{2}G_F F_\pi \left[\bar{u}(p_1,s_1)\gamma^\mu \mathbb{P}_L v(p_2,s_2)\right](p_1+p_2)_\mu \\ &= -\sqrt{2}G_F F_\pi m_\ell \left[\bar{u}(p_1,s_1)\mathbb{P}_L v(p_2,s_2)\right] \ , \end{aligned} \tag{413}$$

where in the last step we used equation of motion for final state spinors, and the fact that neutrino is massless. We also set V_{ud}^{CKM} to 1. Clearly, the decay amplitude is proportional to the charged lepton mass m_ℓ. It is the famous helicity suppression in leptonic charged meson decays.

The amplitude square is

$$\begin{aligned} |\mathcal{M}|^2 &= 2G_F^2 F_\pi^2 m_\ell^2 \mathrm{Tr}\left[(\not{p}_1 + m_\ell)\mathbb{P}_L \not{p}_2 \mathbb{P}_R\right] \\ &= 8G_F^2 F_\pi^2 m_\ell^2 (p_1 \cdot p_2) \\ &= 4G_F^2 F_\pi^2 m_\ell^2 (m_\pi^2 - m_\ell^2) \ . \end{aligned} \tag{414}$$

The resulting decay rate is

$$\Gamma_{\pi^- \to \ell^- \bar{\nu}_\ell} = \frac{1}{4\pi} G_F^2 F_\pi^2 m_\ell^2 m_\pi \left(1 - \frac{m_\ell^2}{m_\pi^2}\right)^2 \ . \tag{415}$$

Due to the charge lepton mass dependence, π^- prefers to decay into $\mu^- \bar{\nu}_\mu$ over into $e^- \bar{\nu}_e$.

Using $m_\pi^- = 139.6\,\mathrm{MeV}$, $m_\mu = 105\,\mathrm{MeV}$, we find charged pion decay rate to be $2.7 \times 10^{-17}\,\mathrm{GeV}$. The lifetime is $\tau_{\pi^\pm} = 2.4 \times 10^{-8}\,\mathrm{sec}$. With such a lifetime, boosted charged pions appear like stable particles at high-energy collider experiments.

The decay branching ratio of π^- into $\mu^- \bar{\nu}_\mu$ is almost 100%. The decay branching ratio into $e^- \bar{\nu}_e$ is approximately

$$\mathrm{Br}_{\pi^- \to e^- \bar{\nu}_e} = \frac{\Gamma_{\pi^- \to e^- \bar{\nu}_e}}{\Gamma_{\pi^- \to \mu^- \bar{\nu}_\mu}} \simeq 1.2 \times 10^{-4} \sim \frac{m_e^2}{m_\mu^2} \ . \tag{416}$$

Another useful branching ratio to know is

$$\mathrm{Br}_{\pi^- \to \mu^- \bar{\nu}_\mu \gamma} \simeq 2 \times 10^{-4} \sim \frac{\alpha}{4\pi} \ , \tag{417}$$

which $1/(16\pi^2)$ comes from additional particle in the final state phase space.

5.3.1 Three-body charged pion decay

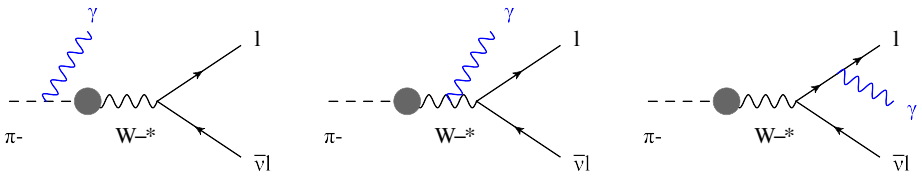
It is an interesting exercise to think about the $\pi^- \to e^- \bar{\nu}_e \gamma$ decay. Naively, it is a three-body decay and the usual helicity suppression argument does not apply here. One would expect it to have a branching ratio comparable to that of the $\pi^- \to \mu^- \bar{\nu}_\mu \gamma$ decay given in Eq. (417). However, the experimentally measured value is

$$\mathrm{Br}_{\pi^- \to e^- \bar{\nu}_e \gamma} \simeq 7.4 \times 10^{-7} \ll \mathrm{Br}_{\pi^- \to \mu^- \bar{\nu}_\mu \gamma} \ , \tag{418}$$

which indicates that the electron mass square suppression is still at work.

The fundamental reason for the appearance of m_e^2 in this decay rate is the Ward identity, which effectively allows us to integrate by parts in Eq. (412) and apply equations of motion, even if the fermion fields are the internal lines of a Feynman diagram and off-shell. In this subsection, we provide a proof of it.

At fundamental level, there are three ways to radiate a photon from the π^- decay diagram, as shown below.

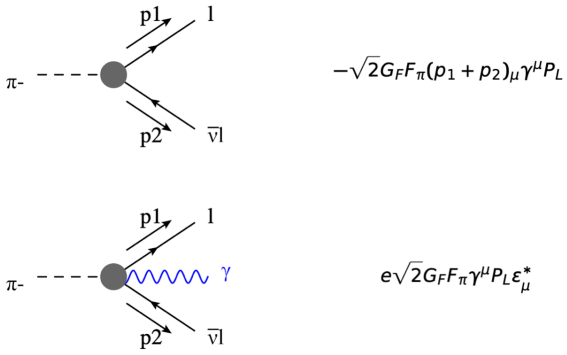


In contrast, in the effective Lagrangian with the W boson integrated out, Eq. (412), we could only radiated the photon from the π^- or ℓ^- lines. The middle diagram would correspond to a contact interaction among $\pi^-, \gamma, \ell^-, \bar{\nu}_\ell$. Where will it come from?

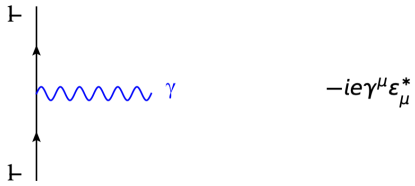
The solution is to impose QED gauge invariance to the effective theory, Eq. (412), and promote the partial derivative on the charged pion field into a covariant derivative,

$$\mathcal{L}_{\rm eff} = -\sqrt{2} G_F F_\pi V_{ud}^{\rm CKM} \left(\bar{\ell} \gamma^\mu \mathbb{P}_L \nu_\ell \right) D_\mu \pi^- \ , \tag{419}$$

where $D_\mu \pi^- = \partial_\mu \pi^- + ie A_\mu \pi^-$. This leads to the following Feynman rules.



In addition, the regular QED vertex for ℓ^--photon interaction is



From these Feynman rules, we can look at the three Feynman diagrams for $\pi^- \to \ell^- \bar{\nu}_\ell \gamma$. The amplitude for the first diagram is obviously proportional to m_ℓ, upon the use of equation of motion for the state fermions (both are on-shell). As said, the middle diagram corresponds to a contact interaction in the heavy W limit. The amplitude is

$$i\mathcal{M}_2 = e\sqrt{2} G_F F_\pi \left[\bar{u}(p_1, s_1) \not{\varepsilon}^* \mathbb{P}_L v(p_2, s_2) \right] \ . \tag{420}$$

The amplitude for the last diagram is

$$
\begin{aligned}
i\mathcal{M}_3 &= \bar{u}(p_1,s_1)\left(-ie\gamma^\mu\varepsilon^*_\mu\right)\frac{i}{\not p_1-\not k-m_\ell}\left(-\sqrt{2}G_F F_\pi(\not p_1+\not p_2+\not k)\mathbb{p}_L\right)v(p_2,s_2)\\
&= \bar{u}(p_1,s_1)\left(-ie\gamma^\mu\varepsilon^*_\mu\right)\frac{i}{\not p_1-\not k-m_\ell}\left(-\sqrt{2}G_F F_\pi(\not p_1+\not k)\mathbb{p}_L\right)v(p_2,s_2)\\
&= \bar{u}(p_1,s_1)\left(-ie\gamma^\mu\varepsilon^*_\mu\right)\frac{i}{\not p_1-\not k-m_\ell}\left(-\sqrt{2}G_F F_\pi(\not p_1+\not k+m_\ell-m_\ell)\mathbb{p}_L\right)v(p_2,s_2)\\
&= -e\sqrt{2}G_F F_\pi\bar{u}(p_1,s_1)\not\varepsilon^*\mathbb{P}_L v(p_2,s_2)\\
&\quad + e\sqrt{2}G_F F_\pi\bar{u}(p_1,s_1)\not\varepsilon^*\frac{m_\ell}{\not p_1-\not k-m_\ell}\mathbb{P}_L v(p_2,s_2)\ ,
\end{aligned}
\tag{421}
$$

where k is the four momentum of the outgoing photon.

Clearly, the first term in the last step cancels $i\mathcal{M}_2$ found above. The sum of all amplitudes is proportional to m_ℓ. Gauge invariance plays an important role here.

5.4 Neutral pion decay

The PCAC relation in Eq. (410) can be generalized to all eight axial vector currents, constructed with the broken generators from $SU(3)_L \times SU(3)_R \to SU(3)_V$. Define

$$
J^\mu_{5i} = \bar{q}\gamma^\mu\gamma_5\frac{\lambda_i}{2}q\ , \quad q = \begin{pmatrix} u\\ d\\ s\end{pmatrix}\ , \tag{422}
$$

the general PCAC relation is

$$
\langle 0|J^\mu_{5i}|\phi_j(q)\rangle = -iF_\pi q^\mu e^{-iq\cdot x}\delta_{ij}\ . \tag{423}
$$

We apply it to π^0 in this subsection,

$$
\langle 0|J^\mu_{53}|\pi^0(q)\rangle = -iF_\pi q^\mu e^{-iq\cdot x}\ . \tag{424}
$$

The following matching is then dictated

$$
J^\mu_{53} = \frac{1}{2}(\bar{u}\gamma^\mu\gamma_5 u - \bar{d}\gamma^\mu\gamma_5 d) \leftrightarrow F_\pi\partial^\mu\pi^0\ . \tag{425}
$$

We mentioned earlier that J^μ_{53} is anomaly free with respect to $SU(3)^2_c$. From the above current composition, we can see it is due to a cancelation between the u and d quark contributions in the triangle diagram (see discussions in sec. 2.4). However, due to the different electric charges of u and d, the current J^μ_{53} is actually anomalous with respect to $U(1)^2_{e.m.}$. In the massless quark limit, the divergence of the current is

$$
\begin{aligned}
\partial_\mu J^\mu_{53} &= -\frac{e^2}{16\pi^2}\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\times\left[\frac{1}{2}\left(\frac{2}{3}\right)^2-\frac{1}{2}\left(-\frac{1}{3}\right)^2\right]N_c\\
&= -\frac{e^2}{32\pi^2}\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = -\frac{\alpha}{4\pi}F_{\mu\nu}\widetilde{F}^{\mu\nu}\ .
\end{aligned}
\tag{426}
$$

Note again that $\widetilde{F}^{\mu\nu} \equiv \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$.

Applying the matching condition Eq. (425), we obtain an equation of motion of π^0 in the massless quark (thus massless π^0) limit,

$$\Box\pi^0 = -\frac{\alpha}{4\pi F_\pi}F_{\mu\nu}\widetilde{F}^{\mu\nu} \ . \tag{427}$$

Such an equation of motion can be obtained by introducing an effective pion-photon interacting term. The low energy effective Lagrangian for π^0 reads

$$\mathcal{L}_{\rm eff}(\pi^0) = \frac{1}{2}\partial_\mu\pi^0\partial^\mu\pi^0 - \frac{\alpha}{4\pi F_\pi}\pi^0 F_{\mu\nu}\widetilde{F}^{\mu\nu} \ . \tag{428}$$

Turning on the pion mass with nonzero quark mass, the second term allows π^0 to decay into two photons. It can be rewritten as

$$\mathcal{L}_{\rm int} = -\frac{\alpha}{2\pi F_\pi}\pi^0\varepsilon^{\mu\nu\rho\sigma}\partial_\mu A_\nu\partial_\rho A_\sigma \ . \tag{429}$$

The matrix element for $\pi^0 \to \gamma(p_1)\gamma(p_2)$ is

$$i\mathcal{M}_{\pi^0\to\gamma\gamma} = 2\times\left(-\frac{\alpha}{2\pi F_\pi}\right)\varepsilon^{\mu\nu\rho\sigma}(ip_{1\mu})(ip_{2\rho})\varepsilon^*_\nu(p_1)\varepsilon^*_\sigma(p_2) \ . \tag{430}$$

The prefactor 2 is a symmetry factor.

Using the identity $\varepsilon^{\mu\nu\rho\sigma}\varepsilon^{\mu'\nu\rho'\sigma} = 2(-g_{\mu\mu'}g_{\rho\rho'} + g_{\mu\rho'}g_{\mu'\rho})$, we get

$$|\mathcal{M}_{\pi^0\to\gamma\gamma}|^2 = \frac{\alpha^2 m_\pi^4}{2\pi^2 F_\pi^2} \ . \tag{431}$$

The decay rate is

$$\Gamma_{\pi^0\to\gamma\gamma} = \frac{1}{2}\times\frac{1}{8\pi}|\mathcal{M}_{\pi^0\to\gamma\gamma}|^2\frac{|\vec{p}_{1\rm cm}|}{m_\pi^2} = \frac{\alpha^2 m_\pi^3}{64\pi^3 F_\pi^2} \ . \tag{432}$$

The prefactor $\frac{1}{2}$ is there for identical photons in the final state. Numerically, using $m_{\pi^0} = 134.98\,\text{MeV}$, we get $\Gamma_{\pi^0\to\gamma\gamma} = 7.6\times10^{-9}\,\text{GeV}$. This is notably much larger than the charged pion decay rate. The lifetime of π^0 is $8.6\times10^{-17}\,\text{sec}$. At collider experiments, π^0 decays promptly into a pair of photons.

Another possible decay channel of π^0 is via virtual Z-boson exchange, whose rate is much lower than the above diphoton decay. The Z-mediated decay of $\pi^0 \to \nu\bar{\nu}$ has not been measured experimentally. The upper bound on its branching ratio is 2.7×10^{-7}.

5.5 Pion-nucleon coupling

For the rest this section, we first discuss the coupling of nucleons (proton and neutron) to the pions in effective theory, and then their connections to the fundamental theory with quarks.

The effective nucleon-pion interacting Lagrangian is

$$\mathcal{L}_{\text{eff}} = \frac{g_{\pi NN}}{M} \bar{N} \gamma^\mu \gamma_5 \frac{\sigma^a}{2} N \partial_\mu \pi^a + \cdots \tag{433}$$

where $\cdots$ represent terms involving two or more pions, σ^a are Pauli matrices, M is the nucleon mass, and

$$N = \begin{pmatrix} p \\ n \end{pmatrix} \ . \tag{434}$$

More explicitly, we have

$$\begin{aligned} \mathcal{L}_{\text{eff}} &= \frac{g_{\pi NN}}{2M} \begin{pmatrix} \bar{p} & \bar{n} \end{pmatrix} \gamma^\mu \gamma_5 \begin{pmatrix} \partial_\mu \pi^0 & \sqrt{2} \partial_\mu \pi^+ \\ \sqrt{2} \partial_\mu \pi^- & -\partial_\mu \pi^0 \end{pmatrix} \begin{pmatrix} p \\ n \end{pmatrix} \\ &= \frac{g_{\pi NN}}{2M} \left(\bar{p} \gamma^\mu \gamma_5 p - \bar{n} \gamma^\mu \gamma_5 n \right) \partial_\mu \pi^0 \\ &\quad + \frac{g_{\pi NN}}{\sqrt{2} M} \bar{p} \gamma^\mu \gamma_5 n \partial_\mu \pi^+ + \frac{g_{\pi NN}}{\sqrt{2} M} \bar{n} \gamma^\mu \gamma_5 p \partial_\mu \pi^- \ . \end{aligned} \tag{435}$$

For on-shell nucleons, the above effective Lagrangian can also be simplified by differentiating by parts and using equations of motion,

$$\mathcal{L}_{\text{eff}} = -g_{\pi NN} \left(\bar{p} i \gamma_5 p - \bar{n} i \gamma_5 n \right) - \sqrt{2} g_{\pi NN} \left(\bar{p} i \gamma_5 n \pi^+ + \text{h.c.} \right) \ . \tag{436}$$

Numerically, the value of $g_{\pi NN}$ is 12.9. It is a huge coupling.

The pion exchange potential between proton and neutron leads to the deuteron nucleus as a bound state.

5.6 W-boson-nucleon coupling

Here we consider how the left-handed quark current matches to the nucleon level operators. The discussion is relevant for calculating neutron beta decay.

The charged current interaction we are interested is

$$\mathcal{L}_{\text{CC}} = \frac{g_2}{\sqrt{2}} \bar{u} \gamma^\mu \mathbb{P}_L d W^+_\mu + \text{h.c.} \ , \tag{437}$$

where we already set $V^{\text{CKM}}_{ud} = 1$. Taking its matrix element between proton and neutron states, we have

$$\begin{aligned} \langle p(p_2) | \mathcal{L}_{\text{CC}} | n(p_1) \rangle &= \frac{g_2}{2\sqrt{2}} W^+_\mu \left[\langle p(p_2) | \bar{u} \gamma^\mu d | n(p_1) \rangle - \langle p(p_2) | \bar{u} \gamma^\mu \gamma_5 d | n(p_1) \rangle \right] \\ &= \frac{g_2}{2\sqrt{2}} W^+_\mu \left[F_1(q^2) \bar{u}(p_2, s_2) \gamma^\mu u(p_1, s_1) + F_2(q^2) \bar{u}(p_2, s_2) \sigma^{\mu\nu} q_\nu u(p_1, s_1) \right. \\ &\quad \left. - G_1(q^2) \bar{u}(p_2, s_2) \gamma^\mu \gamma_5 u(p_1, s_1) - G_2(q^2) \bar{u}(p_2, s_2) \gamma_5 q^\mu u(p_1, s_1) \right] \ , \end{aligned} \tag{438}$$

where $q = p_1 - p_2$ is the momentum exchange to the W boson, $u(p_1, s_1), u(p_2, s_2)$ are the spinors for the neutron and proton, respectively. In the second step, we introduced nucleon level form factors.

In the small q limit, applicable to neutron decay, we obtain the effective Lagrangian

$$\mathcal{L}_{\text{eff}} = \frac{g_2}{2\sqrt{2}} W_\mu^+ \bar{p}\gamma^\mu \left[F_1(0) - G_1(0)\gamma_5\right] n \ , \tag{439}$$

where $F_1(0) = 1$ and

$$g_A \equiv G_1(0) \simeq 1.27 \ . \tag{440}$$

5.7 Goldberger-Treiman relation

The Goldberger-Treiman relation observes a remarkable connection between the nucleon-pion and nucleon-W effective interactions introduced above.

Let us take a closer look at the matrix element of axial current in Eq. (439),

$$\langle p|\bar{u}\gamma^\mu\gamma_5 d|n\rangle = G_1(q^2)\bar{u}(p_2,s_2)\gamma^\mu\gamma_5 u(p_1,s_1) + G_2(q^2)\bar{u}(p_2,s_2)\gamma_5 q^\mu u(p_1,s_1) \ . \tag{441}$$

We first multiply q_μ on both side (corresponding to taking the divergence of the axial current in coordinate space). Because $\bar{u}\gamma^\mu\gamma_5 d$ is a charged current, it is anomaly free. We work in the chiral limit where quark masses are zero. These imply

$$q_\mu \langle p|\bar{u}\gamma^\mu\gamma_5 d|n\rangle = 0 \ . \tag{442}$$

On the right-hand side, we use equation of motion for the nucleon spinors, which gives

$$\begin{aligned} & q_\mu \left[G_1(q^2)\bar{u}(p_2,s_2)\gamma^\mu\gamma_5 u(p_1,s_1) + G_2(q^2)\bar{u}(p_2,s_2)\gamma_5 q^\mu u(p_1,s_1)\right] \\ & = \left[-2MG_1(q^2) + q^2 G_2(q^2)\right] \bar{u}(p_2,s_2)\gamma_5 q^\mu u(p_1,s_1) \ . \end{aligned} \tag{443}$$

In the $q^2 \to 0$ limit, we find the relation

$$G_2(q^2) \simeq \frac{2MG_1(q^2)}{q^2} \ , \quad \text{for } q^2 \to 0 \ . \tag{444}$$

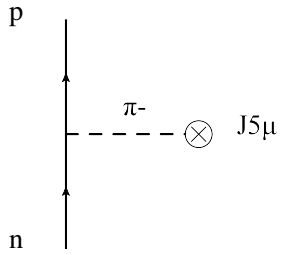
This behaviour means $G_2(q^2)$ has a pole as $q^2 \to 0$. In this regime, it is much larger than $G_1(q^2)$. With this knowledge, we consider the small q^2 limit of Eq. (441), which is approximately

$$\langle p|\bar{u}\gamma^\mu\gamma_5 d|n\rangle \simeq G_2(q^2)\bar{u}(p_2,s_2)\gamma_5 q^\mu u(p_1,s_1) \ , \quad \text{for } q^2 \to 0 \ . \tag{445}$$

Next, for the left-hand side of Eq. (441), we replace the axial current using the matching between quark current and pion field, Eq. (411) (which works at low energy $q^2 \to 0$)

$$\langle p|\bar{u}\gamma^\mu\gamma_5 d|n\rangle = \sqrt{2}F_\pi \langle p|\partial^\mu \pi^-|n\rangle \ . \tag{446}$$

The right-hand side correspond to the following diagram and can be calculated by inserting the $\pi^- pn$ interaction in Eq. (436).



$$\begin{aligned}\langle p|\bar{u}\gamma^\mu\gamma_5 d|n\rangle &= \sqrt{2}F_\pi\left(-i\sqrt{2}g_{\pi NN}\right)\left[\bar{u}(p_2,s_2)i\gamma_5 u(p_1,s_1)\right](-iq^\mu)\frac{i}{q^2-m_\pi^2}\\ &\simeq \frac{2g_{\pi NN}F_\pi q^\mu}{q^2}\left[\bar{u}(p_2,s_2)\gamma_5 u(p_1,s_1)\right]\ .\end{aligned} \tag{447}$$

In the last step, we take the chiral limit where $m_\pi \to 0$. Comparing this with Eq. (445) in the $q^2 \to 0$ limit, we get

$$G_2(q^2) \simeq \frac{2g_{\pi NN}F_\pi}{q^2}\ , \quad \text{for } q^2 \to 0\ . \tag{448}$$

Further comparing with Eq. (444), we obtain

$$g_{\pi NN}F_\pi = MG_1(0)\ , \tag{449}$$

or

$$g_{\pi NN} = \frac{M}{F_\pi}g_A\ , \tag{450}$$

This is the Goldberger-Treiman relation.