Part III: The Standard Model - Revision

Lectures by Christopher Thomas, notes by James Moore

1 Revision of spinors

1.1 The Clifford algebra and spinors

Definition: The *Clifford algebra* is an algebra generated by the objects γ^μ obeying $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}1$.

When we pick a solution to this equation, we say we are picking a *representation* of the Clifford algebra.

Definition: The *chiral* (or *Weyl*) *representation* of the Clifford algebra is the set of matrices:

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix},$$

where σ^i are the Pauli matrices.

Theorem: In the chiral rep, $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$. So $\gamma^{0\dagger} = \gamma^0$ and $\gamma^{i\dagger} = -\gamma^i$.

Proof: Trivial. □

This is useful, because choosing a rep does not constitute any loss of generality:

Theorem: If γ^μ and $\gamma^{\mu\prime}$ are both reps of the Clifford algebra, there exists S such that $\gamma^{\mu\prime} = S\gamma^\mu S^{-1}$.

Proof: Not required. □

Definition: The *chirality operator* is the γ^5 matrix, defined by $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$, which in the chiral representation is

$$\gamma^5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Theorem: γ^5 obeys $(\gamma^5)^2 = I$ and $\{\gamma^5, \gamma^\mu\} = 0$.

Proof: See QFT. □

Definition: A *spinor* is a four-component object $\psi \in \mathbb{C}^4$ transforming in the spinor representation of the Lorentz group (see QFT). The *Dirac adjoint* of a spinor is defined by:

$$\overline{\psi} = \psi^\dagger\gamma^0.$$

Definition: The *Dirac Lagrangian* is $\mathcal{L} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi$.

Theorem: The Euler-Lagrange equations of the Dirac Lagrangian are:

$$(i\partial\!\!\!/ - m)\psi = 0, \qquad \overline{\psi}(-i\overleftarrow{\partial\!\!\!/} - m) = 0,$$

where $A\!\!\!/ = \gamma^\mu A_\mu$, and $\overleftarrow{\partial\!\!\!/}$ mean the derivative acts to the left. The first equation is the *Dirac equation*.

Proof: Vary directly with respect to $\overline{\psi}$ to get first equation. Vary with respect to ψ and integrate by parts twice to get second equation. □

1.2 Chiral spinors

Definition: The *projection operators* are defined by $P_{R,L} = \frac{1}{2}\left(1 \pm \gamma^5\right)$.

Theorem: The projection operators obey:

$$P_{R,L}^2 = P_{R,L}, \qquad P_RP_L = P_LP_R = 0, \qquad P_L + P_R = 1.$$

Proof: Just substitute definitions to check. □

Definition: Given a spinor field $\psi(x)$, the *chiral* (*Weyl*) *spinors* are $\psi_{R,L}(x) = P_{R,L}\psi(x)$.

Theorem: $\overline{\psi}_{L,R}(x) = \overline{\psi}(x)P_{R,L}$.

Proof: Just direct check. □

Theorem: $\gamma^5\psi_{R,L}(x) = \pm\psi_{R,L}(x)$.

Proof: Just direct check. □

Because of this final property, $\psi_{R,L}$ are chiral eigenstates. We say they are *right* and *left-handed* respectively.

Theorem: The Dirac Lagrangian can be written as:

$$\mathcal{L} = i\overline{\psi}_L\partial\!\!\!/\psi_L + i\overline{\psi}_R\partial\!\!\!/\psi_R - m(\overline{\psi}_R\psi_L + \overline{\psi}_L\psi_R).$$

Proof: Use $\psi(x) = (P_R + P_L)\psi(x) = \psi_L(x) + \psi_R(x)$ in original Dirac Lagrangian. □

Hence, in the massless limit the Lagrangian is:

$$\mathcal{L} = i\overline{\psi}_L\partial\!\!\!/\psi_L + i\overline{\psi}_R\partial\!\!\!/\psi_R.$$

We notice this Lagrangian has a global $U(1)_L \times U(1)_R$ symmetry:

Theorem: For massless spinors, the transformation $\psi_{L,R} \mapsto e^{i\alpha_{L,R}}\psi_{L,R}$ is a symmetry of the Lagrangian.

Proof: Trivial. □

In particular, the transformations are *independent*. When $m \neq 0$, there is still a global $U(1)$ symmetry $\psi \mapsto e^{i\alpha}\psi$, equivalent to setting $\alpha_L = \alpha_R$. We say the mass *breaks* the symmetry (see later).

1.3 Quantising spinors

Quantum spinor fields have a mode expansion: $\psi(x) =$

$$\sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_{\mathbf{p}}} \left(b^s(p)u^s(p)e^{-ip\cdot x} + (d^s)^\dagger(p)v^s(p)e^{ip\cdot x}\right)$$

$$=: \sum_{s,p} \left(b^s(p)u^s(p)e^{-ip\cdot x} + (d^s)^\dagger(p)v^s(p)e^{ip\cdot x}\right).$$

where $s \in \{-\frac{1}{2}, \frac{1}{2}\}$ and b, d are mode operators. Note we've used the notation

$$\sum_p = \frac{d^3\mathbf{p}}{(2\pi)^3 2E_{\mathbf{p}}}$$

to save on writing. Also note that $u^s(p)$, $v^s(p)$ are *positive* and *negative* frequency plane-wave solutions of the classical Dirac equation, obeying

$$(\not{p} - m)u = 0, \qquad (\not{p} + m)v = 0,$$

and explicitly given by

$$u^s(p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\xi^s \\ \sqrt{p\cdot\overline{\sigma}}\xi^s \end{pmatrix}, \qquad v^s(p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\eta^s \\ -\sqrt{p\cdot\overline{\sigma}}\eta^s \end{pmatrix},$$

where $\sigma^\mu = (1, \sigma^i)$, $\overline{\sigma}^\mu = (1, -\sigma^i)$, and ξ^s, η^s are two-component spinors. We take $\xi^1 = (1,0)^T$ and $\xi^2 = (0,1)^T$ for spin up and spin down respectively; same holds for η.

Finally, notice that this course uses *relativistic normalisation* of the *operators*, that is,

$$a(p) = \sqrt{2E_{\mathbf{p}}}a_{\mathbf{p}},$$

where $a_{\mathbf{p}}$ are the operators we used in QFT.

In order to deal with the square root of a matrix, i.e. $\sqrt{p\cdot\sigma}$, we rotate the spatial coordinates so that $\mathbf{p}$ lies in the z-direction, i.e. $p = (p^0, 0, 0, p^3)^T$. Then

$$\sqrt{p\cdot\sigma} = \sqrt{p^0\sigma^0 - p^3\sigma^3} = \begin{pmatrix} p^0 - p^3 & 0 \\ 0 & p^0 + p^3 \end{pmatrix}^{1/2}$$

$$= \begin{pmatrix} \sqrt{p^0 - p^3} & 0 \\ 0 & \sqrt{p^0 + p^3} \end{pmatrix}.$$

1.4 Helicity vs chirality

Definition: The *helicity* operator is the projection of angular momentum onto the linear momentum direction: $h = \mathbf{J}\cdot\hat{\mathbf{p}} = (\mathbf{r}\times\mathbf{p} + \mathbf{S})\cdot\hat{\mathbf{p}} = \mathbf{S}\cdot\hat{\mathbf{p}}$, where $\mathbf{S}$ is *spin angular momentum* and $\mathbf{r}\times\mathbf{p}$ is *orbital angular momentum*.

Theorem: In QFT, the spin operator can be written as:

$$S_i = \frac{i}{4}\epsilon_{ijk}\gamma^j\gamma^k.$$

Proof: Recall from quantum mechanics that $\mathbf{S} = \frac{1}{2}\boldsymbol{\sigma}$, where $\boldsymbol{\sigma}$ is the vector of Pauli matrices. So the spin operator should act on spinors via:

$$S_i = \frac{1}{2}\begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} = \frac{i}{4}\epsilon_{ijk}\gamma^j\gamma^k.$$

Now notice that in the chiral rep, we have:

$$[\gamma^j, \gamma^k] = \begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} - \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -[\sigma^j, \sigma^k] & 0 \\ 0 & -[\sigma^j, \sigma^k] \end{pmatrix} = \begin{pmatrix} -2i\epsilon^{ijk}\sigma^i & 0 \\ 0 & -2i\epsilon^{ijk}\sigma^i \end{pmatrix}.$$

Hence $\epsilon^{ijk}S_i = \frac{i}{4}[\gamma^j, \gamma^k]$ (recall $S_i = \frac{1}{2}\mathrm{diag}(\sigma^i, \sigma^i)$). Now multiply through by ϵ^{ijk} to get the result. □

Using the above form of the spin operator, we have:

Theorem: For massless spinors, $hu^s(p) = \frac{1}{2}\gamma^5 u^s(p)$.

Proof: Let $u^s(p)$ be massless, and suppose, by rotating coordinates, that $\mathbf{p}$ points in the positive z-direction, i.e. $p^1 = p^2 = 0$, $p^3 > 0$. Then $p^0 = p^3$ because this is a massless spinor. Thus we have:

$$u^s(p) = \begin{pmatrix} \sqrt{p\cdot\sigma}\xi^s \\ \sqrt{p\cdot\overline{\sigma}}\xi^s \end{pmatrix} = \begin{pmatrix} 0 \\ \sqrt{2p^0}\xi^s \\ \sqrt{2p^0}\xi^s \\ 0 \end{pmatrix}.$$

Now recall that $h = \hat{p}^i S_i = \hat{p}^3 S_3 = S_3$. Hence $hu^s(p) =$

$$S_3 u^s(p) = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ \sqrt{2p^0}\xi^s \\ \sqrt{2p^0}\xi^s \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 \\ -\sqrt{2p^0}\xi^s \\ \sqrt{2p^0}\xi^s \\ 0 \end{pmatrix}.$$

Now simply recall $\gamma^5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}$, and we see that $hu^s(p) = \frac{1}{2}\gamma^5 u^s(p)$. This is Lorentz invariant so holds for all p. □

Slogan: For *massless* spinors, chirality and helicity are the *same* (up to $\frac{1}{2}$). For right-handed massless spinors, spin points in the direction of motion; for left-handed massless spinors, spin points opposite the direction of motion.

2 Discrete symmetries

2.1 Types of symmetry

Definition: Types of symmetry include:

(i) *Intact*. The symmetry is present in both the classical and quantum theory.

(ii) *Anomalous*. The symmetry holds classically but not quantumly. This is not a true symmetry.

(iii) *Explicitly broken*. The symmetry applies to some terms in the Lagrangian, but not others. This indicates the symmetry has been broken from a larger group (see later).

(iv) *Hidden*. The symmetry is respected by the Lagrangian, but *not* the vacuum state. There main type is *spontaneous symmetry breaking*, where we get a vacuum excitation value from one or more scalar fields.

2.2 Wigner's Theorem

In order to implement symmetries in the quantum theory, we need the important result:

Wigner's Theorem: Let ψ, ψ' and ϕ be vectors in Hilbert space. If physics is invariant under the transformation $\psi \mapsto \psi'$, then there exists an operator W such that $\psi' = W\psi$, where W is either

(i) Linear, $W(\alpha\phi + \beta\psi) = \alpha W\phi + \beta W\psi$, and unitary, $(W\phi, W\psi) = (\phi, \psi)$;

(ii) Anti-linear, $W(\alpha\phi + \beta\psi) = \alpha^* W\phi + \beta^* W\psi$, and anti-unitary, $(W\phi, W\psi) = (\phi, \psi)^*$.

2.3 CPT symmetries

In this chapter, we'll care about three symmetries: *charge conjugation*, *parity* and *time-reversal*.

Definition: *Charge conjugation* symmetry maps particles to their respective antiparticles, and vice-versa.

Definition: *Parity* inverts space via $(t, \mathbf{x}) \mapsto (t, -\mathbf{x})$. It is represented in the classical theory by a Lorentz transformation

$$\mathbb{P}^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Definition: *Time-reversal* inverts time via $(t, \mathbf{x}) \mapsto (-t, \mathbf{x})$. It is represented in the classical theory by a Lorentz transformation

$$\mathbb{T}^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note both parity and time-reversal are *improper* Lorentz transformations since they have determinant -1 (so they are not connected to the identity).

Before we start, we should work out whether parity and time-reversal are unitary or anti-unitary.

Notation: Let $W(\Lambda, a)$ denote the quantum operator representing the Poincaré transformation $x^\mu \mapsto \Lambda^\mu{}_\nu x^\nu + a^\mu$, where $\Lambda^\mu{}_\nu$ is a Lorentz transformation.

Lemma: $W(\Lambda_1, a_1)W(\Lambda_2, a_2) = W(\Lambda_1\Lambda_2, \Lambda_1 a_2 + a_1)$.

Proof: W is a group homomorphism. So do calculation in classical space. We have:

$$\begin{aligned} x^\mu &\mapsto {\Lambda_2}^\mu{}_\nu x^\nu + a_2^\mu \\ &\mapsto {\Lambda_1}^\mu{}_\nu {\Lambda_2}^\nu{}_\rho x^\rho + {\Lambda_1}^\mu{}_\nu a_2^\nu + a_1^\mu. \ \square \end{aligned}$$

Lemma: For ω, ϵ infinitesimal, we have

$$W(1 + \omega, \epsilon) = 1 + \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} - i\epsilon_\mu P^\mu,$$

where $J^{\mu\nu}$ generate Lorentz transformations and P^μ generate time and space translations (i.e. P^0 is the Hamiltonian, P^i is the 3-momentum).

Proof: By definition of the infinitesimal operators. $\square$

We can now prove the main result:

Theorem: The operators $\hat{P} = W(\mathbb{P}, 0)$ and $\hat{T} = W(\mathbb{T}, 0)$ representing parity and time-reversal respectively, are unitary & linear, and anti-unitary & anti-linear respectively.

Proof: By the composition law above, $\hat{P}W(\Lambda, a)\hat{P}^{-1} = W(\mathbb{P}\Lambda\mathbb{P}^{-1}, \mathbb{P}a)$ and same for $\mathbb{T}$. Inserting infinitesimal expansion of W on both sides, and comparing coefficients of $-\epsilon_0$, we find

$$\hat{P}iH\hat{P}^{-1} = iH, \qquad \hat{T}iH\hat{T}^{-1} = -iH.$$

Let ψ be an energy eigenstate of H with energy E. Then

$$(\psi, iH\psi) = (\psi, iE\psi) = iE.$$

If $\hat{T}$ is a symmetry, then $\hat{T}\psi$ is an eigenstate of H (since $\hat{T}$ commutes with H when symmetry).

We note $H\hat{T}\psi = E\hat{T}\psi$, so they have the same energy (E is real since H Hermitian, so doesn't matter if linear or antilinear at this stage).

Suppose $\hat{T}$ is linear. Then

$$iE = (\hat{T}\psi, iH\hat{T}\psi) = -(\hat{T}\psi, \hat{T}iH\psi) = -iE,$$

where in the middle equality, we've used $\hat{T}iH\hat{T}^{-1} = -iH$, and in the last equality, we've used linearity. Contradiction (since must hold for all E). So by Wigner's Theorem, $\hat{T}$ is anti-linear and anti-unitary. Similarly can show $\hat{P}$ is linear and unitary. □

In the above Theorem, we only used the coefficient of $-\epsilon_0$ in the infinitesimal expansion. Using other coefficients it's possible to show:

Theorem: Under time-reversal symmetry $\hat{T}$, both angular momentum and linear momentum change sign.

This will be important later when we consider how the Dirac field (which has spin angular momentum $\frac{1}{2}$) transforms under $\hat{T}$.

2.4 Parity symmetry

We want to investigate how quantum fields change under parity, $\hat{P}$. Begin with scalar field:

Theorem: Under parity, the scalar field

$$\phi(x) = \sum_p \left(a(p)e^{-ip\cdot x} + c^\dagger(p)e^{ip\cdot x}\right)$$

maps to

$$\hat{P}\phi(x)\hat{P}^{-1} = \eta_P\phi(x_P),$$

where $x_P = (x^0, -\mathbf{x})$ are the parity-transformed coordinates, and η_P is a complex phase, called the *intrinsic parity* of the field.

Proof: Note $x\cdot p$ is invariant under Lorentz transformations, so $p \mapsto p_P = (p^0, -\mathbf{p})$ under parity. Hence in the quantum theory we require $\hat{P}a^\dagger(p)\hat{P}^{-1} = (\eta^a)^* a^\dagger(p_P)$. Since $\hat{P}$ is unitary, taking the dagger gives $\hat{P}a(p)\hat{P}^{-1} = \eta^a a(p_P)$. Similarly

$$\hat{P}c^\dagger(p)\hat{P}^{-1} = (\eta^c)^* c^\dagger(p_P), \qquad \hat{P}c(p_P)\hat{P}^{-1} = \eta^c c(p_P).$$

Hence

$$\hat{P}\phi(x)\hat{P}^{-1} = \sum_p \left(\eta^a a(p_P)e^{-ip\cdot x} + (\eta^c)^* c^\dagger(p_P)e^{ip\cdot x}\right).$$

Relabel $p \leftrightarrow p_P$ in the sum and use $p_P \cdot x = p \cdot x_P$. Also note integrating over p_P is the same as integrating over p (same range). Hence

$$\hat{P}\phi(x)\hat{P}^{-1} = \sum_p \left(\eta^a a(p)e^{-ip\cdot x_P} + (\eta^c)^* c^\dagger(p)e^{ip\cdot x_P}\right).$$

Now must constrain η^a, η^c. Notice

$$[\phi(x), \hat{P}\phi^\dagger(y)\hat{P}^{-1}] = (2\pi)^3\left((\eta^a)^* D(x - y_P) + \eta^c D(y_P - x)\right),$$

where $D(x-y)$ is the propagator. For causality to be preserved, this must vanish for spacelike separated x and y. Recalling $D(x-y) = -D(y-x)$, we see that $(\eta^a)^* = \eta^c$. The result follows. □

For real ϕ, $a = c$, and so $\eta^a = (\eta^a)^* = \eta_P \in \mathbb{R}$. Hence η_P is ± 1.

Definition: If $\eta_P = 1$, we call the field a *scalar field*. If $\eta_P = -1$, we call it a *pseudoscalar field*.

For complex scalar fields, it's possible $\eta_P \notin \mathbb{R}$. But if there is a conserved charge Q, then $\hat{Q}$, $\hat{P}$, $\hat{H}$ are all mutually commuting, allowing us to define $\hat{P}' = \hat{P}e^{-i\alpha\hat{Q}}$, with α chosen such that $\eta_{P'}$ is real. We then just take $\hat{P}'$ to be the parity operator.

We now want to consider vector fields. Recall that a vector field in the quantum theory has mode expansion:

$$V^\mu(x) = \sum_{p,\lambda}\left(\epsilon^\mu(\lambda,p)a^\lambda(p)e^{-ip\cdot x} + \epsilon^{\mu *}(\lambda,p){c^\lambda}^\dagger(p)e^{ip\cdot x}\right),$$

where $\lambda \in \{-1,0,1\}$ is the *helicity*. The ϵ^μ are the *polarisation vectors*.

Theorem: Under parity, the vector field $V^\mu(x)$ transforms to $\hat{P}V^\mu(x)\hat{P}^{-1} = -\eta_P \mathbb{P}^\mu{}_\nu V^\nu(x_P)$.

Proof: The proof is essentially the same as for a scalar field. However, we must also use the fact $\epsilon^\mu(\lambda, p_P) = -\mathbb{P}^\mu{}_\nu\epsilon^\nu(\lambda,p)$. This result can be shown by working in a specific frame.

In the particle's rest frame, $p = 0$, pick

$$\epsilon^\mu(-1,0) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\-i\\0\end{pmatrix},\ \epsilon^\mu(0,0) = \begin{pmatrix}0\\0\\0\\1\end{pmatrix},\ \epsilon^\mu(1,0) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\i\\0\end{pmatrix}.$$

Boost to frames with 3-momentum $\mathbf{p} = mv\hat{z}$ and $p_P = -mv\hat{z}$ respectively, by using the Lorentz boosts $L^\mu{}_\nu(p)$, $L^\mu{}_\nu(p_P)$; explicitly, this Lorentz boost is given by

$$L^\mu{}_\nu(p) = \begin{pmatrix}\gamma & 0 & 0 & -v\gamma\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ -v\gamma & 0 & 0 & \gamma\end{pmatrix},$$

where γ is the standard Lorentz factor, and to get $L^\mu{}_\nu(p_P)$ we just send $v \mapsto -v$. We find the result holds in this specific frame; by Lorentz invariance, it holds in all frames. The proof is then the same as in the bosonic case. □

Finally, we examine how spinors transform.

Theorem: A Dirac field $\psi(x)$ transforms as

$$\hat{P}\psi(x)\hat{P}^{-1} = \eta_P \gamma^0 \psi(x_P)$$

under parity. The Dirac conjugate transforms as

$$\hat{P}\overline{\psi}(x)\hat{P}^{-1} = \eta_P^* \overline{\psi}(x_P)\gamma^0.$$

Proof: Just as for bosons, we have

$$\hat{P}b^s(p)\hat{P}^{-1} = \eta^b b^s(p_P), \quad \hat{P}d^{s\dagger}(p)\hat{P}^{-1} = (\eta^d)^* d^{s\dagger}(p_P).$$

Hence going through the same steps as for bosons, we get

$$\hat{P}\psi(x)\hat{P}^{-1} = \sum_{p,s} \eta^b b^s(p) u^s(p_P) e^{-ip\cdot x_P} + \eta^{d^*} d^{s\dagger}(p) v^s(p_P) e^{ip\cdot x_P}.$$

We need to evaluate $u^s(p_P)$. From the general form of $u^s(p)$, we have

$$u^s(p_P) = \begin{pmatrix} \sqrt{\sigma\cdot p_P}\xi^s \\ \sqrt{\overline{\sigma}\cdot p_P}\xi^s \end{pmatrix} = \begin{pmatrix} \sqrt{\overline{\sigma}\cdot p}\xi^s \\ \sqrt{\sigma\cdot p}\xi^s \end{pmatrix} = \gamma^0 u^s(p).$$

Similarly, $v^s(p_P) = -\gamma^0 v^s(p)$. Thus $\hat{P}\psi(x)\hat{P}^{-1} =$

$$\sum_{s,p} \left(\eta^b b^s(p)\gamma^0 u^s(p) e^{-ip\cdot x_P} - \eta^{d^*} d^{s\dagger}(p)\gamma^0 v^s(p) e^{ip\cdot x_P} \right).$$

Similar to boson calculation (but using anti-commutator of $\psi(x)$ and $\psi(x_P)$), we require $\eta^b = -(\eta^d)^*$ for preservation of causality. Defining $\eta^b = \eta_P$, we have

$$\hat{P}\psi(x)\hat{P}^{-1} = \eta_P\gamma^0\psi(x_P).$$

Running the whole argument through with $\overline{\psi}$ instead, we get the other result. □

Theorem: For chiral spinors, we have:

$$\hat{P}\psi_{R/L}(x)\hat{P}^{-1} = \eta_P\gamma^0\psi_{L/R}(x),$$

i.e. right and left-handed spinors are swapped.

Proof: Just multiply above result by $\frac{1}{2}(1\pm\gamma^5)$. □

The above result allows us to calculate how fermion bilinears transform under parity:

- Under parity,

  $$\overline{\psi}(x)\psi(x) \mapsto \hat{P}\overline{\psi}(x)\psi(x)\hat{P}^{-1} = \hat{P}\overline{\psi}(x)\hat{P}^{-1}\hat{P}\psi(x)\hat{P}^{-1}.$$

  Inserting the above formulae, we get $\overline{\psi}(x_P)\psi(x_P)$. So $\overline{\psi}\psi$ is a scalar.

- Similarly, $\overline{\psi}(x)\gamma^5\psi(x) \mapsto -\overline{\psi}(x_P)\gamma^5\psi(x_P)$, so this is a pseudoscalar.

- $\overline{\psi}(x)\gamma^\mu\psi(x) \mapsto \mathbb{P}^\mu{}_\nu\overline{\psi}(x_P)\gamma^\nu\psi(x_P)$, so this is a vector.

- $\overline{\psi}(x)\gamma^\mu\gamma^5\psi(x) \mapsto -\mathbb{P}^\mu{}_\nu\overline{\psi}(x_P)\gamma^\nu\gamma^5\psi(x)$, so this is an axial vector.

Theorem: If $\psi(x)$ solves the Dirac equation, so does $\hat{P}\psi(x)\hat{P}^{-1}$.

Proof: The Dirac equation is $(i\not{\partial} - m)\psi(x) = 0$. Send $x \mapsto x_P$. Then $(i\gamma^0\partial_0 - i\gamma^i\partial_i - m)\psi(x_P) = 0$. Then just multiply through by γ^0 to get $(i\not{\partial} - m)(\gamma^0\psi(x_P)) = 0$. □

2.5 Charge conjugation of scalars

Theorem: Under charge conjugation, a scalar field transforms as

$$\hat{C}\phi(x)\hat{C}^{-1} = \eta_C\phi^\dagger(x),$$

where η_C is a complex phase called the *intrinsic charge parity* of the field.

Proof: Charge conjugation should transform a particle to its antiparticle. So we must have:

$$\hat{C}a(p)\hat{C}^{-1} = \eta_C c(p), \quad \hat{C}c(p)\hat{C}^{-1} = \eta_C^* a(p).$$

In principle, the two phases could be unrelated, but in actuality causality constrains these as we've seen before. This immediately gives $\hat{C}\phi(x)\hat{C}^{-1} = \eta_C\phi^\dagger(x)$. □

For a real scalar, $\phi(x) = \phi^\dagger(x)$, so $\eta_C = \pm 1$.

For a complex scalar, we could have $\eta_C = e^{2i\beta}$, say. But then performing the rotation $\phi \mapsto \phi' = e^{-i\beta}\phi$, we find that $\hat{C}\phi'\hat{C}^{-1} = (\phi')^\dagger$. So we can always redefine a complex scalar field to have $\eta_C = 1$.

2.6 Dirac fields: charge conjugation matrix

Let's now deal with Dirac fields; we'll need some motivation first. We might hope the spinor $\psi(x)$ transforms to something of the form $\psi^\dagger(x) = \overline{\psi}^T(x)$ under charge conjugation, as this was the 'antiparticle' in QFT. We thus hope the answer is of the form

$$\hat{C}\psi(x)\hat{C}^{-1} = \eta_C C\overline{\psi}^T(x).$$

where C is a 4×4 matrix.

Why the matrix? To ensure that the Dirac equation is satisfied. If $\psi(x)$ satisfies the Dirac equation, we know $\overline{\psi}(x)(-i\overleftarrow{\not{\partial}} - m) = 0$. Taking the transpose, we have

$$(-i(\gamma^{\mu T})\partial_\mu - m)\overline{\psi}^T(x) = 0,$$

and so multiplying by C, and inserting CC^{-1} appropriately, we have

$$(-iC\gamma^{\mu T}C^{-1}\partial_\mu - m)C\overline{\psi}^T(x) = 0,$$

for any matrix C. Now supposing $C\gamma^{\mu T}C^{-1} = -\gamma^\mu$, $C\overline{\psi}^T(x)$ satisfies the Dirac equation (whereas $\overline{\psi}^T(x)$ on its own didn't work)!

With this in mind, we define:

Definition: The *charge conjugation matrix* is the matrix obeying $C{\gamma^{\mu}}^T C^{-1} = -\gamma^{\mu}$. Equivalently,

$$C{\gamma^{\mu}}^T = -\gamma^{\mu} C,\ {\gamma^{\mu}}^T C^{-1} = -C^{-1}\gamma^{\mu},\ \text{ or }\ \gamma^{\mu} C^T = -C^T {\gamma^{\mu}}^T.$$

Slogan: Can commute C, C^{-1} or C^T past γ^{μ} at cost of minus sign, and introduction of transpose on γ^{μ}.

The charge conjugation matrix C has a lot of properties which make it useful:

Theorem (Properties of C): We have:

(i) C always exists, whatever rep of γ^{μ} is selected;

(ii) C is antisymmetric, $C^T = -C$;

(iii) $C{\gamma^5}^T C^{-1} = \gamma^5$; so above slogan extends to γ^5, but don't pick up minus sign when commuting past;

(iv) In the chiral rep, C is unitary, $C^{\dagger} C = 1$;

(v) A possible choice of C is

$$C = -i\gamma^0\gamma^2 = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix}.$$

Proof: (i) Recall that if $\gamma^{\mu\prime}$ and γ^{μ} both satisfy the Clifford algebra, then $\gamma^{\mu\prime} = S\gamma^{\mu}S^{-1}$ for some matrix S. But $-{\gamma^{\mu}}^T$ satisfies the Clifford algebra, so indeed a C exists.

(ii) First we prove $C^T C^{-1} \propto 1$. Note that $[C^T C^{-1}, \gamma^{\mu}] = 0$, and so by Schur's Lemma, $C^T C^{-1} = c1$, for some c. Since

$$C = (C^T)^T = (cC)^T = c^2 C,$$

we have $c = \pm 1$.

It remains to decide if C is symmetric or antisymmetric. We use an argument based on linear independence.

Note that $C^T = cC$, $(\gamma^{\mu}C)^T = -c\gamma^{\mu}C$, $(\gamma^5 C)^T = c\gamma^5 C$, $(\gamma^{\mu}\gamma^5 C)^T = c\gamma^{\mu}\gamma^5 C$ and $([\gamma^{\mu}, \gamma^{\nu}]C)^T = -c[\gamma^{\mu}, \gamma^{\nu}]C$. Also assert that the matrices C, $\gamma^{\mu}C$, $\gamma^5 C$, $\gamma^{\mu}\gamma^5 C$ and $[\gamma^{\mu}, \gamma^{\nu}]C$ are linearly independent (this is tedious to prove). This is a total of 16 matrices.

From the above calculations, if $c = 1$, we get 6 linearly independent symmetric matrices and 10 linearly independent antisymmetric ones. This is a contradiction. So $c = -1$.

(iii) Follows from calculation in (ii).

(iv) First note that $C{\gamma^{\mu}}^T = -\gamma^{\mu} C \Rightarrow ({\gamma^{\mu}}^{\dagger})^T C^{\dagger} = -C^{\dagger}{\gamma^{\mu}}^{\dagger}$. Considering $\mu = 0, i$ separately and using the chiral rep, we have ${\gamma^{\mu}}^T C^{\dagger} = -C^{\dagger}\gamma^{\mu}$. So we have

$$[\gamma^{\mu}, CC^{\dagger}] = 0,$$

in the same way as (i). Schur's Lemma then implies $CC^{\dagger} = \lambda 1$. Hence

$$C = (C^{\dagger})^{\dagger} = \lambda(C^{-1})^{\dagger} = \lambda(C^{\dagger})^{-1} = \lambda^2 C,$$

so $\lambda = \pm 1$.

If $\lambda = -1$, let $D = iC$. Then D obeys the defining property $D{\gamma^{\mu}}^T D^{-1} = \gamma^{\mu}$, so is a valid charge conjugation matrix with all the properties above (recall, we didn't assume a rep in (i) - (iii)). But $CC^{\dagger} = -1 \Rightarrow DD^{\dagger} = 1$, and we're saved, by working with D instead, WLOG.

(v) We can easily verify the defining property. By (i), the defining property is equivalent to $(\gamma^{\mu}C)^T = \gamma^{\mu}C$, which is easy to verify when $C = -i\gamma^0\gamma^2$. □

2.7 Charge conjugation of Dirac spinors

We are now ready to prove the main result:

Theorem: Under charge conjugation, a spinor transforms as

$$\hat{C}\psi(x)\hat{C}^{-1} = \eta_C C\overline{\psi}^T(x),$$

where C is the charge conjugation matrix.

Proof: Under charge conjugation, we must have

$$\hat{C}b^s(p)\hat{C}^{-1} = \eta_C d^s(p), \quad \hat{C}{d^s}^{\dagger}(p)\hat{C}^{-1} = \eta_C {b^s}^{\dagger}(p),$$

since particles are swapped with antiparticles (in principle, could get different phases, but in actuality, these are the same by causality). Thus

$$\hat{C}\psi(x)\hat{C}^{-1} = \eta_c \sum_{p,s} d^s(p)u^s(p)e^{-ip\cdot x} + {b^s}^{\dagger}(p)v^s(p)e^{ip\cdot x}.$$

Need to get plane-wave spinors in correct form. Recall that:

$$u^s(p) = \begin{pmatrix} \sqrt{\sigma\cdot p}\xi^s \\ \sqrt{\overline{\sigma}\cdot p}\xi^s \end{pmatrix}, \quad v^s(p) = \begin{pmatrix} \sqrt{\sigma\cdot p}\eta^s \\ -\sqrt{\overline{\sigma}\cdot p}\eta^s \end{pmatrix}.$$

WLOG, can choose a basis where $\eta^s = i\sigma^2(\xi^s)^*$, we can choose $C = -i\gamma^0\gamma^2$ and we can rotate space so that $p = (p^0, 0, 0, p^3)^T$. Then $u^s(p) =$

$$\begin{pmatrix} \sqrt{p^0 - p^3}\xi^1 \\ \sqrt{p^0 + p^3}\xi^2 \\ \sqrt{p^0 + p^3}\xi^1 \\ \sqrt{p^0 - p^3}\xi^2 \end{pmatrix} \Rightarrow {\overline{u}^s}^T(p) = {\gamma^0}^T u^{s*}(p) = \begin{pmatrix} \sqrt{p^0 + p^3}\xi^{1*} \\ \sqrt{p^0 - p^3}\xi^{2*} \\ \sqrt{p^0 - p^3}\xi^{1*} \\ \sqrt{p^0 + p^3}\xi^{2*} \end{pmatrix},$$

and finally

$$C{\overline{u}^s}^T(p) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{p^0 + p^3}\xi^{1*} \\ \sqrt{p^0 - p^3}\xi^{2*} \\ \sqrt{p^0 - p^3}\xi^{1*} \\ \sqrt{p^0 + p^3}\xi^{2*} \end{pmatrix} = v^s(p).$$

Similarly we find that $u^s(p) = C{\overline{v}^s}^T(p)$. It follows that $\hat{C}\psi(x)\hat{C}^{-1} = \eta_c C\overline{\psi}^T(x)$, as required. □

As with parity, the results extend to the Dirac conjugate $\overline{\psi}(x)$ and to Weyl spinors:

Theorem: Under charge conjugation we have:

$$\hat{C}\overline{\psi}(x)\hat{C}^{-1} = -\eta_C^* \psi^T(x) C^{-1},$$

and $\hat{C}\psi^T_{L/R}(x)\hat{C}^{-1} = \eta_C C \overline{\psi}^T_{R/L}(x)$.

Proof: $\overline{\psi}$ follows similarly to above. For Weyl spinors, use projection operators and recall $\gamma^5 C = C{\gamma^5}^T$. □

Note that $\psi_L^C(x) := \eta_C C \overline{\psi}^T_R(x)$ is *left-handed.* This can be verified as follows:

$$\gamma^5 C \overline{\psi}^T_R(x) = C{\gamma^5}^T \overline{\psi}^T_R(x) = C(\overline{\psi}_R(x)\gamma^5)^T = -C\overline{\psi}^T_R(x),$$

so we see that the field is indeed left-handed. So left-handed particles get mapped to left-handed antiparticles.

Theorem: If $\psi(x)$ satisfies the Dirac equation, then $\hat{C}\psi(x)\hat{C}^{-1} = \eta_C C \overline{\psi}^T(x)$ satisfies the Dirac equation.

Proof: We already saw this in the motivation for the charge conjugation matrix. □

Definition: A *Majorana fermion* has $b^s(p) = d^s(p)$, i.e. it is its own anti-particle. In this case, we find that $\hat{C}\psi(x)\hat{C}^{-1} = \psi(x)$ (since we can just terminate the above proof before we interfere with the plane wave spinors).

We can now deal with fermion bilinears. Note fermion bilinears are of the form $\overline{\psi}(x) X \psi(x)$ for some matrix X acting on spinors. So we prove the result:

Theorem: $\hat{C}\overline{\psi}(x) X \psi(x) \hat{C}^{-1} = \overline{\psi}(x) X_C \psi(x)$, where $X_C = C X^T C^{-1}$.

Proof: Insert $\hat{C}$ as appropriate (note $[\hat{C}, X] = 0$, since X is a matrix and $\hat{C}$ is an operator):

$$\begin{aligned}\hat{C}\overline{\psi}(x) X \psi(x) \hat{C}^{-1} &= \hat{C}\overline{\psi}(x)\hat{C}^{-1} X \hat{C}\psi(x)\hat{C}^{-1} \\ &= -\psi^T(x) C^{-1} X C \overline{\psi}^T(x).\end{aligned}$$

It's best to write this in index notation. Then we have

$$-\psi_\alpha(x)(C^{-1}XC)_{\alpha\beta}\overline{\psi}_\beta(x) = \overline{\psi}_\beta(x)((C^T X^T (C^T)^{-1})^T)_{\alpha\beta}\psi_\alpha(x)$$

where we lost the minus because fermions anti-commute. Applying the transpose in the middle swaps the indices, and we're left with a matrix $C^T X^T (C^T)^{-1}$ in the middle; recalling $C^T = -C$, we're done. □

Let's apply this to the familiar bilinears:

- $\overline{\psi}\psi$ has $X = I$. So $X_C = I$. Hence invariant under charge conjugation.
- $\overline{\psi}\gamma^5\psi$ has $X = \gamma^5$. So $X_C = C{\gamma^5}^T C^{-1} = \gamma^5 C C^{-1} = \gamma^5$. So invariant under charge conjugation.
- $\overline{\psi}\gamma^\mu\gamma^5\psi$ has $X = \gamma^\mu\gamma^5$. So $X_C = C{\gamma^5}^T{\gamma^\mu}^T C^{-1} = -\gamma^5\gamma^\mu = \gamma^\mu\gamma^5$. So invariant under charge conjugation.
- Finally, $\overline{\psi}\gamma^\mu\psi$ has $X = \gamma^\mu$. So $X_C = C{\gamma^\mu}^T C^{-1} = -\gamma^\mu$, so get minus sign under charge conjugation.

2.8 Time-reversal of scalar fields

Theorem: Under time reversal $\hat{T}$, the scalar field $\phi(x)$ transforms as

$$\hat{T}\phi(x)\hat{T}^{-1} = \eta_T \phi(x_T),$$

where $x_T = (-x^0, \mathbf{x})$, and η_T is a complex phase.

Proof: Same as all other proofs. Start from $\hat{T}a(p)\hat{T}^{-1} = \eta_T a(p_T)$ and $\hat{T}c^\dagger(p)\hat{T}^{-1} = \eta_T c^\dagger(p_T)$ where $p_T = (p^0, -\mathbf{p})$ (for $x \cdot p$ invariant). Only thing we need to be careful of is *anti-linearity* of $\hat{T}$, so complex numbers are conjugated when we pass $\hat{T}$ through them. Also need to use $p_T \cdot x = -p \cdot x_T$ at some point. □

2.9 Time-reversal of Dirac fields

As for charge conjugation, we find we need a matrix, called the *time-reversal matrix*. The argument runs as follows.

Under time reversal, we expect ψ to map to something like $B\psi(x_T)$, where $x_T = (-x^0, \mathbf{x})$. The matrix is needed to help us satisfy the Dirac equation. Recall the Dirac equation for ψ is $(i\partial\!\!\!/ - m)\psi(x) = 0$. Send $\psi(x) \mapsto \hat{T}\psi(x)\hat{T}^{-1}$; then in order to have invariance of the Dirac equation, we need

$$\hat{T}\big((-i\partial\!\!\!/^* - m)\psi(x)\big)\hat{T}^{-1} = 0,$$

recalling that $\hat{T}$ is anti-linear. Remove $\hat{T}$ and $\hat{T}^{-1}$ on both sides, and send $x \mapsto x_T$. We find that

$$(i{\gamma^0}^*\partial_0 - i{\gamma^i}^*\partial_i - m)\psi(x_T) = 0.$$

Inserting $B^{-1}B$ and multiplying through by B, we arrive at the Definition:

Definition: The *time-reversal matrix* is the matrix B satisfying $\gamma^{\mu *} = B(\gamma^0, -\boldsymbol{\gamma})B^{-1}$. Equivalently, this can be written as $B^{-1}\gamma^{\mu *}B = -\mathbb{T}^\mu{}_\nu\gamma^\nu$.

Theorem (Properties of B): We have the following:

(i) B always exists;

(ii) $B^{-1}\gamma^{5*}B = \gamma^5$;

(iii) In the chiral rep, a possible choice of B is $B = C^{-1}\gamma^5$, where C is the charge conjugation matrix.

Proof: (i) $\gamma^{\mu *}$ certainly satisfies the Clifford algebra, as does $(\gamma^0, -\boldsymbol{\gamma})$, so they must be related by a similarity transformation, B.

(ii) Write out γ^5 in full, and insert BB^{-1} in between all factors.

(iii) Just verify it satisfies the defining property for $\mu = 0, i$. □

We are now ready to prove the main result:

Theorem: Under time-reversal, $\psi(x)$ maps to

$$\hat{T}\psi(x)\hat{T}^{-1} = \eta_T B\psi(x_T).$$

Proof: Recall from the very start of this chapter that time-reversal *reverses the sign of angular momentum*. So we must take

$$\hat{T}b^s(p)\hat{T}^{-1} = \eta_T(-1)^{\frac{1}{2}-s}b^{-s}(p_T),$$
$$\hat{T}d^{s\dagger}(p)\hat{T}^{-1} = \eta_T(-1)^{\frac{1}{2}-s}d^{-s\dagger}(p_T).$$

The $(-1)^{\frac{1}{2}-s}$ is there for convenience. By using

$$B = C^{-1}\gamma^5 = \begin{pmatrix} i\sigma^2 & 0 \\ 0 & i\sigma^2 \end{pmatrix},$$

it can be shown similar to the charge conjugation proof that $(-1)^{\frac{1}{2}-s}u^{-s*}(p_T) = -Bu^s(p)$ and $(-1)^{\frac{1}{2}-s}v^{-s*}(p_T) = -Bv^s(p)$. The proof now proceeds through the standard steps (relabelling $s \mapsto -s$ at one point, and $x \cdot p_T = -x_T \cdot p$). □

Similarly, we find that $\hat{T}\overline{\psi}(x)\hat{T}^{-1} = \eta_T^*\overline{\psi}(x_T)B^{-1}$. The fermion bilinears are then dealt with in the same way as charge conjugation:

Theorem: We have $\hat{T}\overline{\psi}(x)X\psi(x)\hat{T}^{-1} = \overline{\psi}(x_T)X_T\psi(x_T)$, where $X_T = B^{-1}X^*B$.

Proof: Same idea as charge conjugation, but simpler; just remember $\hat{T}$ is anti-linear to produce X^*. □

Using this Theorem, we can show $\overline{\psi}\psi$ is invariant under time reversal. Also $\overline{\psi}\gamma^\mu\psi$ transforms to $-\mathbb{T}^\mu{}_\nu\overline{\psi}(x_T)\gamma^\nu\psi(x_T)$; this is sensible since $\mu = 0$ should act like charge density (i.e. invariant), whilst $\mu = i$ should act like current density (i.e. change of sign).

2.10 CPT properties of Maxwell's equation

In applications, we often want the transformations of the photon field $A_\mu(x)$ under $\hat{C}$, $\hat{P}$ and $\hat{T}$. These can be obtained by imposing the invariance of Maxwell's equations.

Theorem: $A_\mu(x)$ transforms under $\hat{C}$, $\hat{P}$ and $\hat{T}$ as:

- $\hat{C}A_\mu(x)\hat{C}^{-1} = -A_\mu(x)$;
- $\hat{P}A_\mu(x)\hat{P}^{-1} = \mathbb{P}^\nu{}_\mu A_\nu(x)$;
- $\hat{T}A_\mu(x)\hat{T}^{-1} = -\mathbb{T}^\nu{}_\mu A_\nu(x)$;

Proof: Use invariance of Maxwell's equation $\partial_\nu F^{\mu\nu} = e\overline{\psi}\gamma^\mu\psi$, which can be derived from the standard QED Lagrangian, where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

PARITY: Under parity, it's quick to see that $e\overline{\psi}\gamma^\mu\psi$ transforms to $e\mathbb{P}^\mu{}_\nu\overline{\psi}(x_P)\gamma^\nu\psi(x_P)$. The derivative transforms as $\partial_\nu \mapsto \mathbb{P}^\mu{}_\nu\partial_\mu$, and therefore Maxwell's equations transform to

$$\mathbb{P}^\rho{}_\nu\partial_\rho\tilde{F}^{\mu\nu}(x_P) = e\mathbb{P}^\mu{}_\nu\overline{\psi}(x_P)\gamma^\nu\psi(x_P).$$

Here, $\tilde{F}^{\mu\nu}$ is the parity-transformed field-strength tensor. In particular, consider the $\tilde{F}^{0i}$ component. The equations have become: $-\partial_i\tilde{F}^{0i}(x_P) = e\overline{\psi}(x_P)\gamma^0\psi(x_P)$, so for invariance of the Maxwell equations, we need $F^{0i}(x) \mapsto -F^{0i}(x_P)$ under parity.

In terms of the photon field, we have:

$$\tilde{F}^{0i}(x_P) = \partial^0\tilde{A}^i(x_P) + \partial^i\tilde{A}^0(x_P) = -F^{0i}(x_P),$$

and so it's clear that $A^i \mapsto -A^i$, whilst $A^0 \mapsto A^0$. Lowering indices, we get the result.

CHARGE CONJUGATION: Under charge conjugation, we see that $e\overline{\psi}\gamma^\mu\psi \mapsto -e\overline{\psi}(x)\gamma^\mu\psi(x)$. The LHS of Maxwell's equations transforms as $\partial_\nu F^{\mu\nu} \mapsto \partial_\nu\tilde{F}^{\mu\nu}$ since the derivatives are left unchanged.

Hence for invariance of Maxwell we need $F^{0i} \mapsto -F^{0i}$, and so $\tilde{F}^{0i} = \partial^0\partial^i - \partial^i\partial^0 = -F^{0i}$; from this we see that $A_\mu(x) \mapsto A_\mu(x)$ under charge conjugation.

TIME REVERSAL: Time reversal can be treated similar to parity to get the result. □

2.11 CPT properties of the scattering matrix

Definition: The S-matrix for a potential $V(t) = -\int d^3x\, \mathcal{L}_{\text{int}}$, is given by

$$S = \mathrm{T}\exp\left(-i\int_{-\infty}^{\infty} dt\, V(t)\right).$$

Consider, for example, $\mathcal{L}_I(x) = -e\overline{\psi}(x)\gamma^\mu A_\mu(x)\psi(x)$, the QED interaction Lagrangian. How does the S-matrix transform under $\hat{C}$, $\hat{P}$ and $\hat{T}$?

The Lagrangian is invariant under $\hat{P}$, $\hat{C}$ and $\hat{T}$. This is a consequence of fermion transformation properties, and the transformation properties of $A_\mu(x)$ which we derived above. It's thus clear that $V(t)$ is invariant under parity and charge, but goes to $V(-t)$ under time reversal.

Finally, it follows the S matrix must be invariant under charge and parity, but under time-reversal it's not so clear, since we have the time-ordered exponential.

Theorem: For the QED Lagrangian, $\hat{T}S\hat{T}^{-1} = S^\dagger$.

Proof: We have by Dyson's formula:

$$S = \sum_{n=0}^{\infty}(-i)^n \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 ... \int_{-\infty}^{t_{n-1}} dt_n V(t_1)V(t_2)...V(t_n),$$

which implies (since $\hat{T}$ is anti-linear): $\hat{T}S\hat{T}^{-1} =$

$$\sum_{n=0}^{\infty} i^n \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 ... \int_{-\infty}^{t_{n-1}} dt_n V(-t_1)V(-t_2)...V(-t_n).$$

Substitute $\tau_i = -t_{n+1-i}$ to get:

$$\int_{-\infty}^{\infty} d\tau_n \int_{-\tau_n}^{\infty} d\tau_{n-1} ... \int_{\tau_2}^{\infty} d\tau_1 V(\tau_n)...V(\tau_1).$$

Now use the identity (which can be proved graphically):

$$\int_{-\infty}^{\infty} d\tau_n \int_{\tau_n}^{\infty} d\tau_{n-1} = \int_{-\infty}^{\infty} d\tau_{n-1} \int_{-\infty}^{\tau_{n-1}} d\tau_n,$$

to swap the order of all the integrals, and swap all the limits at the same time. But notice that the final result is what we would have got had we initially taken the dagger instead (since $(AB)^\dagger = B^\dagger A^\dagger$, and all V are real). □

Theorem: Let $|\xi_T\rangle = \hat{T}|\xi\rangle$ and $|\eta_T\rangle = \hat{T}|\eta\rangle$. Then $\langle\eta_T|S|\xi_T\rangle = \langle\xi|S|\eta\rangle$ (for QED, i.e. $\hat{T}S\hat{T}^{-1} = S^\dagger$). That is, taking the time-reversal of the states and computing the amplitude is the same as watching the process in reverse.

Proof: Since we're working with $\hat{T}$, it's best to ditch Dirac's bra-ket notation. We have

$$\langle\eta_T|S|\xi_T\rangle = (\hat{T}\eta, S\hat{T}\xi) = (\hat{T}\eta, \hat{T}S^\dagger\xi).$$

This follows since $\hat{T}S\hat{T}^{-1} = S^\dagger$. Now use anti-unitary to get:

$$\langle\eta_T|S|\xi_T\rangle = (\eta, S^\dagger\xi)^* = (S^\dagger\xi, \eta) = (\xi, S\eta) = \langle\xi|S|\eta\rangle. \quad \square$$

This result holds generally if $\hat{T}\mathcal{L}_I(x)\hat{T}^{-1} = \mathcal{L}_I(x_T)$, that is, the interaction part of the Hamiltonian is invariant under time reversal. It's clear that the proof that $\hat{T}S\hat{T}^{-1} = S^\dagger$ works in this case just as before.

2.12 The CPT Theorem

CPT Theorem: Any Lorentz-invariant Lagrangian $\mathcal{L}$ with a Hermitian Hamiltonian is invariant under the product $\hat{\Theta} = \hat{C}\hat{P}\hat{T}$.

Proof: There isn't one! Only conjecture. □

Invoking the CPT Theorem can often speed up proofs, since we only need to verify two things are invariant, and invariance of the third follows.

2.13 Baryogenesis

Definition: *Baryogenesis* is the generation of the matter-antimatter asymmetry in the Universe.

There are three conditions necessary for baryogenesis, which come from violation of the symmetries:

The Sakarhov Conditions:

1. *Baryon number violation.* There must exist a process $X \to Y + B$ which yields a baryon excess B.

2. *Non-equilibrium.* The rate $\Gamma(Y + B \to X)$ must be strictly less than the rate $\Gamma(X \to Y + B)$ to prevent the process being undone.

3. *Charge violation and charge-parity violation.* If the Universe starts with equal numbers of X particles and $\overline{X}$ particles (X anti-particles), then the rate of baryon production is

$$\frac{dB}{dt} \propto \Gamma(X \to Y + B) - \Gamma(\overline{X} \to \overline{Y} + \overline{B}).$$

Hence there must be charge violation, else this is zero.

To see CP violation, assume for simplicity that $X \to B$ only, and that B is composed of n quarks. C violation implies that $\Gamma(X \to nq_L) \neq \Gamma(\overline{X} \to n\overline{q}_L)$. Under CP, there is a symmetry $q_L \mapsto \overline{q}_R$, so we can still get:

$$\Gamma(X \to nq_L) + \Gamma(X \to nq_R) = \Gamma(\overline{X} \to n\overline{q}_L) + \Gamma(\overline{X} \to n\overline{q}_R)$$

This would preclude baryogenesis, so we must have CP violation.

3 Spontaneous symmetry breaking

3.1 Discrete and continuous examples

Example 1: Consider a Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - V(\phi),$$

where

$$V(\phi) = \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4}\phi^4,$$

and $\lambda > 0$. The theory has a hidden discrete symmetry $\phi \mapsto -\phi$.

For $m^2 > 0$, the minimum of $V(\phi)$ is at $\phi = 0$. But for $m^2 < 0$, we can complete the square and drop the constant to get:

$$V(\phi) = \frac{\lambda}{4}\left(\phi^2 - v^2\right)^2,$$

where $v = \sqrt{-m^2/\lambda}$. There are now two degenerate minima $\phi = \pm v$. We say that ϕ has acquired a *non-zero vacuum expectation value*.

Choose the vacuum $\phi = v$. Perturbing around it $\phi = v + f$, the Lagrangian becomes:

$$\mathcal{L} = \frac{1}{2}\partial_\mu f \partial^\mu f - \lambda\left(v^2 f + v f^3 + \frac{1}{4}f^4\right).$$

We see f has mass-squared $m_f^2 = 2\lambda v^2$. Note that $f \mapsto -f$ is no longer a symmetry; the symmetry is *broken* by the non-zero vacuum expectation value.

Example 2: The above example generalises to many fields. Consider $\boldsymbol{\phi} = (\phi_1, ..., \phi_N)^T$, with Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu \boldsymbol{\phi} \cdot \partial^\mu \boldsymbol{\phi} - V(\boldsymbol{\phi}).$$

Here,

$$V(\boldsymbol{\phi}) = \frac{1}{2}m^2\boldsymbol{\phi}^2 + \frac{\lambda}{4}\boldsymbol{\phi}^4,$$

and $\lambda > 0$. The theory has a hidden $O(N)$ symmetry of the fields. Exactly as before, $\boldsymbol{\phi} = \mathbf{0}$ is the vacuum for $m^2 > 0$, but for $m^2 < 0$, the minima of the potential are described by $\boldsymbol{\phi}^2 = v^2$ for $v = \sqrt{-m^2/\lambda}$.

Choose the vacuum $\boldsymbol{\phi} = (0, 0, ..., 0, v)^T$ and perturb to $\boldsymbol{\phi} = (\pi_1(x), ..., \pi_{N-1}(x), v + \sigma(x))^T$. The Lagrangian becomes

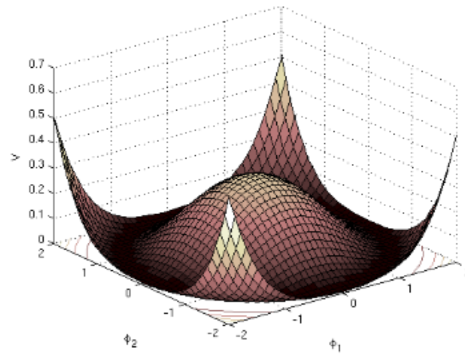
$$\mathcal{L} = \frac{1}{2}\partial_\mu \boldsymbol{\pi} \cdot \partial^\mu \boldsymbol{\pi} + \frac{1}{2}\partial_\mu \sigma \partial^\mu \sigma - V(\boldsymbol{\pi}, \sigma),$$

where

$$V(\boldsymbol{\pi}, \sigma) = \frac{1}{2}m_\sigma^2\sigma^2 + \lambda v(\sigma^2 + \boldsymbol{\pi}^2)\sigma + \frac{\lambda}{4}(\sigma^2 + \boldsymbol{\pi}^2)^2.$$

The σ field acquires a mass-squared $m_\sigma^2 = 2\lambda v^2$, but all the π_i fields are massless.

This makes sense since the potential is like the bottom of a wine-bottle; radial excitations climb the walls and have high energy cost (manifested as mass), but angular excitations have no energy cost.



Note that the final Lagrangian has a hidden $O(N-1)$ from transformation of the $\boldsymbol{\pi}$ fields, but the $O(N)$ symmetry has been broken.

3.2 Classical form of Goldstone's Theorem

The examples above can be generalised dramatically to *Goldstone's Theorem*.

Definition: A transformation $\phi(x) \mapsto g\phi(x)$ of fields is a *symmetry* of the Lagrangian if $V(g\phi) = V(\phi)$. The set of all symmetries is the *symmetry group* of the Lagrangian.

Definition: Let V_{min} be the minimum value of $V(\phi)$. The manifold $\Phi_0 = \{\phi_0 : V(\phi_0) = V_{\text{min}}\}$ is called the *vacuum manifold*.

Definition: Suppose we fix a vacuum ϕ_0. The *invariant subgroup* (or *stability group*) of the vacuum ϕ_0 is the subgroup $H = \{h \in G : h\phi_0 = \phi_0\} \leq G$.

In general, we assume that H is normal in G, $H \trianglelefteq G$, and that G acts *transitively* on Φ_0.

Theorem: The stability groups of all vacua $\phi_0 \in \Phi_0$ are isomorphic.

Proof: Let $\phi_0' = g\phi_0$ (assumes transitivity!!), and suppose H' is the stability group of ϕ_0', H is the stability group of ϕ_0. It's intuitively clear that $H' \cong gHg^{-1}$, since $h'\phi_0' = \phi_0'$ and $ghg^{-1}\phi_0' = gh\phi_0 = g\phi_0 = \phi_0'$. Since H is normal, $gHg^{-1} \cong H$ and we're done. $\square$

Theorem: For any stability group H, we have $\Phi_0 \cong G/H$.

Proof: Fix a vacuum ϕ_0. Let ϕ_0' be any other vacuum. Suppose $\phi_0' = g_1\phi_0 = g_2\phi_0$. Then $g_2^{-1}g_1 \in H$, so g_2 and g_1 are in the same coset. So going from ϕ_0 to any other point is specified by a particular coset, hence $\Phi_0 \cong G/H$. $\square$

Goldstone's Theorem: Under spontaneous symmetry breaking of a continuous symmetry G of the Lagrangian to a subgroup H, there are at least $\dim(G) - \dim(H)$ massless modes.

Proof: Since the symmetry is continuous, write it as $g\phi = \phi + \delta\phi$, where $\delta\phi = i\alpha^a t^a \phi$. Here, α^a are infinitesimal parameters and it^a generate the Lie algebra of G.

Since g is a symmetry, we have

$$V(\phi + \delta\phi) - V(\phi) = i\alpha^a (t^a\phi)_r \frac{\partial V}{\partial \phi_r} = 0. \qquad (*)$$

Let ϕ_0 be the broken vacuum. Expanding V around ϕ_0,

$$V(\phi) - V(\phi_0) = \frac{1}{2}(\phi - \phi_0)_r \left.\frac{\partial^2 V}{\partial\phi_r \partial\phi_s}\right|_{\phi=\phi_0} (\phi - \phi_0)_s + \cdots .$$

There is no first derivative term, since ϕ_0 is a minimum of V. The quadratic term is the *mass* term, as we saw in the examples. Write

$$\mathcal{M}^2_{rs} = \left.\frac{\partial^2 V}{\partial\phi_r \partial\phi_s}\right|_{\phi=\phi_0}$$

to denote the *mass-squared matrix*. Go back to $(*)$ and differentiate wrt ϕ_s, and evaluate at $\phi = \phi_0$:

$$i\alpha^a (t^a\phi)_r \left.\frac{\partial^2 V}{\partial\phi_r \partial\phi_s}\right|_{\phi=\phi_0} = i\alpha^a (t^a\phi)_r \mathcal{M}^2_{rs} = 0.$$

Note there is no other term in the product rule expansion, since $\partial V/\partial\phi_r = 0$ at $\phi = \phi_0$. Since this holds for all α^a, this reduces to $\mathcal{M}^2_{sr}(t^a\phi)_r = 0$.

If the symmetry is unbroken, then $g\phi_0 = \phi_0$ for all choices of vacuum, implying $t^a\phi \equiv 0$ for all a. Else, there exists $g \in G$ such that $g\phi_0 \neq \phi_0$, so that $t^a\phi_0 \neq 0$ for some a values. Then $t^a\phi$ is an eigenvector of $\mathcal{M}^2$ with evalue 0.

Let $\{i\tilde{t}^i, i\theta^{\tilde{a}}\}$ be the basis of the Lie algebra, where $\tilde{t}^i\phi$ preserve the symmetry via $\tilde{t}^i\phi_0 = 0$, and $\theta^{\tilde{a}}\phi$ are eigenvectors of $\mathcal{M}^2$ with zero eigenvalue. WLOG, we can choose the $\theta^{\tilde{a}}$ to be orthogonal to $\tilde{t}^i$ wrt the Killing form, i.e.

$$\mathrm{Tr}(\tilde{t}^i\theta^{\tilde{a}}) = 0.$$

Then $i\tilde{t}^i$ generate the Lie algebra of the stability group H, and hence there are $\dim(H)$ of them. There are $\dim(G)$ generators in total, so it follows there are at least $\dim(G) - \dim(H)$ massless modes, as required. □

If $\mathcal{M}^2$ is $N \times N$, we should generally expect there to be $N - (\dim(G) - \dim(H))$ *massive* modes, but it is possible some of these end up massless anomalously.

Definition: The massless modes coming from Goldstone's Theorem are called *Goldstone bosons*.

3.3 Many examples

Example 1: Let M be an $N \times N$ matrix, and consider:

$$\mathcal{L} = \mathrm{Tr}(\partial^\mu M^\dagger \partial_\mu M) - \frac{1}{2}\lambda \mathrm{Tr}(M^\dagger M M^\dagger M) - k\mathrm{Tr}(M^\dagger M),$$

for $\lambda > 0$. A general symmetry looks like $M \mapsto AMB^{-1}$, and it's clear that $A, B \in U(N)$ for this to stand a chance of being a symmetry. Is the symmetry group $U(N)\times U(N)$?

No. Consider the homomorphism $\phi : U(N) \times U(N) \to G$ where G is the symmetry group of $\mathcal{L}$. Assuming this is surjective, the first isomorphism theorem gives

$$G \cong \frac{U(N) \times U(N)}{\mathrm{Ker}(\phi)}.$$

The kernel of ϕ is $\{(A, B) : \phi(A, B) = \mathrm{id}\}$. Explicitly, (A, B) is in the kernel iff $AMB^\dagger = M$ for all M.

Choose $M = I$. Then $A = B$. So $AMA^\dagger = M \Rightarrow [A, M] = 0$ for all M, and hence $A \propto \lambda I$ by Schur's Lemma. But $A \in U(N)$, so λ must be a phase. Hence $\mathrm{Ker}(\phi) = \{(e^{i\theta}I, e^{i\theta}I)\} \cong U(1)$. Thus

$$G \cong \frac{U(N) \times U(N)}{U(1)}.$$

For $k < 0$, this theory undergoes SSB. Completing the square, we find

$$V(\phi) = -\frac{1}{2}\lambda \mathrm{tr}\left(\left(M^\dagger M + \frac{k}{\lambda}I\right)^2\right) + \frac{1}{2}\frac{k^2 N}{\lambda}.$$

We want the trace's argument to be zero, since $M^\dagger M + \frac{k}{\lambda}I$ is a Hermitian matrix so has real evalues, so $\left(M^\dagger M + \frac{k}{\lambda}I\right)^2$ has non-negative evalues. So the vacuum manifold is $M^\dagger M = -(k/\lambda)I = v^2 I$, where $v = \sqrt{-k/\lambda}$. This can be rewritten as $\Phi_0 = \{vM : M \in U(N)|\}$.

To find the invariant subgroup H, we can pick any $\phi_0 \in \Phi_0$. Select vI, and consider transformations $vI \mapsto AvIB^{-1}$. For $vI \in H$, we need $AvIB^{-1} = vI$, i.e. $A = B$. So

$$H = \{(A, B) \in (U(N) \times U(N))/U(1) : A = B\} \cong \frac{U(N)}{U(1)}.$$

Goldstone's Theorem then implies there are $\dim(G) - \dim(H) = 2N^2 - 1 - N^2 + 1 = N^2$ Goldstone bosons.

Example 2: Consider the same Lagrangian with the additional term

$$\mathcal{L}' = h(\det(M) + \det(M^\dagger)).$$

For (A, B) to give a symmetry, we now need $\det(M) = \frac{\det(A)}{\det(B)}\det(M)$ for all M, i.e. $\det(A) = \det(B)$. Write $A = e^{i\theta}A'$ and $B = e^{i\phi}B'$, where $A', B' \in SU(N)$ (this is possible since $|\det(U)| = 1$ for $U \in U(N)$).

The determinant condition then gives $e^{iN\theta} = e^{iN\phi}$, which implies $\theta = \phi + 2\pi ik/N$. So our symmetry transformations are specified by $(A', B', \theta, k) \in SU(N) \times SU(N) \times U(1) \times \mathbb{Z}_N$. Is this the symmetry group?

Let $\phi : SU(N) \times SU(N) \times U(1) \times \mathbb{Z}_N \to G$. The kernel is given by (A, B, θ, k) such that

$$e^{i\theta} A' M e^{-i\theta} {B'}^{\dagger} e^{2\pi ik/N} = M, \text{ for all } M.$$

Clearly this holds for all θ. We can also absorb $e^{2\pi ik/N}$ into A' to just get another element of $SU(N)$. So can immediately quotient out $U(1) \times \mathbb{Z}_N$.

Left with $A' M {B'}^{\dagger} = M$. Choosing $M = I$ gives $A' = B'$. Then $[A', M] = 0$ so by Schur's Lemma $A' = e^{2\pi ik/N} I$, since $A' \in SU(N)$. Hence the kernel is $\{(e^{2\pi ik/N}, e^{2\pi ik/N})\} \cong \mathbb{Z}_N$. Hence the symmetry group is

$$G = \frac{SU(N) \times SU(N)}{\mathbb{Z}_N}.$$

The vacuum manifold remains $M^\dagger M = v^2 I$, so choose vacuum vI WLOG. Then the invariant subgroup is $(A, B) \in G$ such that

$$AvIB^{-1} = vI \Rightarrow A = B.$$

Hence the invariant subgroup is $H = SU(N)/\mathbb{Z}_N$. So there are $\dim(G) - \dim(H) = 2N^2 - 2 - N^2 + 1 = N^2 - 1$ Goldstone bosons.

Example 3: Consider a theory of five scalars ϕ_a, $a = 1, ..., 5$ expressed as a symmetric traceless 3×3 matrix

$$\Phi = \sum_{a=1}^{5} \phi_a t_a,$$

where t_a are a basis of symmetric traceless matrices obeying $\mathrm{Tr}(t_a t_b) = \delta_{ab}$. The theory has Lagrangian

$$\mathcal{L} = \frac{1}{2}\mathrm{Tr}\left(\partial^\mu \Phi \partial_\mu \Phi\right) - g\left(\frac{1}{4}\mathrm{Tr}(\Phi^4) + \frac{1}{3} b \mathrm{Tr}(\Phi^3) + \frac{1}{2} c \mathrm{Tr}(\Phi^2)\right),$$

where $g > 0$. It's trivial to see this has an $SO(3)$ symmetry $\Phi \mapsto A\Phi A^{-1}$, $A \in SO(3)$.

Let $\mathcal{M}_0$ be the vacuum manifold. If $\Phi_0, \Phi_0' \in \mathcal{M}_0$, then $\Phi_0' = A\Phi_0 A^T$ for some $A \in SO(3)$, assuming transitivity. So Φ_0 and Φ_0' are similar so have the same evalues.

Also $\mathrm{Tr}(\Phi) = 0$, all evalues of Φ sum to zero. Hence all elements of vacuum manifold have same evalues, which sum to zero. Let the eigenvalues be λ_1, λ_2 and λ_3, and choose to work with vacuum

$$\Phi_0 = \mathrm{diag}\{\lambda_1, \lambda_2, \lambda_3\}.$$

There are three cases for the invariant symmetry group:

1. All evalues are equal. Then $\lambda_1 = \lambda_2 = \lambda_3 = 0$. So for all $A \in SO(3)$, we have $A\Phi_0 A^{-1} = \Phi_0$, and it follows that $H = SO(3)$. There is no SSB.
2. Two evalues are equal. We have $\Phi_0 = \mathrm{diag}\{\lambda', \lambda, \lambda\}$. It's straightforward to see that
$$A = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{A} \end{pmatrix} \in H$$
for $\tilde{A} \in SO(2)$. So $H \cong SO(2)$.
3. All evalues are distinct. Then it can be shown that the only $A \in SO(3)$ with $A\Phi_0 A^{-1} = \Phi_0$ is $A = I$. Hence $H = \{e\}$.

To get further constraints on λ_i, we need to calculate $\mathcal{M}_0$ directly using a Lagrange multiplier μ to ensure $\mathrm{Tr}(\Phi) = 0$. We find

$$0 = \Phi^3 + b\Phi^2 + c\Phi - \mu I$$

is the vacuum manifold. Taking the trace, we see

$$3\mu = \mathrm{Tr}(\Phi^3) + b\mathrm{Tr}(\Phi^2) + c\mathrm{Tr}(\Phi).$$

For example, for $H \cong SO(2)$, we set $\Phi = \Phi_0 = \mathrm{diag}(\lambda, \lambda, -2\lambda)$. Inserting this into both equations above, we can eliminate μ to derive $0 = \lambda(3\lambda^2 - b\lambda + c)$. So we get $\lambda = 0$ (i.e. no SSB) or $3\lambda^2 - b\lambda + c = 0$.

Using the discriminant, this has solutions iff $b^2 \geq 12c$. If this is fulfilled, it's possible to get $H \cong SO(2)$.

Example 4: Consider $b = 0$ in Example 3. We use the identity

$$\mathrm{Tr}(M^4) = \frac{1}{2}(\mathrm{Tr}(M^2))^2$$

for traceless matrices to rewrite the Lagrangian as

$$\mathcal{L} = \frac{1}{2}\mathrm{Tr}(\partial^\mu \Phi \partial_\mu \Phi) - \frac{1}{8}(\mathrm{Tr}(\Phi^2))^2 - \frac{1}{2} g c \mathrm{Tr}(\Phi^2).$$

Note $\mathrm{Tr}(\Phi^2) = \mathrm{Tr}(\phi_a \phi_b t_a t_b) = \phi_a \phi_b \delta_{ab} = \boldsymbol{\phi} \cdot \boldsymbol{\phi}$, so we can again rewrite the Lagrangian as:

$$\mathcal{L} = \frac{1}{2}\partial^\mu \boldsymbol{\phi} \cdot \partial_\mu \boldsymbol{\phi} - \frac{1}{8} g (\boldsymbol{\phi} \cdot \boldsymbol{\phi})^2 - \frac{1}{2} c \boldsymbol{\phi} \cdot \boldsymbol{\phi}.$$

It's then evident that the theory has the larger symmetry group $SO(5)$.

Minimising the potential, we find the vacuum manifold is a subset of $0 = \frac{1}{2}\boldsymbol{\phi}|\boldsymbol{\phi}|^2 + c\boldsymbol{\phi}$. If $c > 0$, get $\boldsymbol{\phi} = \mathbf{0}$, but if $c < 0$ we get SSB, with vacuum manifold $|\boldsymbol{\phi}|^2 = -2c$. So clearly $\mathcal{M}_0$ is a 4-sphere. It's possible to show that $SO(5)/SO(4) \cong S^4$, so $H = SO(4)$. There are $\dim(SO(5)) - \dim(SO(4)) = 10 - 6 = 4$ Goldstone bosons.

3.4 Quantum form of Goldstone's Theorem

Theorem: Let G be the Lie group of symmetries of a Lagrangian $\mathcal{L}$, and suppose it is spontaneously broken to the Lie group H by the non-zero VEV $\langle 0|\phi(x)|0\rangle = \phi_0 \neq 0$. That is, $\langle 0|h\phi(x)|0\rangle = \phi_0 \neq 0$ for all $h \in H$, but $\langle 0|g\phi(x)|0\rangle = 0$ for all $g \in G\backslash H$. Assume also that:

1. we are working with a Lorentz covariant theory;
2. all states in our Hilbert space have non-negative norm.

Then there are at least $\dim(G) - \dim(H)$ Goldstone bosons.

Proof: The Lie algebra of G is generated by it^a, $a = 1, ..., \dim(G)$, and the Lie algebra of H is generated by $i\tilde{t^i}$, $i = 1, ..., \dim(H)$, where $t^a, \tilde{t}^i$ are Hermitian.

Since each generator gives a symmetry, by Noether's Theorem we have conserved currents and charges:

$$j^{a\mu}(x) = i\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}t^a\phi,$$
$$Q^a = \int_{\mathbb{R}^3} d^3\mathbf{x}\, j^{a0}(x) = i\int_{\mathbb{R}^3} d^3\mathbf{x}\, \pi(x)t^a\phi(x).$$

Using the relation $[\phi(\mathbf{x},t), \pi(\mathbf{x}',t)] = i\delta(\mathbf{x} - \mathbf{x}')$, we can show that $[Q^a, \phi(0)] = t^a\phi(0)$ (possible to use equal time commutation relations since Q^a is conserved). We relate these to the currents via

$$\langle 0|[Q^a, \phi(0)]|0\rangle = \int d^3\mathbf{x}\, C^{a0}(x),$$

where $C^{a\mu} = \langle 0|[j^{a\mu}(x), \phi(0)]|0\rangle$.

We now try to compute $C^{a\mu}$. Insert a complete set of states to get

$$C^{a\mu} = \sum_n \left(\langle 0|j^{a\mu}(x)|n\rangle\langle n|\phi(0)|0\rangle - \langle 0|\phi(0)|n\rangle\langle n|j^{a\mu}(x)|0\rangle\right).$$

The sum over n here includes integrals over $d^3\mathbf{p}$, etc, depending on the form of $|n\rangle$. Define the *spectral density functions* by

$$i\rho^{a\mu}(k) = (2\pi)^3\sum_n \delta^4(k - p_n)\langle 0|j^{a\mu}(0)|n\rangle\langle n|\phi(0)|0\rangle,$$
$$i\tilde{\rho}^{a\mu}(k) = (2\pi)^3\sum_n \delta^4(k - p_n)\langle 0|\phi(0)|n\rangle\langle n|j^{a\mu}(0)|0\rangle.$$

Recalling $j^{a\mu}(x) = e^{iP\cdot x}j^{a\mu}(0)e^{-iP\cdot x}$ (since momentum generates translations in quantum theory), so that

$$\langle 0|j^{a\mu}(x)|n\rangle = \langle 0|e^{iP\cdot x}j^{a\mu}(0)e^{-iP\cdot x}|n\rangle = \langle 0|j^{a\mu}(x)|n\rangle\, e^{-ip_n\cdot x},$$

where p_n is the momentum of state $|n\rangle$ (assuming $|0\rangle$ has zero momentum). Thus we can then write $C^{a\mu}$ as

$$C^{a\mu} = i\int\frac{d^4k}{(2\pi)^3}\left(\rho^{a\mu}(k)e^{-ik\cdot x} - \tilde{\rho}^{a\mu}(k)e^{ik\cdot x}\right).$$

The spectral densities only depend on k, so if they are to be *Lorentz covariant*, they must be proportional to k^μ. Also, *physical states* have $k^0 > 0$ (i.e. non-negative energy), hence we have:

$$\rho^{a\mu}(k) = k^\mu\Theta(k^0)\rho^a(k^2), \quad \tilde{\rho}^{a\mu}(k) = k^\mu\Theta(k^0)\tilde{\rho}^a(k^2).$$

Thus

$$C^{a\mu} = -\partial^\mu\int\frac{d^4k}{(2\pi)^3}\Theta(k^0)\left(\rho^a(k^2)e^{-ik\cdot x} + \tilde{\rho}^a(k^2)e^{ik\cdot x}\right).$$

Recall the propagator in QFT had the integral expression:

$$D(x-y;\sigma) = \langle 0|\phi(x)\phi(y)|0\rangle = \int\frac{d^3p}{(2\pi)^3p^0}e^{-ip\cdot(x-y)}\bigg|_{p^0=\sqrt{|\mathbf{p}|^2+\sigma}}$$
$$= \int\frac{d^4p}{(2\pi)^3}\Theta(p^0)\delta(p^2-\sigma)e^{-ip\cdot(x-y)},$$

where σ is the mass-squared of the ϕ field. Writing

$$\rho(k^2) = \int d\sigma\rho(\sigma)\delta(k^2-\sigma),$$

we can reduce $C^{a\mu}$ to a propagator:

$$C^{a\mu} = -\partial^\mu\int d\sigma\big(\rho^a(\sigma)D(x;\sigma) + \tilde{\rho}^a(\sigma)D(-x;\sigma)\big).$$

For all spacelike x, i.e. $x^2 < 0$, we have $D(x;\sigma) = D(-x;\sigma)$. But when $x^2 < 0$, $C^{a\mu}$ is a commutator of spacelike-separated operators, and hence must vanish. So $\rho^a(\sigma) = -\tilde{\rho}^a(\sigma)$ holds when x is spacelike.

But neither ρ^a or $\tilde{\rho}^a$ depend on x. So this relation holds everywhere. Hence

$$C^{a\mu} = -\partial^\mu\int d\sigma\, \rho^a(\sigma)i\Delta(x;\sigma),$$

where $i\Delta(x;\sigma) = D(x;\sigma) - D(-x;\sigma)$ is the Feynman propagator. Conservation of the current implies $\partial_\mu j^{a\mu} = 0$, so $\partial_\mu C^{a\mu} = 0$. Thus:

$$-\partial^2\int d\sigma\, \rho^a(\sigma)i\Delta(x;\sigma) = -\int d\sigma\, \rho^a(\sigma)i\partial^2\Delta(x;\sigma) = 0.$$

Recall the Feynman propagator satisfies Klein-Gordon, so $(\partial^2+\sigma)\Delta = 0$, and we can rewrite the condition as

$$\int d\sigma\, \sigma\rho^a(\sigma)i\Delta(x;\sigma) = 0.$$

This must hold for all x, including timelike x where $\Delta(x;\sigma) \neq 0$. Hence $\sigma\rho^a(\sigma) = 0$ (assuming $\rho(\sigma) > 0$, which is true since the original spectral densities were defined in terms of norms of states). There are 2 cases:

1. $\rho^a(\sigma) = 0$. Then $C^{a\mu} = 0$, and so $\langle 0|t^a\phi(0)0\rangle = \langle 0|[Q^a, \phi(0)]|0\rangle = 0$. So t^a is an unbroken generator.
2. $\rho^a(\sigma) = N^a\delta(\sigma)$, for $N^a \neq 0$. Then $C^{a\mu} = -iN^a\partial^\mu\Delta(x;\sigma)$, and so
$$\langle 0|[Q^a, \phi(0)]|0\rangle = -iN^a\int d^3\mathbf{x}\, \partial^0\Delta(x;0).$$

Need to evaluate the RHS. Recall

$$i\Delta(x;0) = D(x;\sigma) - D(-x;\sigma) = \int \frac{d^4k}{(2\pi)^3}\delta(k^2-\sigma)\epsilon(k^0)e^{-ik\cdot x},$$

where $\epsilon(k^0) = +1$ for $k^0 > 0$ and -1 for $k^0 < 0$. Integrating over $d^3\mathbf{x}$, we can convert $e^{-i\mathbf{k}\cdot\mathbf{x}}$ into a delta function. We also introduce a limit as a technical trick:

$$\begin{aligned}\int d^3\mathbf{x}\; i\Delta(x;0) &= \lim_{\sigma\to 0}\int dk^0\delta((k^0)^2-\sigma)\epsilon(k^0)e^{-ik^0x_0} \\ &= \lim_{\sigma\to 0}\int dk^0\left(\frac{\delta(k^0-\sqrt{\sigma})}{|2\sqrt{\sigma}|}+\frac{\delta(k^0+\sqrt{\sigma})}{|2\sqrt{\sigma}|}\right)e^{-ik^0x_0} \\ &= -\lim_{\sigma\to 0}\frac{1}{\sqrt{\sigma}}\sin(x_0\sqrt{\sigma}) = -x_0.\end{aligned}$$

Therefore, we have

$$\langle 0|[Q^a,\phi(0)]|0\rangle = -iN^a\int d^3\mathbf{x}\;\partial^0\Delta(x;0) = iN^a \neq 0.$$

So t^a is a broken generator.

We still need to count the Goldstone bosons. We'll do so by going back to the spectral density functions $\rho^{a\mu}$, $\tilde{\rho}^{a\mu}$. Label the states with non-zero contribution to ϕ_0 by $B(p)$, and let (permitted by Lorentz covariance):

$$\langle 0|j^{a\mu}(0)|B(p)\rangle = iF^a_B p^\mu, \quad \langle B(p)|\phi(0)|0\rangle = Z^B.$$

Note that $|B(p)\rangle$ is spinless, since $\phi(0)\,|0\rangle$ and Z^B are rotationally invariant. Note also that $|B(p)\rangle$ is massless, since in the case where we have a ϕ_0 contribution, we must have $\rho(\sigma) = N^a\delta(\sigma)$, which only contributes when the ϕ field mass is $\sigma = 0$ (this carries across to $\rho^{a\mu}(k) = k^\mu\Theta(k^0)\rho^a(k^2) = k^\mu\Theta(k^0)N^a\delta(k^2)$, showing we only get a contribution to $C^{a\mu}$, and hence to ϕ_0, when $k^2 = 0$, i.e. the field is massless).

We need to count the $|B(p)\rangle$, which we are now certain are the Goldstone bosons. Recall that

$$i\rho^{a\mu}(k) = ik^\mu\Theta(k^0)N^a\delta(k^2) = \int\frac{d^3\mathbf{p}}{2|\mathbf{p}|}\delta^4(k-p)ik^\mu N^a,$$

and

$$\begin{aligned}i\rho^{a\mu}(k) &= \sum_B\int\frac{d^3\mathbf{p}}{2|\mathbf{p}|}\delta^4(k-p)\,\langle 0|j^{a\mu}(0)|B(p)\rangle\,\langle B(p)|\phi(0)|0\rangle \\ &= \int\frac{d^3\mathbf{p}}{2|\mathbf{p}|}\delta^4(k-p)ik^\mu\sum_B F^a_B Z^B,\end{aligned}$$

by definitions. These are equal for arbitrary k, and so

$$N^a = \sum_B F^a_B Z^B.$$

The unbroken generators span H, so there are $\dim(H)$ unbroken generators, and $n = \dim(G) - \dim(H)$ broken generators, i.e. n components of N^a. So F^a_B has rank n. Since row rank is the same as column rank, and column rank is the dimension of the space spanned by the columns, we must have at least n B labels. So there are at least n Goldstone bosons. □

3.5 The Abelian Higgs mechanism

Gauge theories can violate the conditions of Goldstone's Theorem. For example, they can contain states with negative norm (*ghosts*) or can have non-Lorentz invariant gauge conditions.

Example: Consider scalar electrodynamics, described by Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi)^*(D^\mu\phi) - V(\phi^*\phi),$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $D_\mu = \partial_\mu + iqA_\mu$. This theory has a $U(1)$ gauge symmetry, given by

$$\phi(x)\mapsto e^{i\alpha(x)}\phi(x), \qquad A_\mu \mapsto A_\mu - \frac{1}{q}\partial_\mu\alpha(x).$$

Take $V(\phi^*\phi) = \mu^2|\phi|^2 + \lambda|\phi|^4$, with $\lambda > 0$, and consider how μ^2 changes.

- If $\mu^2 > 0$, then $\phi = 0$ is the unique vacuum and $\mu^2|\phi|^2$ is the usual mass term for the complex scalar ϕ. The photon is massless (no quadratic A_μ term) and ϕ has mass-squared μ^2.
- If $\mu^2 < 0$, then the vacuum manifold is $|\phi|^2 = -\mu^2/2\lambda = v^2/2$, for $v^2 = -\mu^2/\lambda$. WLOG, set $\phi = v/\sqrt{2}$, and expand around the minimum:

$$\phi = \frac{1}{\sqrt{2}}e^{i\theta(x)/v}(v+\eta(x)) = \frac{1}{\sqrt{2}}(v+\eta(x)+i\theta(x)).$$

Here, $\eta(x)$ is real and $\theta(x)$ is real. The Lagrangian becomes:

$$\begin{aligned}\mathcal{L} = \frac{1}{2}(\partial_\mu\eta\partial^\mu\eta - 2v^2\lambda\eta^2) + \frac{1}{2}(\partial_\mu\theta)(\partial^\mu\theta) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ + qvA_\mu\partial^\mu\theta + \frac{1}{2}q^2v^2A_\mu A^\mu + \mathcal{L}_{\text{int}},\end{aligned}$$

where $\mathcal{L}_{\text{int}}$ denotes interaction terms. We see η has mass $\sqrt{2v^2\lambda}$, A_μ has mass qv (i.e. the photon has become massive!) and θ is massless.

Now note we can factorise the Lagrangian as:

$$\begin{aligned}\mathcal{L} = \frac{1}{2}(\partial_\mu\eta\partial^\mu\eta - 2v^2\lambda\eta^2) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ +\frac{1}{2}q^2v^2\left(A_\mu + \frac{1}{vq}\partial_\mu\theta\right)\left(A^\mu + \frac{1}{vq}\partial^\mu\theta\right) + \mathcal{L}_{\text{int}},\end{aligned}$$

Since this is a gauge theory, we're free to first pick a specific gauge. Choose to transform to *unitary gauge*, $\alpha = -\theta/v$. Then

$$\phi\mapsto\phi' = e^{-i\theta/v}\phi = \frac{1}{\sqrt{2}}(v+\eta), \quad A_\mu\mapsto A'_\mu = A_\mu + \frac{1}{vq}\partial_\mu\theta.$$

This gives the Lagrangian

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu\eta\partial^\mu\eta - 2\lambda v^2\eta^2\right) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}q^2v^2A'_\mu A'^\mu + \mathcal{L}_{\text{int}}.$$

The Goldstone boson, θ, is no longer present!

Definition: If choosing a gauge eliminates a Goldstone boson, we say the gauge field *eats* the Goldstone boson.

The above SSB process is called the *Abelian Higgs mechanism*. It gives a photon of mass $m_\gamma = qv$. It clearly doesn't happen in real life.

3.6 Review of non-Abelian gauge theory

In the Standard Model, we have a similar SSB process to the above, but for a non-Abelian gauge theory. Hence, let's recap non-Abelian gauge theories.

Definition: In a *gauge theory* with *gauge group* G, fields $\psi_i(x)$ transform as $\psi_i(x) \mapsto U_{ij}(x)\psi_j(x)$, where $U_{ij}(x)$ is in some representation of G.

If it^a are the generators the Lie algebra of G, this transformation law becomes $\psi_i(x) \mapsto \exp(it^a\theta^a(x))_{ij}\psi_j(x)$. The conjugate field transforms as $\overline{\psi}_i(x) \mapsto \overline{\psi}_j(x)(U^\dagger)_{ji} = \overline{\psi}_j \exp(-it^a\theta^a(x))_{ji}$.

Recall the generators it^a obey $[t^a, t^b] = if^{abc}t^c$, where f^{abc} are called the *structure constants* of the Lie algebra. The generators may also be chosen to obey the orthogonality condition

$$\mathrm{Tr}(t^a t^b) = T(R)\delta^{ab},$$

where $T(R)$ is a constant dependent on the representation, called the *Dynkin index*. For the fundamental representation, $T(R) = \frac{1}{2}$.

In order for a theory to be invariant under gauge transformations, we need to promote the derivatives to *covariant derivatives*.

Definition: The *covariant derivative* of a gauge theory is

$$(D_\mu)_{ij} = \partial_\mu\delta_{ij} + ig(t^a A^a_\mu)_{ij},$$

where A^a_μ is a set of *gauge fields*, which transform under gauge transformations as

$$(t^a A^a_\mu)_{ij} \mapsto U t^a A^a_\mu U^{-1} + \frac{i}{g}(\partial_\mu U)U^{-1},$$

or infinitesimally as

$$(t^a A^a_\mu)_{ij} \mapsto (t^a A^a_\mu)_{ij} - \frac{1}{g}\partial_\mu\theta^a - f^{abc}\theta^b A^c_\mu.$$

The gauge fields require a kinetic term to be included in the Lagrangian. Define the *field-strength tensor* by

$$igt^a F^a_{\mu\nu} = [D_\mu, D_\nu] \;\Rightarrow\; F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - gf^{abc}A^b_\mu A^c_\nu.$$

The required kinetic term is then

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} = -\frac{1}{4T(R)}\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}).$$

Showing gauge invariance: This is simple, as long as we use *finite* gauge transformations, and the results:

Theorem: Under a gauge transformation, $D_\mu\phi \mapsto UD_\mu\phi$.

Proof: We have $D_\mu\phi = \partial_\mu\phi + igA_\mu\phi$

$$\mapsto \partial_\mu(U\phi) + ig\left(UA_\mu U^{-1} + \frac{i}{g}\partial_\mu(U)U^{-1}\right)U\phi = UD_\mu\phi. \quad \square$$

Theorem: $F^a_{\mu\nu}F^{a\mu\nu}$ is gauge invariant.

Proof: Note $igF_{\mu\nu}\phi = [D_\mu, D_\nu]\phi$. Under a gauge transformation then, we have

$$igF_{\mu\nu}U\phi = U[D_\mu, D_\nu]\phi,$$

so that $F_{\mu\nu} \mapsto UF_{\mu\nu}U^{-1}$. Thus $F^a_{\mu\nu}F^{a\mu\nu} \propto \mathrm{Tr}(F_{\mu\nu}F^{\mu\nu})$ is gauge invariant (using cyclicity of the trace). $\square$

3.7 The non-Abelian Higgs mechanism

We are now ready to discuss SSB of non-Abelian gauge theories. In general, we work with a Lagrangian of the form

$$\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \frac{1}{2}(D^\mu\phi)^\dagger(D_\mu\phi) - V(\phi),$$

with gauge group G, where ϕ is a *real* multiplet.

Theorem: Assuming:

(i) the generators t^a of the Lie algebra of G obey $\tilde{\phi}^\dagger(t^a\phi) = (t^a\tilde{\phi})^\dagger\phi$, for all $\tilde{\phi}$, ϕ;

(ii) $V'(\phi)^\dagger t^a\phi = 0$;

we have that this Lagrangian is gauge-invariant.

Proof: The kinetic term is gauge invariant by the above.

The given condition is $\tilde{\phi}^\dagger t^a\phi = \tilde{\phi}^\dagger t^{a\dagger}\phi$ for all $\tilde{\phi}$, ϕ, so $t^a = t^{a\dagger}$. This implies the finite transformation $U = \exp(it^a\theta^a(x))$ is unitary. So under a gauge transformation $(D^\mu\phi)^\dagger D_\mu\phi \mapsto (UD^\mu\phi)^\dagger UD_\mu\phi = (D^\mu\phi)^\dagger D_\mu\phi$, since U is unitary.

For the potential, it's actually best to work infinitesimally. Note $V(\phi) \mapsto V(\phi + it^a\theta^a\phi) = V(\phi) + iV'(\phi)^\dagger(t^a\theta^a\phi) = V(\phi)$, by assumption (ii). $\square$

Let's consider SSB of this Lagrangian. Recall that we can get the gauge bosons to *eat* some of the Goldstone bosons by a clever gauge choice. As before, use:

Definition: Let $\phi_0 \neq 0$ be the vacuum. Then *unitary gauge* is defined by $(t^a\phi_0)^\dagger\phi = 0$.

What is the point of this gauge choice? Note $\tilde{t}^i\phi_0 = 0$ automatically holds for any unbroken generators, so this condition says nothing about those.

For broken generators $\theta^{\tilde{a}}$, we recall some differential geometry. The vectors $t^a\phi_0$ span the *tangent space* to ϕ_0, so $\theta^{\tilde{a}}\phi_0$ span a subspace of the tangent space, corresponding to directions associated with the broken generators.

Therefore, the gauge choice tells us that perturbations to ϕ have no component in the broken generators' direction (compare with the Abelian Higgs, when the gauge choice causes $\theta(x)$ to vanish).

This is formalised in the following Theorem:

Theorem: Let $\phi_0 \neq 0$ minimise $V(\phi)$, and let $\phi = \phi_0 + f$ be a perturbation to ϕ_0. Then in unitary gauge, $(t^a\phi_0)^\dagger f = 0$.

Proof: We automatically have $(t^a\phi_0)^\dagger(\phi_0 + f) = 0$. So just need $(t^a\phi_0)^\dagger\phi_0 = \phi_0^\dagger t^a\phi_0 = 0$.

Recall that ϕ is a *real* multiplet. So under any gauge transformation $\phi \mapsto \phi' = \phi + i\alpha^a(x)t^a\phi$, the field must remain real. So $\phi' - \phi$ is real, so $i\alpha^a(x)t^a$ is real. Since we work the real Lie algebra, $\alpha^a(x) \in \mathbb{R}$, and it^a is real. Hence $(it^a)^* = it^a$, which implies $t^{a*} = -t^a$.

The hermiticity condition $t^{a\dagger} = t^a$ implies $t^{aT} = -t^a$, i.e. the generators are *antisymmetric*. So write $\phi_0 t^a\phi_0 = (\phi_0)_i(\phi_0)_j t^a_{ij} = 0$, and we're done. □

Theorem: Suppose $V(\phi)$ is minimised at $\phi = \phi_0 \neq 0$. In unitary gauge, SSB of the above Lagrangian via the expansion $\phi = \phi_0 + f$ gives a mass to the gauge bosons corresponding to the broken generators, and gives no Goldstone bosons (they are all eaten).

Proof: As usual, expand $\phi = \phi_0 + f$. In unitary gauge, we have $(t^a\phi_0)^\dagger f = 0$ from above. Since this condition is meaningless for unbroken generators, split everything up into broken/unbroken.

Split $t^a = (\tilde{t}^i, \theta^{\tilde{a}})$, where $\tilde{t}^i$ are unbroken and $\theta^{\tilde{a}}$ are broken. Also split the gauge fields:

$$A_\mu = {A'_\mu}^i\tilde{t}^i + \hat{A}^{\tilde{a}}_\mu\theta^{\tilde{a}},$$

so that A'_μ are the 'unbroken gauge fields' and $\hat{A}_\mu$ are the 'broken gauge fields'. The covariant derivative splits as:

$$D_\mu\phi = \partial_\mu f + igA^a_\mu t^a(\phi_0 + f) = D'_\mu f + ig\hat{A}^{\tilde{a}}_\mu\theta^{\tilde{a}}(\phi_0 + f),$$

where $D'_\mu = \partial_\mu + ig{A'_\mu}^i\tilde{t}^i$ (note $\tilde{t}^i\phi_0 = 0$ by definition). Hence $(D_\mu\phi)^\dagger D^\mu\phi =$

$$\left((D'_\mu f)^\dagger - ig(\hat{A}^{\tilde{a}}_\mu)^\dagger(\phi_0 + f)^\dagger\theta^{\tilde{a}\dagger}\right)\left(D'^\mu f + ig\hat{A}^{\tilde{b}\mu}\theta^{\tilde{b}}(\phi_0 + f)\right).$$

Simplifying, and keeping only the quadratic terms, we have

$$(D'_\mu f)^\dagger D'^\mu f + ig(D'_\mu f)^\dagger\hat{A}^\mu\phi_0 - ig\phi_0^\dagger\hat{A}^\dagger_\mu D'^\mu f + g^2\phi_0^\dagger\hat{A}^\dagger_\mu\hat{A}^\mu\phi_0.$$

The middle two terms combine to zero, which we show as follows:

$$\begin{aligned}&(D'_\mu f)^\dagger\hat{A}^\mu\phi_0 - \phi_0^\dagger\hat{A}^\dagger_\mu D'^\mu f\\ &= (\partial_\mu f^\dagger - igf^\dagger A'_\mu)\hat{A}^\mu\phi_0 - \phi_0^\dagger\hat{A}_\mu(\partial^\mu f + igA^\mu f)\\ &= \partial_\mu f^\dagger\hat{A}^\mu\phi_0 - \phi_0^\dagger\hat{A}_\mu\partial^\mu f.\end{aligned}$$

Hmm we still need something to save us. But we haven't used the gauge condition! Recall $\phi_0^\dagger\theta^{\tilde{a}}f = 0$, so $\phi_0^\dagger\hat{A}_\mu f = 0$, by multiplying through by $\hat{A}^{\tilde{a}}_\mu(x)$. Taking the derivative, we have $\phi_0^\dagger\partial_\mu\hat{A}^\mu f = -\phi_0^\dagger\hat{A}^\mu\partial_\mu f$.

But recall that the x dependence is entirely in the coefficients of the generators, i.e. $\hat{A}^{\tilde{a}}_\mu(x)$. So we've shown that

$$\phi_0^\dagger\hat{A}^\mu\partial_\mu f = -(\partial_\mu\hat{A}^{\tilde{a}\mu})\phi_0^\dagger\theta^{\tilde{a}}f = 0,$$

by the gauge condition. Similarly, the other term is zero. Thus we're left with:

$$(D'_\mu f)^\dagger D'^\mu f + g^2\phi_0^\dagger\hat{A}^\dagger_\mu\hat{A}^\mu\phi_0.$$

We also expand $V(\phi)$ to get $V(\phi_0)$ (a constant, which we forget about) and $\frac{1}{2}f^\dagger\mathcal{M}f$, where $\mathcal{M}$ is the mass matrix.

Hence the kinetic part of the Lagrangian is:

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \frac{1}{2}(D'_\mu f)^\dagger D'^\mu f + \frac{1}{2}g^2\phi_0^\dagger\hat{A}_\mu\hat{A}^\mu\phi_0 - \frac{1}{2}f^\dagger\mathcal{M}f.$$

In principle, components of f could acquire mass, and components of $\hat{A}^\mu$ could acquire mass. It depends on the matrices $\mathcal{M}$ and $\hat{\mathcal{M}}_{\tilde{a}\tilde{b}} = g^2(\theta^{\tilde{a}}\phi_0)^\dagger\theta^{\tilde{b}}\phi_0$.

- $\hat{A}^\mu$ MASSES: All of the broken gauge fields acquire a mass from eating the Goldstone bosons. This is because the matrix $g^2(\theta^{\tilde{a}}\phi_0)^\dagger\theta^{\tilde{b}}\phi_0$ has no zero eigenvalues; if it did, say $v^{\tilde{b}}$, they would have to obey $v^{\tilde{b}}\theta^{\tilde{b}}\phi_0 = 0$. But then $v^{\tilde{b}}\theta^{\tilde{b}}$ would be an unbroken generator, contradiction.

- f MASSES: Recall that near any ϕ, we can expand:

  $$V(\phi + \delta\phi) - V(\phi) = i\alpha^a(t^a\phi)_r\frac{\partial V}{\partial\phi_r} + O(\alpha^2).$$

  Differentiate to obtain:

  $$\frac{\partial^2 V}{\partial\phi_s\partial\phi_r}(t^a\phi)_r + \frac{\partial V}{\partial\phi_r}(t^a)_{rs}.$$

  Evaluating at $\phi = \phi_0$ gives $\mathcal{M}_{sr}(t^a\phi_0)_r = 0$. This shows that each broken generator, i.e. $\theta^{\tilde{a}}\phi_0 \neq 0$, gives a zero eigenvector of $\mathcal{M}_{sr}$. But the unitary gauge condition implies that f is orthogonal to such evectors, and thus there are no necessary Goldstone bosons in this case. □

The final technique of the proof above also allows us to prove the following useful result:

Theorem: The mass eigenstates form multiplets of the gauge group H (i.e. they are eigenvectors of the generators of the Lie algebra of H).

Proof: Recall we have:

$$V(\phi+\delta\phi)-V(\phi)=i\alpha^a(t^a\phi)_r\frac{\partial V}{\partial\phi_r}+O(\alpha^2)$$

and so differentiating, we get:

$$\frac{\partial^2 V}{\partial\phi_s\partial\phi_r}(t^a\phi)_r+\frac{\partial V}{\partial\phi_r}(t^a)_{rs}.$$

Differentiating a second time and evaluating at $\phi=\phi_0$, we obtain:

$$\left.\frac{\partial^3 V}{\partial\phi_t\partial\phi_s\partial\phi_r}\right|_{\phi=\phi_0}(t^a\phi_0)_r+\mathcal{M}_{sr}(t^a)_{rt}+\mathcal{M}_{tr}(t^a)_{rs}=0.$$

For unbroken generators, we have $\tilde{t}^a\phi_0=0$, so the first term vanishes. We're left with:

$$\mathcal{M}\tilde{t}^a+(\tilde{t}^a)^T\mathcal{M}=0.$$

(Notice $\mathcal{M}^T=\mathcal{M}$ by definition). Recall that $\tilde{t}^a$ were *anti-symmetric* and $\{i\tilde{t}^a\}$ generates H. Therefore:

$$[\mathcal{M},it^a]=0.$$

It follows that the mass matrix and the generators of H are simultaneously diagonalisable; thus the result follows. □

3.8 Examples of the Higgs mechanism

Example 1: Consider an $SU(2)$ gauge theory coupled to a complex 2-component scalar field ϕ via the Lagrangian:

$$\mathcal{L}=-\frac{1}{4}\mathbf{F}^{\mu\nu}\cdot\mathbf{F}_{\mu\nu}+(D^\mu\phi)^\dagger D_\mu\phi-\frac{1}{2}\lambda(\phi^\dagger\phi-\frac{1}{2}v^2)^2.$$

Here, the generators of $SU(2)$ are $i\boldsymbol{\tau}=\frac{1}{2}i\boldsymbol{\sigma}$, where $\boldsymbol{\sigma}$ are the Pauli matrices. The structure constants are quickly computed to be $[\tau^a,\tau^b]=i\epsilon^{abc}\tau^c\Rightarrow f^{abc}=\epsilon^{abc}$.

Notice that in this example we've used vector notation. We can discern its meaning as follows; recall that

$$F^a_{\mu\nu}=\partial_\mu A^a_\nu-\partial_\nu A^a_\mu-g\epsilon^{abc}A^b_\mu A^c_\nu,$$

so in vector notation, we have

$$\mathbf{F}_{\mu\nu}=\partial_\mu\mathbf{A}_\nu-\partial_\nu\mathbf{A}_\mu-g\mathbf{A}_\mu\times\mathbf{A}_\nu.$$

Vacuum and unitary gauge: It's clear that for $v^2>0$, we get spontaneous symmetry breaking, with the vacuum manifold $\phi^\dagger\phi=\frac{1}{2}v^2$. As usual, we can WLOG pick a vacuum, say $\phi_0=\frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix}$, and as usual it will be desirable to impose unitary gauge.

Let's see another example of how to construct it explicitly. In general, a perturbation to ϕ_0 is of the form:

$$\begin{aligned}\phi&=\frac{1}{\sqrt{2}}\begin{pmatrix}\theta^1+i\theta^2\\v+h+i\theta^3\end{pmatrix}\\&=\frac{1}{\sqrt{2}}\begin{pmatrix}0\\v+h\end{pmatrix}+\frac{1}{\sqrt{2}}\frac{1}{(v+h)}\begin{pmatrix}(v+h)(\theta^1+i\theta^2)\\(v+h)i\theta^3\end{pmatrix}\\&=\frac{1}{\sqrt{2}}\begin{pmatrix}0\\v+h\end{pmatrix}+\frac{1}{\sqrt{2}}\frac{1}{(v+h)}\begin{pmatrix}-i\theta^3&\theta^1+i\theta^2\\-\theta^1+i\theta^2&i\theta^3\end{pmatrix}\begin{pmatrix}0\\v+h\end{pmatrix}.\end{aligned}$$

where θ^1, θ^2, θ^3 and h are real fields. Notice that

$$\begin{pmatrix}-i\theta_3&\theta_1+i\theta_2\\-\theta_1+i\theta_2&i\theta_3\end{pmatrix}=2i\theta^2\tau^1+2i\theta^1\tau^2-2i\theta^3\tau^3.$$

and hence we may write

$$\begin{aligned}\phi&=\frac{1}{\sqrt{2}}\begin{pmatrix}\theta^1+i\theta^2\\v+h+i\theta^3\end{pmatrix}\\&=\frac{1}{\sqrt{2}}\exp\left(\frac{1}{v+h}(2i\theta^2\tau^1+2i\theta^1\tau^2-2i\theta^3\tau^3)\right)\begin{pmatrix}0\\v+h\end{pmatrix}.\end{aligned}$$

Since the exponential is an element of $SU(2)$, we can safely gauge transform ϕ such that the perturbation is (WLOG) of the form:

$$\phi=\frac{1}{\sqrt{2}}\begin{pmatrix}0\\v+h\end{pmatrix},$$

which is unitary gauge, as required.

Unbroken symmetry: Notice that the vacuum manifold can be written $a^2+b^2+c^2+d^2=\frac{1}{2}v^2$, where $\phi=(a+ib,c+id)^T$, and hence it follows the vacuum manifold is of dimension 3. Now by Goldstone's Theorem, we have $\dim(H)=\dim(SU(2))-\dim(\Phi_0)=3-3=0$, so the symmetry is completely destroyed.

Particle masses: Expanding ϕ near the vacuum as above, and keeping only terms that are quadratic in h or quadratic in A_μ (and ignoring any couplings between fields), we see that h acquires a mass $\sqrt{\lambda v^2}$ and the A^a_μ fields all acquire a mass $\frac{1}{2}gv$.

Example 2: Now consider an $SU(2)$ gauge theory

$$\mathcal{L}=-\frac{1}{4}\mathbf{F}^{\mu\nu}\cdot\mathbf{F}_{\mu\nu}+\frac{1}{2}(D^\mu\boldsymbol{\phi})\cdot(D_\mu\boldsymbol{\phi})-\frac{1}{8}\lambda(\boldsymbol{\phi}^2-v^2)^2.$$

where $\boldsymbol{\phi}$ is a real triplet (i.e. transforms in the adjoint rep, so that the generators are $(t^a)_{jk}=-i\epsilon_{ajk}$). Notice that the covariant derivative of this theory is:

$$D_\mu\boldsymbol{\phi}=\partial_\mu\boldsymbol{\phi}-e\mathbf{A}_\mu\times\boldsymbol{\phi}.$$

When $v^2>0$, the vacuum manifold is $\boldsymbol{\phi}^2=v^2$, this is clearly the 2-sphere, S^2. Now use the isomorphism $S^2\cong SU(2)/U(1)$ to conclude that the unbroken symmetry group is $U(1)$.

WLOG choose the vacuum $\phi_0 = (0,0,v)^T$ and choose a perturbation $\phi = (0,0,v+h)^T$ by moving to unitary gauge (here, we can remove the two degrees of freedom in the first two components because there are $\dim(SU(2)) - \dim(U(1)) = 3-1 = 2$ broken generators).

Inserting into the Lagrangian, we find that the mass of the h field becomes $\sqrt{\lambda v^2}$, and the masses of A^1_μ, A^2_μ becomes ev. However, A^3_μ remains massless (it was protected by the $U(1)$ invariant subgroup).

We also see there are h^4, h^3 self-interactions, and gauge boson-h interactions via $(A^a_\mu)^2 h$ and $(A^a_\mu)^2 h^2$.

Example 3: Finally, consider an $SU(2)$ gauge theory coupled to a complex triplet field $\boldsymbol{\phi}$ with Lagrangian:

$$\mathcal{L} = -\frac{1}{4}\mathbf{F}^{\mu\nu}\cdot\mathbf{F}_{\mu\nu} + (D^\mu\boldsymbol{\phi})\cdot(D_\mu\boldsymbol{\phi}) + \frac{1}{2}g^2(\boldsymbol{\phi}^*\times\boldsymbol{\phi})^2.$$

Notice that if we write $\boldsymbol{\phi} = \boldsymbol{\theta} + i\boldsymbol{\chi}$, then the potential term becomes:

$$V(\boldsymbol{\phi}) = -\frac{1}{2}g^2(\boldsymbol{\phi}^*\times\boldsymbol{\phi})^2 = 2g^2(\boldsymbol{\theta}\times\boldsymbol{\chi})^2 \geq 0.$$

So we can minimise the potential by choosing $\boldsymbol{\phi}_0 = v\mathbf{e}_3/\sqrt{2}$ to be our ground state.

To find the invariant subgroup, recall that we're in the adjoint rep, so we want to relate $SU(2)$ to 3×3 matrices. Recall $SU(2)$ is a double cover of $SO(3)$, so we can immediately do this. Then it's clear that matrices of the form:

$$\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}$$

preserve the vacuum, with $A \in SO(2)$. So the unbroken symmetry group is $SO(2)$.

It follows that by Goldstone's Theorem that there are $\dim(SU(2)) - \dim(SO(2)) = 2$ massless modes; these can be eaten by the gauge bosons by imposing unitary gauge. Since there are two massless modes, we can remove two degrees of freedom from the perturbation in unitary gauge, e.g. choose the conditions:

$$\mathrm{Re}(v^*\boldsymbol{\phi}\cdot\mathbf{e}_1) = \mathrm{Re}(v^*\boldsymbol{\phi}\cdot\mathbf{e}_2) = 0.$$

In this gauge choice, perturbations are of the form $\phi = \frac{1}{\sqrt{2}}(f_1 - i\frac{v_1}{v_2}f_1, f_3 - \frac{v_1}{v_2}f_3, v + g_1 + ig_2)^T$. Substituting into the Lagrangian then gives the mass terms in the usual way.

However... the above shows that the masses will *depend on* $|v|$, which means that changing $|v|$ will change the theory itself. In particular, theories with different v are inequivalent - the choice of vacuum mattered here!

This is because for this theory, the gauge transformations do not act transitively on the vacuum manifold.

4 The electroweak theory

4.1 Gauge boson-ϕ coupling

Definition: The *electroweak theory* is an $SU(2)\times U(1)$ gauge theory, together with a complex scalar field ϕ, called the *Higgs field* in the fundamental representation of $SU(2)$. Later, we'll also include both *quarks* and *leptons*.

The gauge boson-Higgs coupling is of the standard non-Abelian Higgs form:

$$\mathcal{L}_{\text{gauge},\phi} = -\frac{1}{2}\mathrm{Tr}(F^W_{\mu\nu}F^{W\mu\nu}) - \frac{1}{4}F^B_{\mu\nu}F^{B\mu\nu} + (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi),$$

where the covariant derivative is given by

$$D_\mu\phi = \partial_\mu\phi + igW^a_\mu\tau^a\phi + \frac{1}{2}ig'B_\mu\phi.$$

Here, $i\tau^a = \frac{1}{2}i\sigma^a$ are the generators of $SU(2)$, and W^a_μ, B_μ are the gauge boson fields. The factor of $\frac{1}{2}$ next to B_μ means that we say the Higgs has *hypercharge* $\frac{1}{2}$. From the covariant derivative, we can read off the transformation law for the Higgs under $SU(2)\times U(1)$ transformations as:

$$\phi(x) \mapsto \exp(i\alpha^a(x)t^a(x))\exp\left(\frac{1}{2}i\beta(x)\right)\phi(x)$$

The potential of the theory is

$$V(\phi) = \mu^2|\phi|^2 + \lambda|\phi|^4,$$

where $\mu^2 < 0$, so that SSB occurs.

The vacuum state of ϕ obeys $|\phi|^2 = -\mu^2/2\lambda = v^2/2$, so WLOG choose the vacuum

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix},$$

The unbroken subgroup H is $U(1)$, since the only transformations preserving ϕ_0 are those with $\alpha^1(x) = \alpha^2(x) = 0$ and $\alpha^3(x) = \beta(x)$:

$$\phi_0 \mapsto \exp\left(\frac{1}{2}i\beta(x)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right)\exp\left(\frac{1}{2}i\beta(x)\right)\phi_0$$
$$= \frac{1}{\sqrt{2}}\begin{pmatrix} e^{i\beta} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ v \end{pmatrix} = \phi_0.$$

As we will see, this is the familiar gauge group $U(1)$ of *electromagnetism*, mediated by the photon.

We now wish to perturb around this vacuum. Let

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix},$$

where $h(x)$ is a real scalar field. We're allowed to do this because we've imposed *unitary gauge*.

As in the Abelian Higgs mechanism (where we used a complex ϕ), this amounts to removing the $U(1)$ phase, and any matrix $U \in SU(2)$ which would rotate our field.

After spontaneous symmetry breaking, it turns out that some combinations of the gauge bosons W_μ^a, B_μ play an important role.

Definition: Define the following linear combinations of the gauge bosons:

$$Z^0 = \frac{1}{\sqrt{g^2+g'^2}}(gW^3 - g'B), \quad W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp iW^2),$$
$$A = \frac{1}{\sqrt{g^2+g'^2}}(g'W^3 + gB).$$

These are called the Z-*boson*, the $W^\pm$-*bosons* and the *photon* respectively. It's convenient also to define the *Weinberg angle* θ_W, given by

$$\cos(\theta_W) = \frac{g}{\sqrt{g^2+g'^2}}, \quad \sin(\theta_W) = \frac{g'}{\sqrt{g^2+g'^2}}.$$

with which we can rephrase the Z and photon definitions as:

$$\begin{pmatrix} Z^0 \\ A \end{pmatrix} = \begin{pmatrix} \cos(\theta_W) & -\sin(\theta_W) \\ \sin(\theta_W) & \cos(\theta_W) \end{pmatrix} \begin{pmatrix} W^3 \\ B \end{pmatrix}.$$

Definition: Define the field-strength tensors:

$$F_{\mu\nu}^A = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad F_{\mu\nu}^Z = \partial_\mu Z_\nu^0 - \partial_\nu Z_\mu^0,$$
$$F_{\mu\nu}^{W^\pm} = \frac{1}{\sqrt{2}}\left(F_{1\mu\nu}^W - iF_{2\mu\nu}^W\right).$$

Theorem: In unitary gauge, expanding $\phi = \phi_0 + (0, h/\sqrt{2})^T$, the Lagrangian becomes: $\mathcal{L}_{\text{gauge},\phi} =$

$$\underbrace{-\frac{1}{2}F_{\mu\nu}^{W^\pm\dagger}F^{W^\pm\mu\nu} - \frac{1}{4}F_{\mu\nu}^A F^{A\mu\nu} - \frac{1}{4}F_{\mu\nu}^Z F^{Z\mu\nu} + \frac{1}{2}|\partial h|^2}_{\text{kinetic terms}}$$
$$\underbrace{-igW^{+\mu}W^{-\nu}\left(\sin(\theta_W)F_{\mu\nu}^A + \cos(\theta_W)F_{\mu\nu}^Z\right)}_{W\text{ boson-photon and }W\text{ boson-}Z\text{ boson interactions}}$$
$$\underbrace{+\frac{1}{2}g^2\left(W^{+2}W^{-2} - (W^{+\mu}W_\mu^-)^2\right)}_{W\text{ boson self-interactions}}$$
$$\underbrace{+\frac{1}{4}g^2(v+h)^2\left(W_\mu^+W^{-\mu} + \frac{1}{2}\sec^2(\theta_W)Z_\mu^0 Z^{0\mu}\right)}_{\text{gauge boson masses and gauge boson-Higgs interactions}}$$
$$\underbrace{+\frac{1}{2}\mu^2(v+h)^2 + \frac{1}{4}\lambda(v+h)^4}_{\text{Higgs mass and Higgs self-interactions}}.$$

Proof: Expanding the covariant derivative of the Higgs, we have: $D_\mu\phi = \partial_\mu\phi + igW_\mu\phi + \frac{1}{2}ig'B_\mu\phi$

$$= \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{1}{2}ig(W_\mu^1 - iW_\mu^2)(v+h) \\ \partial_\mu h + \frac{1}{2}i(g'B_\mu - gW_\mu^3)(v+h) \end{pmatrix}$$

Therefore, $(D_\mu\phi)^\dagger(D^\mu\phi)$ has expansion:

$$\frac{1}{8}g^2(v+h)^2|W^1 - iW^2|^2 + \frac{1}{2}|\partial h|^2 + \frac{1}{8}(v+h)^2|g'B - gW^3|^2.$$

Tidying, we get the gauge boson-Higgs interactions, and the Higgs kinetic term.

Now consider expanding $V(\phi)$; this clearly gives the Higgs self-interactions and mass term as in the final result.

Finally, we deal with the kinetic terms $F_{a\mu\nu}^W F^{Wa\mu\nu}$ and $F_{\mu\nu}^B F^{B\mu\nu}$. Notice that

$$F_{a\mu\nu}^W F^{Wa\mu\nu} = F_{1\mu\nu}^W F^{W1\mu\nu} + F_{2\mu\nu}^W F^{W2\mu\nu} + F_{3\mu\nu}^W F^{W3\mu\nu}$$
$$= 2F_{\mu\nu}^{W^\pm\dagger}F^{W^\pm\mu\nu} + F_{3\mu\nu}^W F^{W3\mu\nu},$$

so we immediately get the $W^\pm$ kinetic term. Now, using the structure constants of the fundamental rep of $SU(2)$ given by $f^{abc} = \epsilon^{abc}$, we have

$$F_{3\mu\nu}^W = \partial_\mu W_\nu^3 - \partial_\nu W_\mu^3 - g\epsilon^{3bc}W_\mu^b W_\nu^c$$

$$= \partial_\mu(\cos(\theta_W)Z_\nu^0 + \sin(\theta_W)A_\nu) - \partial_\nu(\cos(\theta_W)Z_\mu^0 + \sin(\theta_W)A_\mu)$$
$$-g\left(W_\mu^1W_\nu^2 - W_\mu^2W_\nu^1\right)$$

$$= \cos(\theta_W)F_{\mu\nu}^Z + \sin(\theta_W)F_{\mu\nu}^A + ig\left(W_\mu^+W_\nu^- - W_\mu^-W_\nu^+\right).$$

Squaring this, we have

$$F_{3\mu\nu}^W F^{W3\mu\nu} = \cos^2(\theta_W)F_{\mu\nu}^Z F^{Z\mu\nu} + \sin^2(\theta_W)F_{\mu\nu}^A F^{A\mu\nu}$$
$$-g^2\left(W_\mu^+W_\nu^- - W_\mu^-W_\nu^+\right)^2 + 2\sin(\theta_W)\cos(\theta_W)F_{\mu\nu}^Z F^{A\mu\nu} +$$
$$2ig\left(W_\mu^+W_\nu^- - W_\mu^-W_\nu^+\right)\cos(\theta_W)F^{Z\mu\nu} +$$
$$2ig\left(W_\mu^+W_\nu^- - W_\mu^-W_\nu^+\right)\sin(\theta_W)F^{A\mu\nu}.$$

By antisymmetry of $F^{A\mu\nu}$ and $F^{Z\mu\nu}$, the last two terms can be simplified completely to:

$$4igW_\mu^+W_\nu^-\left(\cos(\theta_W)F^{Z\mu\nu} + \sin(\theta_W)F^{A\mu\nu}\right).$$

giving the expected W boson-photon and W boson-Z boson interactions. Simplifying the W^+, W^- squared term immediately gives the W-boson self-interactions.

Finally, we also need to consider contributions from $F_{\mu\nu}^B$. We have

$$F_{\mu\nu}^B = \partial_\mu B_\nu - \partial_\nu B_\mu = -\sin(\theta_W)F_{\mu\nu}^Z + \cos(\theta_W)F_{\mu\nu}^A.$$

Squaring gives: $F_{\mu\nu}^B F^{B\mu\nu} =$

$$\sin^2(\theta_W)F_{\mu\nu}^Z F^{Z\mu\nu} + \cos^2(\theta_W)F_{\mu\nu}^A F^{A\mu\nu} - 2\sin(\theta_W)\cos(\theta_W)F_{\mu\nu}^Z F^{A\mu\nu}.$$

This combines exactly with $F_{3\mu\nu}^W F^{W3\mu\nu}$ to give the required Lagrangian. □

4.2 Analysis of $\mathcal{L}_{\text{gauge},\phi}$

Let's analyse the parts of the Lagrangian above:

- MASSES: We find the photon is massless, the W boson has mass $m_W = \frac{1}{2}gv$ (note W complex), the Z boson has mass $m_Z = \frac{1}{2}gv\sec(\theta_W)$ (note Z real), and the Higgs has mass $m_\phi = \sqrt{3\lambda v^2 + \mu^2} = \sqrt{2\lambda v^2}$.

- Relation of Z and W masses: We see that the Z and W masses are related by the formula:
$$m_W = m_Z \cos(\theta_W).$$
- Interactions: Everything interacts with everything else, *except* the Z boson and Higgs do not interact with the photon. This reflects the fact that the Z boson and Higgs have zero electric charge.

4.3 Gauge boson-lepton coupling

Leptons are a type of fermion. Leptons are again given mass by the Higgs mechanism, so there are two forms of coupling we need to consider:

- Gauge boson-lepton coupling through terms like $\overline{\psi} i \not{D} \psi$. Call the full coupling $\mathcal{L}^{\text{EW}}_{\text{lept}}$.
- Higgs-lepton coupling, through a part of the Lagrangian which we'll call $\mathcal{L}_{\text{lept},\phi}$.

Begin with coupling to the gauge bosons. The action of the covariant derivative D_μ on a fermion is given by
$$D_\mu \psi = (\partial_\mu + igW^a_\mu T^a + ig' Y B_\mu)\psi$$
$$= \Bigg(\partial_\mu + \frac{1}{\sqrt{2}} ig\left(W^+_\mu T^+ + W^-_\mu T^-\right) + \frac{igZ^0_\mu}{\cos(\theta_W)}\left(T^3 - \sin^2(\theta_W)(T^3+Y)\right) + ig\sin(\theta_W) A_\mu (T^3+Y)\Bigg)\psi,$$
where we've expanded in terms of the physical W, Z and photon fields. Note also that iT^a are the generators for the $SU(2)$ rep of ψ, and that Y is the hypercharge of the lepton field ψ. We define $T^\pm = T^1 \pm iT^2$.

We note immediately that:

- $g\sin(\theta_W)$ is the coupling of A_μ to the lepton field ψ. Hence $e = g\sin(\theta_W)$ is the familiar *electric charge*.
- Depending on the rep, $Q = T^3 + Y$ is a matrix which when applied to eigenvectors ψ_λ, returns $\lambda\psi_\lambda$, so that ψ_λ has an actual charge $e\lambda$. Thus Q is called the *charge matrix*.

We now input the experimental observation:

Observation: The two lightest lepton fields in Nature are the *electron field* $e(x)$ and the *electron neutrino field* $\nu_e(x)$. The left-handed fields form an $SU(2)$ doublet:
$$L(x) = \begin{pmatrix} \nu_{e_L}(x) \\ e_L(x) \end{pmatrix},$$
where $e_L(x) = P_L e(x)$ and $\nu_{e_L}(x) = P_L \nu_e(x)$. Thus, the left-handed particles couple to the W boson.

The right-handed neutrino does not couple to any bosons, so we omit further discussion of it. The right-handed electron forms an $SU(2)$ singlet:
$$R(x) = e_R(x),$$
and hence does not couple to the W boson.

We also have from experiment that electrons have charge $-e$ and neutrinos are neutral. Therefore, for the right-handed electron field, we have $-1 = Q = T^3 + Y = 0 + Y = Y$, and hence $e_R(x)$ has hypercharge -1.

For the left-handed fields, we have
$$\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} = Q = T^3 + Y = \begin{pmatrix} \frac{1}{2}+Y & 0 \\ 0 & -\frac{1}{2}+Y \end{pmatrix}.$$
Hence $\nu_{e_L}(x)$ and $e_L(x)$ have hypercharge $-\frac{1}{2}$.

Using the above experimental observations, we have
$$\mathcal{L}^{\text{EW}}_{\text{lept}} = \overline{L} i\not{D} L + \overline{R} i\not{D} R.$$
Note the covariant derivatives are different! The first has $iT^a = \frac{1}{2}i\sigma^a$, while the second has $T^a \equiv 0$.

From this Lagrangian, we can identify the interactions between gauge bosons and leptons in the electroweak theory. Writing out the Lagrangian in full, we have
$$\mathcal{L}^{\text{EW}}_{\text{lept}} = \underbrace{\overline{\nu}_{e_L} i\not{\partial}\nu_{e_L} + \overline{e}_L i\not{\partial} e_L + \overline{e}_R i\not{\partial} e_R}_{\text{kinetic terms}} + \underbrace{e\overline{e}_L \not{A} e_L + e\overline{e}_R \not{A} e_R}_{\text{electron-photon interactions}}$$
$$\underbrace{-\frac{g}{\sqrt{2}}\left(\overline{\nu}_{e_L} \not{W}^+ e_L + \overline{e}_L \not{W}^- \nu_{e_L}\right)}_{\text{electron-neutrino-}W\text{ boson interactions}}$$
$$\underbrace{-\frac{g}{2\cos(\theta_W)}\Big(\overline{\nu}_{e_L} \not{Z}^0 \nu_{e_L} + (2\sin^2(\theta_W)-1)\overline{e}_L \not{Z}^0 e_L + 2\sin^2(\theta_W)\overline{e}_R \not{Z}^0 e_R\Big)}_{\text{electron-neutrino-}Z\text{ boson interactions}}.$$
We can tidy this up a little by noticing that
$$\overline{\psi}\gamma^\mu\psi = (\overline{\psi}_L + \overline{\psi}_R)\gamma^\mu(\psi_L + \psi_R)$$
$$= \overline{\psi}_L\gamma^\mu\psi_L + \overline{\psi}_R\gamma^\mu\psi_R + \overline{\psi}_L\gamma^\mu\psi_R + \overline{\psi}_R\gamma^\mu\psi_L.$$
Now note that $\overline{\psi}_L\gamma^\mu\psi_R = \overline{\psi}P_R\gamma^\mu P_R\psi_R = 0$ (since P_R changes into a P_L when it passes through the γ^μ). Similarly $\overline{\psi}_R\gamma^\mu\psi_L = 0$.

So the whole Lagrangian can be simplified to:
$$\mathcal{L}^{\text{EW}}_{\text{lept}} = \underbrace{\overline{\nu}_{e_L} i\not{\partial}\nu_{e_L} + \overline{e} i\not{\partial} e}_{\text{kinetic terms}} + \underbrace{e\overline{e}\not{A}e}_{e\text{-photon ints}} \underbrace{-\frac{g}{\sqrt{2}}\left(\overline{\nu}_{e_L}\not{W}^+ e_L + \overline{e}_L \not{W}^- \nu_{e_L}\right)}_{\text{electron-neutrino-}W\text{ boson interactions}}$$
$$\underbrace{-\frac{g}{2\cos(\theta_W)}\left(\overline{\nu}_{e_L}\not{Z}^0\nu_{e_L} - \frac{1}{2}\overline{e}\not{Z}^0(1-\gamma^5 - 4\sin^2(\theta_W))e\right)}_{\text{electron-neutrino-}Z\text{ boson interactions}}.$$

To tidy this up even further, we make the following definitions:

Definition: The *leptonic electromagnetic current* is:

$$J^\mu_{\text{lept,EM}} = -\overline{e}\gamma^\mu e.$$

The *leptonic charged weak current* is:

$$J^\mu_{\text{lept}} = \overline{\nu}_{e_L}\gamma^\mu(1-\gamma^5)e.$$

The *leptonic neutral weak current* is:

$$J^\mu_{\text{lept},n} = \frac{1}{2}\left(\overline{\nu}_{e_L}\gamma^\mu(1-\gamma^5)\nu_e - \overline{e}\gamma^\mu(1-\gamma^5 - 4\sin^2(\theta_W))e\right).$$

With this notation (omitting the lept's), the Lagrangian reduces to:

$$\mathcal{L}^{\text{EW}}_{\text{lept}} = \overline{\nu}_{e_L} i\partial\!\!\!/ \nu_{e_L} + \overline{e} i D\!\!\!\!/\, e - eA_\mu J^\mu_{\text{EM}} - \frac{g}{2\sqrt{2}}\left(W^+_\mu J^\mu + W^-_\mu J^{\mu\dagger}\right) - \frac{g}{2\cos(\theta_W)} J^\mu_n Z^0_\mu.$$

4.4 Higgs boson-lepton coupling

A useful result for this section is:

Theorem: Let ψ be a spinor. Then $\overline{\psi}\psi = \overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L$.

Proof: We have

$$\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L = \overline{\psi}P_RP_R\psi + \overline{\psi}P_LP_L\psi = \overline{\psi}(P_R+P_L)\psi. \square$$

We now want to give the leptons masses. We can't do this directly, because fermion mass terms such as

$$m_e\overline{e}e = m_e\left(\overline{e}_Le_R + \overline{e}_Re_L\right)$$

explicitly break gauge invariance. Thus to give the leptons masses, we *must* couple them to the Higgs boson and use SSB.

Definition: The Higgs boson-lepton coupling before SSB is

$$\mathcal{L}_{\text{lept},\phi} = -\sqrt{2}\lambda_e\left(\overline{L}\phi R + \overline{R}\phi^\dagger L\right),$$

where λ_e is called the *Yukawa coupling*.

Theorem: $\mathcal{L}_{\text{lept},\phi}$ is gauge invariant.

Proof: Under a gauge transformation, $L \mapsto e^{-\frac{1}{2}i\alpha}UL$, $\phi \mapsto e^{\frac{1}{2}i\alpha}U\phi$ and $R \mapsto e^{-i\alpha}R$. It's then straightforward to substitute these formulae in and check. $\square$

Theorem: In unitary gauge, expanding the Higgs field as $\phi = (0,(v+h)/\sqrt{2})^T$ gives: $\mathcal{L}_{\text{lept},\phi} = -m_e\overline{e}e - \lambda_e h\overline{e}e$, where $m_e = \lambda_e v$ is the electron's acquired mass.

Proof: Very simple calculation. $\square$

4.5 Generations

The electron has some heavier cousins, called the *muon*, μ, and *tau*, τ. The electron neutrino also has heavier muon neutrino and tau neutrino counterparts. These are included in the Standard Model in exactly the sam way as the electron and electron neutrino, now with three L and R fields:

$$L^1 = \begin{pmatrix}\nu_{e_L}\\ e_L\end{pmatrix}, \quad L^2 = \begin{pmatrix}\nu_{\mu_L}\\ \mu_L\end{pmatrix}, \quad L^3 = \begin{pmatrix}\nu_{\tau_L}\\ \tau_L\end{pmatrix},$$

and

$$R^1 = e_R, \quad R^2 = \mu_R, \quad R^3 = \tau_R.$$

Each of the three sets is referred to as a *generation*.

We can generalise the couplings to the gauge bosons and the Higgs as follows.

$\mathcal{L}^{\text{EW}}_{\text{lept}}$ just gets three copies of itself, with $e \mapsto \mu \mapsto \tau$ for each generation, so remains simple.

$\mathcal{L}_{\text{lept},\phi}$ becomes more complicated. It's possible we get something of the form

$$\mathcal{L}_{\text{lept},\phi} = -\sqrt{2}\left(\lambda^{ij}\overline{L}^i\phi R^j + (\lambda^\dagger)^{ij}\overline{R}^i\phi^\dagger L^j\right),$$

where λ^{ij} is a matrix responsible for mixing generations.

Theorem: λ may be diagonalised, so that there is no mixing between lepton generations.

Proof: Note $\lambda\lambda^\dagger$ is Hermitian, so there exists unitary K such that $\lambda\lambda^\dagger = K\Lambda^2K^\dagger$, where Λ is diagonal with real entries (note we can write Λ^2, since if v is an evector of $\lambda\lambda^\dagger$, we have $0 \leq ||\lambda^\dagger v||^2 = v\lambda^\dagger\lambda v = ||v||^2\alpha$, where α is the evalue, i.e. all evalues are non-negative).

Let $S = \lambda^\dagger K\Lambda^{-1}$. Then S is unitary, since

$$S^\dagger S = \Lambda^{-1}K^\dagger\lambda\lambda^\dagger K\Lambda^{-1} = \Lambda^{-1}\Lambda^2\Lambda^{-1} = I.$$

Therefore, $\lambda = K\Lambda S^\dagger$.

Let $L^i \mapsto K^{ij}L^j$ and $R^i \mapsto S^{ij}R^j$. This diagonalises $\mathcal{L}_{\text{lept},\phi}$, but leaves

$$\mathcal{L}^{\text{EW}}_{\text{lept}} = \sum_{i=1}^{3}\left(\overline{L}^i iD\!\!\!\!/\, L^i + \overline{R}^i iD\!\!\!\!/\, R^i\right)$$

invariant. So we can assume WLOG no mixing between lepton generations (by taking the lepton generations to be the vectors $K^{ij}L^j$, $S^{ij}R^j$, i.e. after diagonalisation). $\square$

4.6 Coupling to quarks

Observation: *Quark* fields are fermion fields. There are six *flavours* of quark in Nature: up, down, charm, strange, top and bottom. Their behaviour depends on whether they are left-handed or right-handed:

- *Right-handed quarks* are in $SU(2)$ singlets:
$$u_R^i = (u_R, c_R, t_R), \quad d_R^i = (d_R, s_R, b_R).$$
Here, i labels the generation of quark. Experiment says that up, charm and top quarks have charge $+2/3$, so they have hypercharge $+2/3$ too. Down, strange and bottom quarks have charge $-1/3$, so they have hypercharge $-1/3$ also.
- *Left-handed quarks* are in $SU(2)$ doublets:
$$Q_L^i = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix} = \left(\begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \right).$$
To get the charges right, we need a charge matrix:
$$\begin{pmatrix} 2/3 & 0 \\ 0 & -1/3 \end{pmatrix} = Q = T^3 + Y = \begin{pmatrix} \frac{1}{2} + Y & 0 \\ 0 & -\frac{1}{2} + Y \end{pmatrix}.$$
So left-handed quarks have hypercharge $Y = 1/6$.

Quarks have the usual coupling to the gauge bosons, via
$$\mathcal{L}_{\text{quark}}^{\text{EW}} = \sum_{i=1}^{3} \left(\overline{Q}_L^i i\not{D} Q_L^i + \overline{u}_R^i i\not{D} u_R^i + \overline{d}_R^i i\not{D} d_R^i \right).$$
Just as for leptons, we can expand the covariant derivatives to find all the interactions:

Theorem: The above Lagrangian can be written as:
$$\begin{aligned}\mathcal{L}_{\text{quark}}^{\text{EW}} = \sum_{i=1}^{3} \Bigg(& \underbrace{i\overline{u}^i \not{\partial} u^i + i\overline{d}^i \not{\partial} d^i}_{\text{kinetic terms}} - \underbrace{\frac{2}{3} e\overline{u}^i \not{A} u^i + \frac{1}{3} e\overline{d}^i \not{A} d^i}_{\text{quark-photon couplings}} \\ & \underbrace{- \frac{g}{\sqrt{2}} \left(\overline{u}_L^i \not{W}^+ d_L^i + \overline{d}_L^i \not{W}^- u_L^i \right)}_{\text{quark-}W\text{ boson couplings}} \\ & \underbrace{- \frac{g}{2\cos(\theta_W)} \left(\frac{1}{2} \overline{u}^i \not{Z}^0 \left(1 - \gamma^5 - \frac{8}{3} \sin^2(\theta_W) \right) u^i + \frac{1}{2} \overline{d}^i \not{Z}^0 \left(\frac{4}{3} \sin^2(\theta_W) - 1 + \gamma^5 \right) d^i \right)}_{\text{quark-}Z\text{ boson coupling}}\Bigg).\end{aligned}$$

Proof: By a similar calculation to $\mathcal{L}_{\text{lept}}^{\text{EW}}$. $\square$

We can tidy this up by defining currents.

Definition: For quarks, the *hadronic electromagnetic current* is defined by
$$J_{\text{had,EM}}^{\mu} = \frac{2}{3} \overline{u}^i \gamma^\mu u^i - \frac{1}{3} \overline{d}^i \gamma^\mu d^i.$$
The *hadronic charged weak current* is
$$J_{\text{had}}^{\mu} = \overline{u}^i \gamma^\mu (1 - \gamma^5) d^i.$$
The *hadronic neutral weak current* is $J_{\text{had},n}^{\mu}$ where $2J_n^\mu =$
$$\overline{u}^i \gamma^\mu \left(1 - \gamma^5 - \frac{8}{3} \sin^2(\theta_W) \right) u^i + \overline{d}^i \gamma^\mu \left(\frac{4}{3} \sin^2(\theta_W) - 1 + \gamma^5 \right) d^i$$
The Lagrangian then reduces to (omitting the had's):
$$\begin{aligned}\mathcal{L}_{\text{quark}}^{\text{EW}} = \sum_{i=1}^{3} \Bigg(& i\overline{u}^i \not{\partial} u^i + i\overline{d}^i \not{\partial} d^i - eA_\mu J_{\text{EM}}^{\mu} \\ & - \frac{g}{2\sqrt{2}} \left(W_\mu^+ J^\mu + W_\mu^- J^{\mu\dagger} \right) - \frac{g}{2\cos(\theta_W)} J_n^\mu Z_\mu^0 \Bigg).\end{aligned}$$

For convenience, we define the *full currents* as the sum of the leptonic and hadronic pieces:

Definition: The full *electromagnetic current* is:
$$J_{\text{EM}}^{\mu} = \sum_f q_f \overline{f} \gamma^\mu f,$$
where the sum is over all fermion species f (both quarks and leptons), and q_f is the charge of the species f. The full *charged weak current* is
$$J^\mu = \sum_{i=1}^{3} \overline{\nu}_{e_i} \gamma^\mu (1 - \gamma^5) e_i + \sum_{i=1}^{3} \overline{u}^i \gamma^\mu (1 - \gamma^5) d^i,$$
where $e_1 = e$, $e_2 = \mu$ and $e_3 = \tau$ are the lepton generations. Finally, the full *neutral weak current* is
$$J_n^\mu = \frac{1}{2} \sum_f \left[2I_f \overline{f} \gamma^\mu (1 - \gamma^5) f - 4\sin^2(\theta_W) q_f \overline{f} \gamma^\mu f \right],$$
where I_f is the fermion f's *weak isospin* (equal to $\frac{1}{2}$ for ν_{e_i}, u, c, t, and equal to $-\frac{1}{2}$ for e_i, d, s, b).

With this Definition, we can completely characterise the gauge boson-fermion interactions in the electroweak theory using the Lagrangian: $\mathcal{L}_{\text{ferm}}^{\text{EW}} =$
$$\sum_f \overline{f} i\not{\partial} f - eA_\mu J_{\text{EM}}^{\mu} - \frac{g}{2\sqrt{2}} \left(W_\mu^+ J^\mu + W_\mu^- J^{\mu\dagger} \right) - \frac{g}{2\cos(\theta_W)} J_n^\mu Z_\mu^0.$$

As for leptons, we give the quarks mass by coupling them to the Higgs boson. The interaction Lagrangian is:
$$\begin{aligned}\mathcal{L}_{\text{quark},\phi} = -\sqrt{2} \Big(& \lambda_d^{ij} \overline{Q}_L^i \phi d_R^j + \lambda_u^{ij} \overline{Q}_L^i \phi^c u_R^j \\ & + (\lambda_d^*)^{ij} \overline{d}_R^i \phi^\dagger Q_L^j + (\lambda_u^*)^{ij} \overline{u}_L^i (\phi^c)^\dagger Q_R^j \Big),\end{aligned}$$
where $(\phi^c)^\alpha = \epsilon^{\alpha\beta} (\phi^\dagger)^\beta$, where $\epsilon^{\alpha\beta}$ is the Levi-Civita symbol. We need this odd field in the second term to make the hypercharges sum to 0. The $\lambda_{u/d}$ matrices are called *Yukawa matrices*.

Theorem: $\mathcal{L}_{\text{quark},\phi}$ is gauge invariant.

Proof: First and third terms are trivial. For second and fourth terms, just need to work out transformation of ϕ^c. Under $U(1)$, we have $\phi^c = \epsilon\phi^* \mapsto \epsilon(e^{\frac{1}{2}i\beta}\phi)^* = e^{-\frac{1}{2}i\beta}\phi^c$. So ϕ^c has hyperchage $-1/2$, and thus the second term transforms correctly under $U(1)$.

Now under $SU(2)$, we have

$$\begin{aligned}\phi^c = \epsilon\phi^* &\mapsto \epsilon e^{-i(\alpha^1\tau^1 - \alpha^2\tau^2 + \alpha^3\tau^3)}\phi^* \\ &= \epsilon\phi^* - i\epsilon\left(\alpha^1\tau^1 - \alpha^2\tau^2 + \alpha^3\tau^3\right)\phi^*\end{aligned}$$

where we've expanded infinitesimally. Now recall that $\tau^a = \frac{1}{2}\sigma^a$, and

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = i\sigma^2.$$

Now using properties of the Pauli matrices, we can commute things to find

$$\phi^c = e^{i(\alpha^1\tau^1 + \alpha^2\tau^2 + \alpha^3\tau^3)}\phi^c.$$

So ϕ^c transforms in the fundamental rep of $SU(2)$, and hence we're done. □

Theorem: Under SSB, expanding $\phi = (0, (v+h)/\sqrt{2})^T$:

$$\begin{aligned}\mathcal{L}_{\text{quark},\phi} = &-v\lambda_d^{ij}\overline{d}_L^i d_R^j - v\lambda_u^{ij}\overline{u}_L^i u_R^j - h\lambda_d^{ij}\overline{d}_L^i d_R^j - h\lambda_u^{ij}\overline{u}_L^i u_R^j \\ &-v(\lambda_d^*)^{ij}\overline{d}_R^i d_L^j - v(\lambda_u^*)^{ij}\overline{u}_R^i u_L^j - h(\lambda_d^*)^{ij}\overline{d}_R^i d_L^j - h(\lambda_u^*)^{ij}\overline{u}_R^i u_L^j.\end{aligned}$$

Proof: Trivial exercise. □

4.7 The CKM matrix

Naturally, we wish to diagonalise the mass terms as in the case of leptons. Using exactly the same proof as we did for leptons, we can diagonalise:

$$\lambda_u = K_u\Lambda_u S_u^\dagger, \quad \lambda_d = K_d\Lambda_d S_d^\dagger,$$

where $K_{u/d}$ and $S_{u/d}$ are unitary matrices. Transform the quark fields as

$$u_L \mapsto K_u u_L, \quad u_R \mapsto S_u u_R, \quad d_L \mapsto K_d d_L, \quad d_R \mapsto S_d d_R.$$

Then the mass terms become completely diagonal and we're left with:

$$\mathcal{L}_{\text{quark},\phi} = -\sum_{i=1}^{3}\left(m_d^i\overline{d}^i d^i + m_u^i\overline{u}^i u^j + h\Lambda_d^{ii}\overline{d}^i d^i - h\Lambda_u^{ii}\overline{u}^i u^i\right),$$

where $m_{u/d}^i = v\Lambda_{u/d}^{ii}$.

Unlike leptons, $\mathcal{L}_{\text{quark}}^{\text{EW}}$ is *not* invariant under diagonalisation. The charged quark current transforms as

$$\begin{aligned}J^\mu = \overline{u}^i\gamma^\mu(1-\gamma^5)d^i &\mapsto \overline{u}^i\gamma^\mu(1-\gamma^5)d^i \\ = 2\overline{u}_L^i\gamma^\mu d_L^i \mapsto 2\overline{u}_L^i\gamma^\mu(K_u^\dagger K_d)^{ij}d_L^j &= \overline{u}^i\gamma^\mu(1-\gamma^5)(K_u^\dagger K_d)^{ij}d^j.\end{aligned}$$

Definition: The matrix $V_{\text{CKM}} = K_u^\dagger K_d$ is called the *Cabibbo-Kobyashi-Maskawa* (CKM) matrix.

A non-diagonal V_{CKM} leads to the $W^\pm$ boson mediating intergenerational quark couplings.

4.8 Cabibbo mixing

Theorem: If there are only two generations, WLOG the CKM matrix can have the form:

$$V_{\text{CKM}} = \begin{pmatrix} \cos(\theta_C) & \sin(\theta_C) \\ -\sin(\theta_C) & \cos(\theta_C) \end{pmatrix},$$

where θ_C is called the *Cabibbo angle*. Mixing between generations in this case is called *Cabibbo mixing*.

Proof: Note V_{CKM} is unitary. A general unitary matrix may be expressed in the form

$$V_{\text{CKM}} = \begin{pmatrix} \cos(\theta_C)e^{i\alpha} & \sin(\theta_C)e^{i\beta} \\ -\sin(\theta_C)e^{i(\alpha+\gamma)} & \cos(\theta_C)e^{i(\beta+\gamma)} \end{pmatrix}$$

where θ_C is some angle, and $e^{i\alpha}$, $e^{i\beta}$, $e^{i\gamma}$ are phases.

We now perform an operation called *quark rephasing*. Using $U(1)$ transformations of the form $q^i \mapsto e^{i\phi}q^i$, we can eliminate the phases $e^{i\alpha}$, $e^{i\beta}$, $e^{i\gamma}$. This is possible, since there are four quark fields, and we are allowed to remove as many *relative* phases as we like (there are 3 relative phases for 4 fields). This leaves the desired matrix. □

Example: Consider a two-generation model where the quark mass Lagrangian is off the form:

$$\mathcal{L}_m = -\frac{1}{2}\left(\overline{q}_+ m_+(1+\gamma^5)q_+ + \overline{q}_- m_-(1+\gamma^5)q_- + \text{h.c.}\right),$$

where h.c. denotes Hermitian conjugate, and

$$q_+ = \begin{pmatrix} u' \\ c' \end{pmatrix}, \; q_- = \begin{pmatrix} d' \\ s' \end{pmatrix}, \; m_+ = \begin{pmatrix} 0 & a \\ a^*b & \end{pmatrix}, \; m_- = \begin{pmatrix} 0 & c \\ c^* & d \end{pmatrix},$$

where b and d are assume real. Define the matrix $R(\theta)$ by

$$R(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix},$$

and define the angle θ_+ and the masses m_u, m_c by

$$R(\theta_+)\begin{pmatrix} 0 & |a| \\ |a| & b \end{pmatrix}R(\theta_+)^{-1} = \begin{pmatrix} m_u & 0 \\ 0 & -m_c \end{pmatrix}.$$

This is possible because the matrix we are trying to diagonalise is real symmetric. Similarly define θ_- for m_s, m_d.

Note that if we write $a = |a|e^{i\phi}$ for some ϕ, then

$$m_+ = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & |a| \\ |a| & b \end{pmatrix}\begin{pmatrix} e^{-i\phi} & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence we see that rephasing the quark fields as

$$q_+ \mapsto \begin{pmatrix} e^{i\phi} & 0 \\ 0 & 1 \end{pmatrix} q_+,$$

and then applying the rotation $q_+ \mapsto R(\theta_+)^{-1} q_+$, we can diagonalise the $+$ mass terms. Similarly for $-$ terms.

Recall the V_{CKM} matrix is obtained from the charged weak current:

$$\overline{q}_+ \gamma^\mu (1-\gamma^5) q_- \mapsto \overline{q}_+ R(\theta_+) R(\theta_-)^{-1} \gamma^\mu (1-\gamma^5) q_-.$$

So $V_{\text{CKM}} = R(\theta_C) = R(\theta_+)R(\theta_-)^{-1}$ and hence the Cabibbo angle is given by $\theta_C = \theta_+ - \theta_-$. By expanding the definitions of θ_+ and θ_-, we find that

$$\theta_- = \arctan\left(\frac{m_d}{m_s}\right), \quad \theta_+ = \arctan\left(\frac{m_u}{m_c}\right),$$

so we have found θ_C explicitly in this case.

For a three-generation model, we can repeat the proof of the Theorem. This time, 9 parameters describe V_{CKM}, split as 3 angles and 6 phases. There are 6 quark fields, so 5 relative phases, so we can remove 5 phases by quark rephasing.

Therefore, V_{CKM} is parametrised by 3 angles and 1 phase, hence is *not* real. It follows that the Yukawa matrices are not real; from here, it is clear that CP symmetry *must* be violated in the Standard Model by intergenerational quark coupling. In particular, by the CPT Theorem, T symmetry must also be violated by intergenerational quark coupling.

4.9 Neutrino oscillations and mass

In some solar neutrino experiments in the 2000s, the number of electron neutrinos detected was smaller than predicted. It was theorised that electron neutrinos *oscillated* into muon and tau neutrinos.

This phenomenon can be explained by two possible models:

1. **Neutrinos are Dirac fermions.** If neutrinos are Dirac, then there *must* be right-handed neutrinos. We write

$$N^i = \nu_R^i = (\nu_{e_R}, \nu_{\mu_R}, \nu_{\tau_R}).$$

The lepton-Higgs coupling then takes the same form as the quark version:

$$\mathcal{L}_{\text{lept},\phi} = -\sqrt{2}\left(\lambda^{ij}\overline{L}^i \phi R^j + \lambda_\nu^{ij}\overline{L}^i \phi^c N^j + \text{h.c.}\right)$$

We diagonalise this matrix in exactly the same way as the quark version, and get the same mixing term in the charged weak current. This time the mixing matrix is called the *Pontecorvo-Maki-Nakagawa-Sakata matrix*, and is denoted U_{PMNS}.

2. **Neutrinos are Majorana fermions.** Since neutrinos are neutral, they could be their own antiparticles, i.e. they are Majorana fermions. The mode expansion of a Majorana fermion is of the form:

$$\nu(x) = \sum_{s,p} b^s(p) u^s(p) e^{-ip\cdot x} + b^{s\dagger}(p) v^s(p) e^{ip\cdot x}.$$

Indeed, under charge conjugation, this field transforms as $\hat{C}\nu(x)\hat{C}^{-1} = C\overline{\nu}^T(x) = \nu(x)$, showing it is its own antiparticle.

Furthermore, for Majorana fermions, the left and right-handed fields are *not* independent. Indeed, we have

$$\nu_R(x) = \hat{C}\nu_L(x)\hat{C}^{-1} = C\overline{\nu}_L^T(x).$$

Therefore Majorana mass terms look like:

$$\mathcal{L}_{m_\nu,\text{Majorana}} = -\frac{1}{2}\sum_i m_\nu^i \left(\overline{\nu}_L^{i,C}\nu_L^i + \overline{\nu}_L^i \nu_L^{i,C}\right),$$

where $\psi^C = \hat{C}\psi\hat{C}^{-1}$. What terms in the unbroken Lagrangian generate such mass terms?

It turns out the right operator (which is $SU(2)\times U(1)$ invariant) to introduce is

$$\mathcal{L}_{\nu\phi,\text{Majorana}} = -\frac{Y^{ij}}{M}(L^{iT}\tilde{\phi})C(\tilde{\phi}^T L^j) + \text{h.c.},$$

where $\tilde{\phi}^\alpha = \epsilon^{\alpha\beta}\phi^\beta$.

4.10 The full electroweak Lagrangian

The full electroweak Lagrangian is composed of all the parts we've discussed above. Before spontaneous symmetry breaking, the Lagrangian is:

$$\mathcal{L}_{\text{EW}} = \underbrace{-\frac{1}{2}\mathrm{Tr}(F^W_{\mu\nu}F^{W\mu\nu}) - \frac{1}{4}F^B_{\mu\nu}F^{B\mu\nu}}_{\text{gauge boson kinetic terms}} + \underbrace{(D_\mu\phi)^\dagger D^\mu\phi}_{\text{Higgs kinetic term}} - \underbrace{\mu^2|\phi|^2 - \lambda|\phi|^4}_{\text{Higgs self-interactions}} + \sum_{i=1}^{3}\left(\underbrace{\overline{Q}^i_L i\not{D}Q^i_L + \overline{u}^i_R i\not{D}u^i_R + \overline{d}^i_R i\not{D}d^i_R}_{\text{quark kinetic terms}} + \underbrace{\overline{L}^i i\not{D}L^i + \overline{R}^i i\not{D}R^i}_{\text{lepton kinetic terms}}\right)$$

$$-\sqrt{2}\left(\underbrace{\lambda^{ij}_d\overline{Q}^i_L\phi d^j_R + \lambda^{ij}_u\overline{Q}^i_L\phi^c u^j_R + \text{h.c.}}_{\text{quark-Higgs couplings}}\right) - \sqrt{2}\left(\underbrace{\lambda^{ij}\overline{L}^i\phi R^j + \text{h.c.}}_{\text{lepton-Higgs couplings}}\right)$$

Recall that the λ matrices are called *Yukawa couplings* and $(\phi^c)^\alpha = \epsilon^{\alpha\beta}(\phi^\dagger)^\beta$. After spontaneous symmetry breaking, and diagonalisation of the quark and lepton fields, we have:

$$\mathcal{L}_{\text{EW}} = \underbrace{-\frac{1}{2}F^{W^\pm\dagger}_{\mu\nu}F^{W^\pm\mu\nu} - \frac{1}{4}F^A_{\mu\nu}F^{A\mu\nu} - \frac{1}{4}F^Z_{\mu\nu}F^{Z\mu\nu} + \frac{1}{2}|\partial h|^2 + \sum_f \overline{f}i\not{\partial}f}_{\text{kinetic terms}} + \underbrace{m_W^2 W^+_\mu W^{-\mu} + \frac{1}{2}m_Z^2 Z^0_\mu Z^{0\mu} - \frac{1}{2}m_h^2 h^2 - \sum_f m_f\overline{f}f}_{\text{mass terms}}$$

$$\underbrace{-eA_\mu J^\mu_{\text{EM}} - \frac{g}{2\sqrt{2}}\left(W^+_\mu J^\mu + W^-_\mu J^{\mu\dagger}\right) - \frac{g}{2\cos(\theta_W)}J^\mu_n Z^0_\mu}_{\text{gauge boson-fermion interactions}} \underbrace{- igW^{+\mu}W^{-\nu}\left(\sin(\theta_W)F^A_{\mu\nu} + \cos(\theta_W)F^Z_{\mu\nu}\right)}_{W\text{ boson-photon and }W\text{ boson-}Z\text{ boson interactions}} + \underbrace{\frac{1}{2}g^2\left(W^{+2}W^{-2} - (W^{+\mu}W^-_\mu)^2\right)}_{W\text{ boson self-interactions}}$$

$$\underbrace{+(gm_W h + \frac{1}{4}g^2h^2)\left(W^+_\mu W^{-\mu} + \frac{1}{2}\sec^2(\theta_W)Z^0_\mu Z^{0\mu}\right)}_{\text{gauge boson-Higgs interactions}} + \underbrace{\frac{gm_h^2}{4m_W^2}h^3 + \frac{g^2m_h^2}{8m_W^2}h^4}_{\text{Higgs self-interactions}} - \underbrace{\sum_f \frac{gm_f}{2m_W}\overline{f}fh}_{\text{Higgs-fermion interactions}} .$$

where sums over f indicate sums over all fermion flavours (where m_f is the mass of the fermion of flavour f), i.e. including both lepton and quark fields. The currents are given by:

$$J^\mu_{\text{EM}} = \sum_f q_f\overline{f}\gamma^\mu f, \qquad J^\mu = \sum_{i=1}^{3}(\overline{\nu}_{e_i}\gamma^\mu(1-\gamma^5)e_i) + \sum_{i=1,j=1}^{3}\left(\overline{u}^i\gamma^\mu(1-\gamma^5)V^{ij}_{\text{CKM}}d^j\right), \qquad \text{and} \qquad J^\mu_n = \frac{1}{2}\sum_f\left[2I_f\overline{f}\gamma^\mu(1-\gamma^5)f - 4\sin^2(\theta_W)q_f\overline{f}\gamma^\mu f\right],$$

where q_f is the charge of fermion f and I_f is the isospin of fermion f ($\frac{1}{2}$ for right handed particles, and $-\frac{1}{2}$ for left handed particles). Notice that throughout we have eliminated certain couplings in favour of masses, the couplings g, e, and θ_W using the relationships:

$$-\frac{\mu^2}{2\lambda} = \frac{v^2}{2}, \quad \cos(\theta_W) = \frac{g}{\sqrt{g^2+g'^2}}, \quad m_W = \frac{1}{2}gv, \quad m_Z = \frac{1}{2}gv\sec(\theta_W), \quad m_h = \sqrt{2\lambda v^2}, \quad e = g\sin(\theta_W).$$

From these we can derive results like $v = 2m_W/g$ and $\lambda = m_h^2g^2/8m_W^2$, allowing us to achieve this task.

5 Weak decays

5.1 Fermi effective theory

We'll mainly consider processes where energies and momenta are lower than the masses of the W and Z boson, m_W and m_Z. We then use *Fermi effective theory*.

To derive the effective theory, we need the facts:

Theorem: The W and Z propagators, given by,

$$D^W_{\mu\nu}(x-x') = \langle 0|T\{W^-_\mu(x)W^+_\nu(x')\}|0\rangle\,,$$
$$D^Z_{\mu\nu}(x-x') = \langle 0|T\{Z^0_\mu(x)Z^0_\nu(x')\}|0\rangle\,.,$$

have the Fourier transform:

$$\tilde{D}^{Z/W}_{\mu\nu}(p) = \frac{i}{p^2 - m^2_{Z/W} + i\epsilon}\left(-\eta_{\mu\nu} + \frac{p_\mu p_\nu}{m^2_{Z/W}}\right).$$

Proof: Recall from QFT the propagator is i times the Green's function of the EL equation for the *free* field.

The kinetic part of the electroweak Lagrangian for the Z boson is

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}(\partial_\mu Z_\nu - \partial_\nu Z_\mu)(\partial^\mu Z^\nu - \partial^\nu Z^\mu) + \frac{1}{2}m_Z^2 Z_\mu Z^\mu.$$

Hence the Euler-Lagrange equations are

$$\partial^2 Z_\rho - \partial_\rho \partial\cdot Z + m_Z^2 Z_\rho = 0.$$

If we add a source $j^\mu(x)$, the Lagrangian is appended by $Z_\mu j^\mu$ and the equation of motion becomes:

$$\partial^2 Z_\rho - \partial_\rho \partial\cdot Z + m_Z^2 Z_\rho = -j_\rho.$$

Note that taking the divergence of the above equation, we have $m_Z^2\partial\cdot Z = -\partial\cdot j$; substituting this back in, we have

$$(\partial^2 + m_Z^2)Z_\mu = -\left(\eta_{\mu\nu} + \frac{\partial_\mu\partial_\nu}{m_Z^2}\right)j^\nu.$$

Take the Fourier transform of this equation to obtain:

$$\tilde{Z}_\mu(p) = \frac{1}{p^2 - m_Z^2}\left(\eta_{\mu\nu} - \frac{p_\mu p_\nu}{m_Z^2}\right)\tilde{j}^\nu(p).$$

By the definition of a Green's function, we have

$$Z_\mu(x) = i\int d^4x'\; D^Z_{\mu\nu}(x-x')j^\nu(x').$$

Taking the Fourier transform (recall convolution becomes product):

$$\tilde{Z}_\mu(p) = i\tilde{D}^Z_{\mu\nu}(p)\tilde{j}^\nu(p),$$

then just compare to get result. Exactly the same for the W boson (just replace $Z \mapsto W$ and $m_Z \mapsto m_W$). □

Theorem: When a process has energies and momenta much less than m_W and m_Z, we may replace the weak interaction part of the Lagranian, $\mathcal{L}_W$, with the *Fermi effective Lagrangian*:

$$\mathcal{L}_W^{\text{eff}} = -\frac{G_F}{\sqrt{2}}\left(J^{\mu\dagger}(x)J_\mu(x) + \rho J_n^{\mu\dagger}(x)J_{n\mu}(x)\right),$$

where

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2}, \qquad \rho = \frac{m_W^2}{m_Z^2\cos^2(\theta_W)}.$$

Proof: The weak part of the Lagrangian in the electroweak theory is:

$$\mathcal{L}_W = -\frac{g}{2\sqrt{2}}\left(J^\mu W^+_\mu + J^{\mu\dagger}W^-_\mu\right) - \frac{g}{2\cos(\theta_W)}J_n^\mu Z^0_\mu.$$

The scattering matrix is

$$S = \mathrm{Texp}\left(-i\int d^4x\; \mathcal{L}_W(x)\right).$$

Since g is a small coupling constant at low energies, we can expand the matrix element using Dyson's formula:

$$\langle f|S|i\rangle = \langle f|I|i\rangle - \frac{g^2}{8}\langle f|\int d^4x d^4x'\; T\Big\{J^{\mu\dagger}D^W_{\mu\nu}(x-x')J^\nu(x')$$
$$+\frac{1}{\cos^2(\theta_W)}J_n^{\mu\dagger}(x)D^Z_{\mu\nu}(x-x')J_n^\nu(x')\Big\}|i\rangle + O(g^4).$$

where we've assume the W and Z bosons are not in the initial or final states, so that there's no $O(g)$ term, and no $O(g^2)$ cross term. We've also used Wick's Theorem to obtain the W and Z propagators, $D^W_{\mu\nu}$ and $D^Z_{\mu\nu}$.

From the propagator proof above, we know that for $m^2_{Z/W} \gg p^2$, we have

$$\tilde{D}^{Z/W}_{\mu\nu}(p) \approx \frac{i\eta_{\mu\nu}}{m^2_{Z/W}} \;\Rightarrow\; D^{Z/W}_{\mu\nu}(x-x') = \frac{i\eta_{\mu\nu}}{m^2_{Z/W}}\delta^4(x-x').$$

Substituting into the matrix element, this immediately gives the result. (i can be removed, since this is only an overall phase.) □

Definition: G_F is called the *Fermi coupling* and ρ is called the *rho-parameter*.

Recall that in the classical electroweak theory, we showed that $m_W^2 = m_Z^2\cos^2(\theta_W)$. Therefore, in the quantum theory we may write $\rho = 1 + \Delta\rho$, where $\Delta\rho$ comes from quantum loop effects.

Finally, note that $[G_F] = -2$, to compensate for $[J^{\mu\dagger}J_\mu] = 6$. Thus this theory is *non-renormalisable*; it does not hold to arbitrarily high energies. However, we were expecting this, since we assumed that we working well below the energy scale $O(m_W^2, m_Z^2)$.

5.2 Cross sections and decay rates review

Decay rates

Definition: The *decay rate* Γ_X is:

$$\Gamma_X = \frac{\text{number of decays of } X \text{ observed}}{\text{time taken} \times \text{number of } X \text{ in sample}}.$$

Definition: The *lifetime* τ_X is defined by $1/\Gamma_X$.

Theorem: For a process where X can decay into any of a set of final states f, the decay rate is given by

$$\Gamma_X = \frac{1}{2m_X} \sum_f \int |\mathcal{M}_{fX}|^2 \, d\rho_f,$$

where m_X is the mass of X, $\mathcal{M}_{fX}$ is the *invariant amplitude*, defined in QFT by

$$\langle f|S - I|X\rangle = (2\pi)^4 \delta^4 \left(p_X - \sum_r p_r \right) i\mathcal{M}_{fX},$$

and $d\rho_f$ is the *invariant integration measure*, given by

$$d\rho_f = (2\pi)^4 \delta^4 \left(p_X - \sum_r p_r \right) \prod_r \left(\frac{d^3 p_r}{(2\pi)^3 \cdot 2p_r^0} \right),$$

where r ranges across the particles in one of the final states f, which have momenta p_r.

Proof: We use $\langle f|S - I|X\rangle$, since we want to exclude the possibility that no decay occurs. Then the probability of the decay $X \to f$ is

$$\mathbb{P}(X \to f) = \frac{|\langle f|S - I|X\rangle|^2}{\langle f|f\rangle \langle X|X\rangle}.$$

In QFT, we used wavepackets to deal with the possibility that $\langle f|f\rangle$, $\langle X|X\rangle$ could be infinite. Here, we'll instead work in finite spatial volume V and temporal extent T. In particular, the delta functions become

$$(2\pi)^3 \delta^3(\mathbf{0}) \mapsto V, \quad (2\pi)^4 \delta^4(0) \mapsto VT.$$

Recall that with relativistic normalisation, we have

$$\langle X|X\rangle = (2\pi)^3 2p_X^0 \delta^3(0) \mapsto 2p_X^0 V, \quad \langle f|f\rangle = \prod_r \left(2p_r^0 V\right).$$

In the rest frame, $p_X^0 =_X$. Hence

$$\mathbb{P}(X \mapsto f) = \frac{|\mathcal{M}_{fX}|^2 (2\pi)^4 \delta^4 \left(p_i - \sum_r p_r \right) VT}{2m_X V \cdot \prod_r (2p_r^0 V)}$$

Note the VT in the numerator comes from the δ function squared in $|\langle f|S - I|X\rangle|^2$.

We now convert this to the decay rate $\Gamma(X \to f)$, then sum over final states f to get the answer. The decay rate is, by definition: $\Gamma(X \to f) =$

$$\frac{\mathbb{P}(X \to f)}{T} = \frac{|\mathcal{M}_{fX}|^2 (2\pi)^4}{2m_X} \delta^4 \left(p_i - \sum_r p_r \right) \prod_r \left(\frac{1}{2p_r^0 V} \right).$$

We can't ever measure the momenta p_r with exact precision, though. So we have to integrate over all possible 1-particle states in a box V with momentum p_r, for each r, which is given by

$$\prod_r \left(\frac{V d^3 p_r}{(2\pi)^3} \right).$$

So we're left with the final expression, as required. □

Cross-sections

Definition: Suppose we fire a beam of particle at a target. Let n be the number of scattering events per unit time, divided by the number of target particles. Let the *incident flux* be F, i.e. the number of incoming particles per unit area per unit time. Then the *cross-section* is defined by $\sigma = n/F$.

Theorem: The cross section for the process $i \to f$ (with two initial particles in i) is given by

$$d\sigma = \frac{|M_{fi}|^2}{\mathcal{F}} d\rho_f,$$

where $\mathcal{F}$ is the *flux factor*, given by $\mathcal{F} = 4E_1 E_2 |\mathbf{v}_a - \mathbf{v}_b|$. Here, $|\mathbf{v}_a - \mathbf{v}_b|$ is the relative velocity of the incident beam to the target, and E_1, E_2 are the energies of the two initial particles.

Proof: The total number of scattering events per unit time is $N = n\rho_b V$, where ρ_b is the density of the target, and V is the finite spatial volume. The incident flux is $F = |\mathbf{v}_a - \mathbf{v}_b|\rho_a$, where ρ_a is the density of the incident beam, and $\mathbf{v}_a - \mathbf{v}_b$ is the relative velocity of the incident beam to the target. Hence

$$N = n\rho_b V = F\sigma\rho_b V = |\mathbf{v}_a - \mathbf{v}_b|\rho_a \rho_b V\sigma.$$

The normalisation $\langle i|i\rangle = 2p_i^0 V$ corresponds to having one particle per unit volume; similarly for $\langle f|f\rangle$. Hence $\rho_a = \rho_b = 1/V$. So

$$N = \frac{|\mathbf{v}_a - \mathbf{b}|\sigma}{V} \quad \Rightarrow \quad dN = \frac{|\mathbf{v}_a - \mathbf{v}_b|}{V} d\sigma. \qquad (*)$$

Now use decay rate derivation to find dN, except since we have two incoming particles, we get $\langle i|i\rangle = (2E_1 V)(2E_2 V)$ in the denominator. This gives

$$dN = \frac{1}{(2E_1)(2E_2)V} |\mathcal{M}_{fi}|^2 d\rho_f.$$

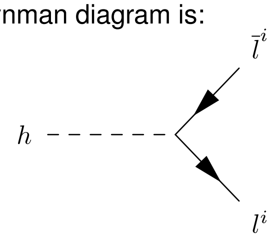
Substituting $(*)$ and rearranging, we get the result. □

5.3 Higgs decay

Example: Consider Higgs to lepton decay $h \to \overline{l}^i l^i$, where l^i is a specific lepton (either e, μ or τ). The coupling of the Higgs to leptons is given in the electroweak Lagrangian by

$$\mathcal{L}_{\text{int}} = -\sum_i \lambda_i h \overline{l}^i l^i,$$

where λ_i is a Yukawa coupling. Thus at tree-level the only contributing Feynman diagram is:



So by the Feynman rules, the tree-level amplitude for $h \to \overline{l}l$ decay is $i\mathcal{M} = -i\lambda_i \overline{u}^s(q)v^r(k)$.

To get the probability, we need to sum over final spins and average over initial spins. The Higgs boson is spinless, so the probability of decay is

$$\begin{aligned}\sum_{s,r} |\mathcal{M}|^2 &= \lambda_i^2 \sum_{s,r} [\overline{v}^r_\alpha(k)u^s_\alpha(q)][\overline{u}^s(q)_\beta v^r_\beta(k)] \\ &= \lambda_i^2 \mathrm{Tr}((\not{q} + m_i)(\not{k} - m_i)).\end{aligned}$$

Here, we've used results about the plane wave spinors from QFT. Now use results on traces of γ matrices from QFT to simplify to

$$\sum_{s,r} |\mathcal{M}|^2 = 4\lambda_i^2(k \cdot q - m_i^2).$$

It's now quickest to recognise that there is a 4-momentum conserving δ function in the decay rate formula; hence we can impose 4-momentum conservation now. Squaring the conservation law $p = k + q$, the probability simplifies to $2\lambda_i^2(m_h^2 - 4m_i^2)$, where m_h is the mass of the Higgs.

To get the decay rate, the only non-trivial integral we need to do is the integral over $d\rho_f$, which we perform in the rest frame, so that $p = (m_h, \mathbf{0})^T$. Calculating, we have:

$$\begin{aligned}&\frac{1}{4\pi^2}\int \frac{d^3\mathbf{k}}{2k^0}\frac{d^3\mathbf{q}}{2q^0}\delta^4(p - k - q) \\ &= \frac{1}{4\pi}\int_0^\infty \frac{d|\mathbf{k}|}{m_i^2 + |\mathbf{k}|^2}\,|\mathbf{k}|^2 \delta\left(m_h - 2\sqrt{m_i^2 + |\mathbf{k}|^2}\right).\end{aligned}$$

It's now simple to use a standard δ function identity to do the integral. The answer we obtain is

$$\Gamma = \frac{\lambda_i^2}{8\pi m_h^2}(m_h^2 - 4m_i^2)^{3/2}.$$

We can write this in terms of G_F, the Fermi coupling, using the fact that the Yukawa couplings are given by

$$\lambda_i^2 = \frac{m_i^2}{v^2} = \frac{m_i^2 g^2}{4m_W^2} = 2m_i^2 \cdot \frac{G_F}{\sqrt{2}},$$

where we've also used the fact that $m_W^2 = v^2 g^2/4$, from a long time ago when we studied the electroweak theory. Therefore, the final answer may be written as

$$\Gamma = \frac{G_F}{\sqrt{2}} \cdot \frac{1}{4\pi} \cdot \frac{m_i^2}{m_h^2}(m_h^2 - 4m_i^2)^{3/2}.$$

5.4 Z boson decay

Example: Consider Z boson decay into two specific leptons: $Z \to \overline{l}l$. We can't use Fermi effective theory, as this wipes away the existence of the Z and W bosons. So we return back to the coupling the electroweak Lagrangian, given by:

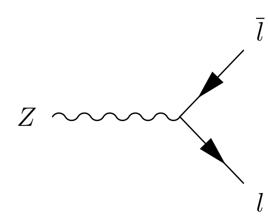
$$-\frac{g}{2\cos(\theta_W)} J_n^\mu Z_\mu,$$

where J_n^μ is the leptonic neutral weak current. Recall that this can be written as:

$$J_n^\mu = \sum_f \overline{f}\gamma^\mu \left[I_f(1-\gamma^5) - 2\sin^2(\theta_W)q_f\right] f = \sum_f \overline{f}\gamma^\mu(v - a\gamma^5)f,$$

where $v = I_f - 2\sin^2(\theta_W)q_f$ and $a = I_f$; here v and a depend on the species of lepton in question (whether it is an electron, neutrino, etc).

The only contributing Feynman diagram is:



which by the Feynman rules has amplitude:

$$i\mathcal{M} = \frac{ig}{2\cos(\theta_W)}[\overline{u}_l(k)\gamma^\mu(v - a\gamma^5)v_l(q)]\epsilon_\mu(p, \lambda).$$

Here, $\epsilon_\mu(p, \lambda)$ is a *polarisation vector*. Now's a good time to review what this means.

Recall that the quantum Z field has mode expansion:

$$Z_\mu = \sum_{p,\lambda}\left(a_Z(p,\lambda)\epsilon_\mu(p,\lambda)e^{-ip\cdot x} + a_Z^\dagger(p,\lambda)\epsilon_\mu^*(p,\lambda)e^{ip\cdot x}\right),$$

where λ is the *polarisation*, running over $\lambda = -1, 0, 1$. The operators a_Z and $a_Z^\dagger$ obey

$$[a_Z(p,\lambda), a_Z^\dagger(p',\lambda')] = \delta(p - p')\delta_{\lambda,\lambda'}.$$

Similarly, the quantum W field has mode expansion:

$$W_\mu = \sum_{p,\lambda} \left(a_W(p,\lambda)\epsilon_\mu(p,\lambda)e^{-ip\cdot x} + c_W^\dagger(p,\lambda)\epsilon_\mu^*(p,\lambda)e^{ip\cdot x} \right),$$

where $a_W^\dagger$ creates a W^+ particle and $c_W^\dagger$ creates a W^- particle.

The polarisation vectors satisfy the identities:

Theorem: We have the following:

(i) $\epsilon(p,\lambda)\cdot p = 0$.

(ii) $\epsilon^*(p,\lambda)\cdot\epsilon(p,\lambda') = -\delta_{\lambda,\lambda'}$.

(iii) $\displaystyle\sum_\lambda \epsilon_\mu(p,\lambda)\epsilon_\nu^*(p,\lambda) = -\eta_{\mu\nu} + \frac{p_\mu p_\nu}{m_{Z/W}^2}$.

Proof: (i) and (ii) can be arranged by Definition of polarisation vectors (see QED in QFT). To prove (iii), simply note that $\{\epsilon^*(p,\lambda), p\}$, when taken over $\lambda = -1,0,1$, forms a basis. So we can contract both sides with each basis element to check the identity holds. Contracting with $\epsilon^*(p,\lambda')$, On the LHS, we have:

$$\sum_\lambda \epsilon^{\mu *}(p,\lambda')\epsilon_\mu(p,\lambda)\epsilon_\nu^*(p,\lambda) = -\epsilon_\nu^*(p,\lambda')$$

and on the RHS, we have

$$-\epsilon_\nu^*(p,\lambda'),$$

since $\epsilon^*(p,\lambda)\cdot p = 0$ (since p is real, this follows from (i)). Just by looking at the equation, it's trivial the contraction with p also works. So we're done. □

BACK TO Z BOSON DECAY...

To calculate the probability, we sum $|\mathcal{M}|^2$ over the final spins, and average over the initial *polarisations*. There are three polarisations of the Z boson, so we want to find:

$$\frac{1}{3}\sum_{\substack{\text{spins}\\ \text{polarisations}}} |\mathcal{M}|^2$$

$$= \frac{g^2}{12\cos^2(\theta_W)}\sum_{r,s,\lambda}[\overline{v}_l^s(q)\gamma^\mu(v-a\gamma^5)u_l^r(k)]\epsilon_\mu^*(p,\lambda)\cdot$$
$$[\overline{u}_l^r(k)\gamma^\nu(v-a\gamma^5)v_l^s(q)]\epsilon_\nu(p,\lambda)$$

$$= \frac{g^2}{12\cos^2(\theta_W)}\mathrm{Tr}((\not{k}\gamma^\nu(v-a\gamma^5)\not{q}\gamma^\mu(v-a\gamma^5))\left(-\eta_{\mu\nu} + \frac{p_\mu p_\nu}{m_Z^2}\right)$$

where we've used the identity for the polarisation vectors. We've also *neglected fermion masses*, which are small compared to the mass of the Z boson.

Now use some trace identities from QFT:

$$\mathrm{Tr}(\gamma^{\mu_1}...\gamma^{\mu_{2n+1}}) = 0,$$
$$\mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) = 4(\eta^{\mu\nu}\eta^{\rho\sigma} - \eta^{\mu\rho}\eta^{\nu\sigma} + \eta^{\mu\sigma}\eta^{\nu\rho}),$$
$$\mathrm{Tr}(\gamma^5\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) = -4i\epsilon^{\mu\nu\rho\sigma},$$

These reduce the probability to

$$\frac{g^2}{3\cos^2(\theta_W)}(v^2+a^2)\left(k\cdot q + \frac{2(k\cdot p)(q\cdot p)}{m_Z^2}\right).$$

Pre-emptively use conservation of 4-momentum, $p = q+k$, to obtain:

$$p^2 = q^2 + 2q\cdot k + k^2 = 2q\cdot k \quad \Rightarrow \quad q\cdot k = \frac{1}{2}m_Z^2.$$

Similarly, by considering $(p-k)^2$ and $(p-q)^2$, we obtain $p\cdot k = \frac{1}{2}m_Z^2$ and $p\cdot q = \frac{1}{2}m_Z^2$. So the probability reduces to

$$\frac{g^2(v^2+a^2)m_Z^2}{3\cos^2(\theta_W)}.$$

It's now simple to insert into the decay rate formula, perform the standard integrals (in the *rest frame* of the Z boson) and get the answer:

$$\Gamma = \frac{g^2 m_Z(v^2+a^2)}{48\pi\cos^2(\theta_W)} = \frac{G_F}{\sqrt{2}}\cdot\frac{m_Z^2(v^2+a^2)}{6\pi},$$

using $G_F/\sqrt{2} = g^2/8m_W^2 = g^2/8m_Z^2\cos^2(\theta_W)$.

5.5 Muon decay

Example: Consider muon decay $\mu^-(p) \to e^-(k)\overline{\nu}_e(q)\nu_\mu(q')$. Since $m_\mu \approx 106$ MeV $\ll m_W \approx 80$ GeV, we can use Fermi effective theory.

Since this is a flavour-changing interaction (a muon turns into an electron), it is mediated by the W boson, and hence we only need to consider

$$\mathcal{L}_{\text{int}} = -\frac{G_F}{\sqrt{2}}J^{\mu\dagger}J_\mu,$$

where the relevant part of the charged weak current is

$$J^\alpha = \overline{\nu}_e\gamma^\alpha(1-\gamma^5)e + \overline{\nu}_\mu\gamma^\alpha(1-\gamma^5)\mu.$$

Instead of using Feynman diagrams to calculate the amplitude, we'll use a different technique:

Theorem: To tree-level, and up to some phase, for any process $i \to f$ with Lagrangian $\mathcal{L}(x)$, we have

$$\mathcal{M}_{fi} = \langle f|\mathcal{L}(0)|i\rangle.$$

Here, $\mathcal{L}(0)$ is the Lagrangian *evaluated at zero*.

Proof: Recall:

$$\langle f|S-I|i\rangle = i(2\pi)^4\delta^4(p_f-p_i)\mathcal{M}_{fi} = i\int d^4x\ \langle f|\mathcal{L}(x)|i\rangle\,.$$

Writing $\mathcal{L}$ in momentum space, we have

$$i(2\pi)^4\delta^4(p_f-p_i)\mathcal{M}_{fi} = i\int \frac{d^4p}{(2\pi)^4}\left(\int d^4x\ e^{-ip\cdot x}\right)\langle f|\tilde{\mathcal{L}}(p)|i\rangle\,.$$

We know from QFT that commuting operators in $\langle f|\tilde{\mathcal{L}}(p)|i\rangle$ will eventually enforce global momentum conservation, at least up to a phase and up to tree level, so that $p = p_f - p_i$ in the exponent. Thus integrating out x, we have:

$$i(2\pi)^4\delta^4(p_f-p_i)\mathcal{M}_{fi} = i(2\pi)^4\delta^4(p_f-p_i)\,\langle f|\tilde{\mathcal{L}}(0)|i\rangle\,,$$

up to a phase. The result follows. □

Thus in our case, we want to compute:

$$\begin{aligned}\mathcal{M} &= -\frac{G_F}{\sqrt{2}}\langle e^-(k)\overline{\nu}_e(q)\nu_\mu(q')|J^{\alpha\dagger}J_\alpha(0)|\mu^-(p)\rangle \\ &= -\frac{G_F}{\sqrt{2}}\langle e^-(k)\overline{\nu}_e(q)|\overline{e}(0)\gamma^\alpha(1-\gamma^5)\nu_e(0)|0\rangle \\ &\quad\cdot\langle\nu_\mu(q')|\overline{\nu}_\mu(0)\gamma_\alpha(1-\gamma^5)\mu(0)|\mu^-(p)\rangle\,.\end{aligned}$$

To get the final line, we've inserted J^α and $J^{\alpha\dagger}$, and considered mode expansions of the fields. To proceed, we insert the full mode expansions; this gives, for example:

$$\begin{aligned}\mu(0)\,|\mu^-(p)\rangle &= \sum_{s,\tilde{p}}\left(u^s(\tilde{p})b(\tilde{p}) + d^\dagger(\tilde{p})v^s(\tilde{p})\right)\sqrt{2E_{\mathbf{p}}}b^\dagger(p)\,|0\rangle \\ &= u^s(p) + (d^\dagger \text{ term which cancels to the left})\end{aligned}$$

Carrying out the whole computation, we find

$$\mathcal{M} = -\frac{G_F}{\sqrt{2}}[\overline{u}_e(k)\gamma^\alpha(1-\gamma^5)v_{\nu_e}(q)][\overline{u}_{\nu_\mu}(q')\gamma_\alpha(1-\gamma^5)u_\mu(p)],$$

where we've suppressed the spin indices, but there is a different one for each spinor. Since the initial state has two spin states, we must average and sum to get:

$$\begin{aligned}\frac{1}{2}\sum_{\text{spins}}|\mathcal{M}|^2 &= \frac{G_F^2}{4}\mathrm{Tr}((\not{k}+m_e)\gamma^\alpha(1-\gamma^5)\not{q}\gamma^\beta(1-\gamma^5)) \\ &\quad\cdot\mathrm{Tr}((\not{q}'\gamma_\alpha(1-\gamma^5)(\not{p}+m_\mu)\gamma_\beta(1-\gamma^5)).\end{aligned}$$

We used some trace identities from above, and we also needed to use the identity:

$$\epsilon^{\alpha\beta\sigma\rho}\epsilon_{\alpha\beta\lambda\tau} = -2(\delta^\sigma_\lambda\delta^\rho_\tau - \delta^\sigma_\tau\delta^\rho_\lambda).$$

The result is:

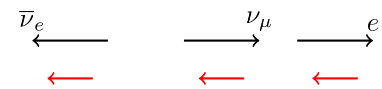
$$\frac{1}{2}\sum_{\text{spins}}|\mathcal{M}|^2 = 64G_F^2(p\cdot q)(k\cdot q').$$

Interlude: Consider the case when e, ν_μ go out along the $+z$ axis, and $\overline{\nu}_e$ goes out along the $-z$ direction. Then

$$k\cdot q' = \sqrt{m_e^2+k_z^2}q_z' - k_z q_z'.$$

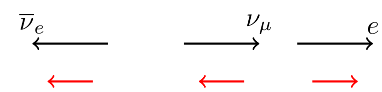
We see that if $m_e = 0$, the above probability is zero - this process never occurs!

The reason is because angular momentum conservation is violated by this process. For massless spinors, recall that *helicity* (spin in the direction of travel) is the same as *chirality*. Since the weak interaction only couples to left-handed particles, we must have spins as shown:



Spins are in direction of travel for particles, and opposite direction of travel for anti-particles, since they are all left-handed. This gives total spin $|S_z| = \frac{3}{2}$ after the interaction. But the total spin of the muon is $\frac{1}{2}$, so we have a contradiction.

When $m_e \neq 0$, the process can occur, since there is no longer a one to one correspondence between spin and helicity. Left and right handed spinors are coupled through a mass term so that this (valid) setup is possible instead:



We say that the process that *does* occur has its *helicity suppressed*.

Back to muon decay...

The decay rate is now given by:

$$\Gamma = \frac{(2\pi)^4\cdot 64G_F^2}{2\mu\cdot(2\pi)^9}\int\frac{d^3\mathbf{k}}{2k^0}\frac{d^3\mathbf{q}}{2q^0}\frac{d^3\mathbf{q}'}{2q'^0}\delta^4(p-k-q-q')(p\cdot q)(k\cdot q').$$

There is a clever trick for the evaluation of this integral. We do it in two pieces.

Piece 1: First, define

$$I_{\mu\nu}(p-k) = \int\frac{d^3\mathbf{q}}{|\mathbf{q}|}\frac{d^3\mathbf{q}'}{|\mathbf{q}'|}\delta^4(p-k-q-q')q_\mu q_\nu'.$$

Note $q^0 = |\mathbf{q}|\ q'^0 = |\mathbf{q}'|$, since neutrinos are assumed massless here. Since this is Lorentz covariant, and symmetric on μ, ν, it must be of the form:

$$I_{\mu\nu} = a\underbrace{(p-k)_\mu(p-k)_\nu}_{\substack{\text{RHS is Lorentz covariant,}\\ \text{and only depends on } p-k}} + b\eta_{\mu\nu}\underbrace{(p-k)\cdot(p-k)}_{\text{for convenience}}$$

To find a and b, we contract $I_{\mu\nu}$ with various things. Firstly,

$$(p-k)^2(a+4b) = \eta^{\mu\nu} I_{\mu\nu} = \int \frac{d^3\mathbf{q}}{|\mathbf{q}|}\frac{d^3\mathbf{q}'}{|\mathbf{q}'|}\delta^4(p-k-q-q')q\cdot q'.$$

The δ function allow us to compute:

$$(p-k)^2 = q^2 + 2q\cdot q' + {q'}^2 = 2q\cdot q',$$

since neutrinos are massless. So we find

$$a + 4b = \frac{1}{2}\int \frac{d^3\mathbf{q}}{|\mathbf{q}|}\frac{d^3\mathbf{q}'}{|\mathbf{q}'|}\delta^4(p-k-q-q')$$

Now simply do the integral on the RHS. Choose a frame where $\mathbf{p}-\mathbf{k}=\mathbf{0}$ (possible since RHS is a Lorentz scalar). Then $\mathbf{q}=-\mathbf{q}'$ from the δ function, and we're left with

$$\int \frac{d^3\mathbf{q}}{|\mathbf{q}|^2}\delta(p^0-k^0-2|\mathbf{q}|) = 2\pi\int_0^\infty \delta\left(\frac{1}{2}(p^0-k^0)-|\mathbf{q}|\right)d|\mathbf{q}| = 2\pi.$$

Hence $a+4b=\pi$. To get another equation, contract $I_{\mu\nu}$ with $(p-k)^\mu(p-k)^\nu$. Then

$$(p-k)^4(a+b) = \int \frac{d^3\mathbf{q}}{|\mathbf{q}|}\frac{d^3\mathbf{q}'}{|\mathbf{q}'|}\delta^4(p-k-q-q')q\cdot(p-k)q'\cdot(p-k).$$

Again, use the δ function to write

$$(p-k-q)^2 = {q'}^2 = 0 \quad\Rightarrow\quad (p-k)^2 = 2(p-k)\cdot q,$$

and similarly $(p-k)^2 = 2(p-k)\cdot q'$. Then $a+b=\frac{1}{2}\pi$ using the integral we worked out earlier.

Solving our two equation simultaneously, we find that

$$a=\frac{\pi}{3},\quad b=\frac{\pi}{6}.$$

PIECE 2: Substitute $I_{\mu\nu}$ back into the decay rate to find:

$$\Gamma = \frac{G_F^2}{(2\pi)^4 3m_\mu}\int\frac{d^3\mathbf{k}}{k^0}\left(2p\cdot(p-k)k\cdot(p-k)+(p\cdot k)(p-k)^2\right)$$

Now work in the rest frame of μ. So $p=(m_\mu,\mathbf{0})^T$, implying that $p\cdot k = m_\mu E$, where E is the energy of the electron. Note also that $p\cdot p = m_\mu^2$, $k\cdot k = m_e^2$.

We make one final approximation: since $m_e/m_\mu \approx 0.0048$, we neglect the mass of the electron: $k\cdot k=0$. Then the whole thing reduces to $\Gamma =$

$$\frac{m_\mu G_F^2}{3(2\pi)^4}\int d^3\mathbf{k}\,(3m_\mu-4E) = \frac{4\pi m_\mu G_F^2}{3(2\pi)^4}\int_0^{\frac{1}{2}m_\mu} dE\,E^2(3m_\mu-4E).$$

Why the limits? We need to consider maximum and minimum electron energies. Since $E=|\mathbf{k}|$, the energy is a minimum when the electron is at rest, $E=0$.

The maximum energy occurs when ν_μ, $\overline{\nu}_e$ are in the same direction, opposite to the electron; by conservation of momentum, the electron then has its largest possible momentum, and hence largest energy. In this scenario, by energy conservation:

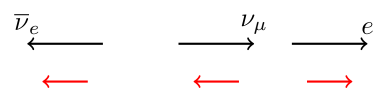
$$E + (E_{\overline{\nu}_e}+E_{\nu_\mu}) = m_\mu,$$

and by momentum conservation, $E-(E_{\overline{\nu}_e}+E_{\nu_\mu})=0$. Hence $E=m_\mu/2$ at its maximum.

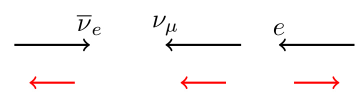
Performing the final, simple integral, we have

$$\Gamma = \frac{G_F^2 m_\mu^5}{192\pi^3}.$$

From our earlier discussion, we know the final state when everything is aligned in the z-direction looks like:



But under a parity transformation, the particles change direction, but the spins do not (spin is an axial vector). Thus we get:



Since neutrinos are massless, helicity is the same as chirality for them. So they must be right-handed! Contradiction, as weak interaction only couples left-handed particles. Thus *muon decay violates parity symmetry*.

5.6 Polarised muon decay

Example: It's possible to repeat the above calculation with a *polarised muon*. We can represent the polarisation with a spin 4-vector, $s^\mu=(0,\mathbf{s})$ (in the muon rest frame), obeying

$$s\cdot p = 0,\quad u_\mu(p)\overline{u}_\mu(p) = (\not{p}+m_\mu)\cdot\frac{1}{2}(1+\gamma^5\not{s}).$$

Following a very similar calculation to muon decay above, we arrive at the formula

$$\sum_{\text{spins}}|\mathcal{M}|^2 = 64G_F^2[q'\cdot k][q\cdot(p-m_\mu s)].$$

Note, it is useful to define $r=p-m_\mu s$ throughout this example. Using very similar steps to the above (i.e. defining a suitable $I_{\mu\nu}$ and using Lorentz covariance to determine its form, then contracting to find unknown constants), we find that the differential decay rate is given by

$$d\Gamma = \frac{G_F^2 m_\mu^5}{24(2\pi)^4}x^2(3-2x-(2x-1)\hat{\mathbf{k}}\cdot\mathbf{s})\,dx\,d\Omega(\hat{\mathbf{k}}),$$

where $x=2E/m_\mu$, E is the energy of the electron, and x may range from 0 to 1.

5.7 Pion decay

Definition: A *pion* is one of a family of three composite particles made of a *valence* quark and valence anti-quark. The possibilities are:

$\pi^- = \overline{u}d, \quad \pi^+ = \overline{d}u, \quad \pi^0 =$ a superposition of $\overline{u}u$ and $\overline{d}d$.

Pions are bound together by *gluons* (see QCD later). We use the term *valence* quarks to distinguish from the virtual quark-anti quark pairs popping in and out of existence in the *sea* of hadrons (again, see QCD).

Consider the process $\pi^-(p) \to e^-(k)\overline{\nu}_e(q)$, i.e. a pion decaying into an electron and an anti-neutrino.

This time, the interactions are governed by both the leptonic part of the charged weak current

$$J^\alpha_{\text{lept}} = \overline{\nu}_e\gamma^\alpha(1-\gamma^5)e + \text{(irrelevant stuff)}$$

and the hadronic (quark) part of the charged weak current:

$$\begin{aligned} J^\alpha_{\text{had}} &= \overline{u}\gamma^\alpha(1-\gamma^5)(V_{ud}d + V_{us}s + V_{ub}b) + \text{(irrelevant stuff)} \\ &= \overline{u}\gamma^\alpha(1-\gamma^5)V_{ud}d + \text{(irrelevant stuff)} \end{aligned}$$

Here, V_{ud} comes from the V_{CKM} matrix which governs the coupling of different generations of quarks. We normally write $J^\alpha_{\text{had}} = V^\alpha_{\text{had}} - A^\alpha_{\text{had}}$, where

$$V^\alpha_{\text{had}} = \overline{u}\gamma^\alpha V_{ud}d, \quad A^\alpha_{\text{had}} = \overline{u}\gamma^\alpha\gamma^5 V_{ud}d.$$

The first piece is called the *vector-like hadronic current* and the second part is called the *axial hadronic current* for obvious reasons.

In Fermi effective theory then, the amplitude for the process is

$$\begin{aligned} \mathcal{M} &= \langle e^-(k)\overline{\nu}_e(q)|\mathcal{L}^{\text{eff}}_W(0)|\pi^-(p)\rangle \\ &= -\frac{G_F}{\sqrt{2}}\langle e^-(k)\overline{\nu}_e(q)|\overline{e}(0)\gamma_\alpha(1-\gamma^5)\nu_e(0)|0\rangle\langle 0|J^\alpha_{\text{had}}(0)|\pi^-(p)\rangle \\ &= -\frac{G_F}{\sqrt{2}}(\overline{u}_e(k)\gamma_\alpha(1-\gamma^5)v_{\overline{\nu}_e}(q))\langle 0|V^\alpha_{\text{had}}(0) - A^\alpha_{\text{had}}(0)|\pi^-(p)\rangle . \end{aligned}$$

Note that we *cannot expand the second matrix element*. This is because the pion is a *bound QCD state*, and cannot be expanded perturbatively, since QCD is a strongly-coupled theory.

To skirt round the issue, we define:

Definition: The *pion decay constant*, F_π, is defined by $\langle 0|A^\alpha_{\text{had}}(0)|\pi^-(p)\rangle = \langle 0|V_{ud}\overline{u}\gamma^\alpha\gamma^5 d|\pi^-(p)\rangle = i\sqrt{2}F_\pi p^\alpha$. This definition is allowed by Lorentz covariance.

Do we need another constant for the vector-like piece? No, it is excluded on the grounds of parity. By Lorentz covariance, the only thing we can have on the RHS for the vector-like piece is

$$\langle 0|\overline{u}\gamma^\alpha d|\pi^-(p)\rangle = Ap^\alpha.$$

Experimentally, we know that the pion is a *pseudoscalar*, i.e. it has intrinsic parity -1 and spin zero. Under a parity transformation then, we then have:

$$\begin{aligned} Ap^\alpha &= \langle 0|\overline{u}\gamma^\alpha d|\pi^-(p)\rangle = \langle 0|\hat{P}^{-1}\hat{P}\overline{u}\gamma^\alpha d\hat{P}^{-1}\hat{P}|\pi^-(p)\rangle \\ &= -\langle 0|\mathbb{P}^\alpha{}_\beta\overline{u}\gamma^\beta d|\pi^-(p_P)\rangle = -\mathbb{P}^\alpha{}_\beta Ap^\alpha = -Ap^\alpha_P. \end{aligned}$$

We get a contradiction unless $A = 0$, so we can just ignore the vector-like piece. Thus the amplitude is

$$\mathcal{M} = iG_F F_\pi \overline{u}_e(k)\not{p}(1-\gamma^5)v_{\overline{\nu}_e}(q)$$

We're ready to calculate now. Begin by pre-emptively applying 4-momentum conservation $p = k+q$. Then $\not{p} = \not{k}+\not{q}$, and recalling the Dirac equation for plane wave spinors: $\overline{u}_e(k)\not{k} = m_e\overline{u}_e(k)$, $\not{q}v_{\overline{\nu}_e}(q) = 0$ (since neutrino massless), the amplitude simplifies to:

$$\mathcal{M} = iG_F F_\pi m_e\overline{u}_e(k)(1-\gamma^5)v_{\overline{\nu}_e}(q)$$

Since the pion has spin zero, after some calculation we find the probability of decay is

$$\sum_{\text{spins}}|\mathcal{M}|^2 = 8|G_F F_\pi m_e V_{ud}|^2(k\cdot q).$$

The best thing to do now is impose momentum conservation early, via $m_\pi^2 = p^2 = (k+q)^2 = k^2 + 2k\cdot q = m_e^2 + 2k\cdot q$. Then we can write $k\cdot q$ in terms of constant quantities, and pull it out of any integrals.

Doing the remaining integrals in the decay rate calculation in the standard way, we find

$$\Gamma = \frac{|G_F F_\pi m_e V_{ud}|^2}{\pi m_\pi}\left(\frac{m_\pi^2 - m_e^2}{2m_\pi}\right)^2.$$

An identical calculation for $\pi \to \mu\overline{\nu}_\mu$ decay gives

$$\Gamma = \frac{|G_F F_\pi m_\mu V_{ud}|^2}{\pi m_\pi}\left(\frac{m_\pi^2 - m_\mu^2}{2m_\pi}\right)^2.$$

Dividing the two results, we find

$$r = \frac{\Gamma(\pi^- \to e^-\overline{\nu}_e)}{\Gamma(\pi^- \to \mu^-\overline{\nu}_\mu)} = \frac{m_e^2}{m_\mu^2}\left(\frac{m_\pi^2 - m_e^2}{m_\pi^2 - m_\mu^2}\right)^2 \approx 1.28\times 10^{-4}.$$

This is very small. This shows that helicity is much less suppressed for the muon because of its greater mass.

5.8 Kaon decay

Definition: A *kaon* is a composite particle containing a strange valence quark or strange valence antiquark, together with one other valence quark. The lightest kaons are: $K^0 = \overline{s}d, \quad \overline{K}^0 = \overline{d}s, \quad K^+ = \overline{s}u, \quad K^- = \overline{u}s.$

Like pions, kaons are *pseudoscalar particles*, i.e. they have intrinsic parity -1 and are spinless.

Example: Consider the decay $K^-(p) \to \mu^-(k)\overline{\nu}_\mu(q)$. This time, the relevant parts of the currents are

$$J^\alpha_{\text{lept}} = \overline{\nu}_\mu \gamma^\alpha (1-\gamma^5)\mu,$$
$$J^\alpha_{\text{had}} = V_{us}\overline{u}\gamma^\alpha(1-\gamma^5)s = V^\alpha_{\text{had}} - A^\alpha_{\text{had}},$$

where $V^\alpha_{\text{had}} = V_{us}\overline{u}\gamma^\alpha s$, $A^\alpha_{\text{had}} = V_{us}\overline{u}\gamma^\alpha\gamma^5 s$. In exactly the same way as π^- decay, we have that K^- is a pseudoscalar, so only A^α_{had} is relevant to the decay. As per usual, the amplitude is given by:

$$\begin{aligned}\mathcal{M} &= \langle \mu^-(p)\overline{\nu}_\mu(q)|\mathcal{L}^{\text{eff}}_W(0)|K^-(p)\rangle \\ &= \frac{G_F}{\sqrt{2}}\langle \mu^-(p)\overline{\nu}_\mu(q)|\overline{\mu}\gamma^\alpha(1-\gamma^5)\nu_\mu|0\rangle \langle 0|A_\alpha(0)|K^-(p)\rangle .\end{aligned}$$

Again, to skirt round the issue of QCD, we define:

Definition: The *kaon decay constant* F_K is defined by
$$\langle 0|A_\alpha(0)|K^-(p)\rangle = iV_{us}\sqrt{2}F_K p_\alpha.$$

Thus we're left with

$$\mathcal{M} = iG_F V_{us}F_K\overline{u}_\mu(k)\not{p}(1-\gamma^5)v_{\overline{\nu}_\mu}(q).$$

Again, pre-emptively using $p = k+q$, we can reduce this to:
$$\mathcal{M} = iG_F V_{us}F_K m_\mu \overline{u}_\mu(k)(1-\gamma^5)v_{\overline{\nu}_\mu}(q).$$

Now calculating as usual, we find the decay rate:

$$\Gamma = \frac{G_F^2|F_K|^2\sin^2(\theta_C)}{4\pi}m_\mu^2 m_K\left(1-\frac{m_\mu^2}{m_K^2}\right)^2,$$

where $V_{us} = \sin(\theta_C)$, and θ_C is the Cabibbo angle (assuming we are working at sufficiently low energies such that top and bottom quarks may be neglected).

5.9 Neutral kaon mixing

The *neutral kaons* are K^0 and $\bar{K}^0$. Neutral kaon mixing is important because it gives empirical evidence of CP violation in the Standard Model.

Since K^0 and $\bar{K}^0$ are a particle-antiparticle pair, they are $\hat{C}$ conjugates. They are also pseudoscalars so both have intrinsic parity -1. Thus we can arrange for the phases to be such that

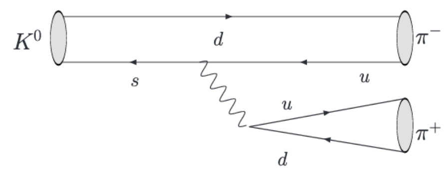
$$\hat{C}\hat{P}\,|K^0\rangle = -\,|\bar{K}^0\rangle\,, \quad \hat{C}\hat{P}\,|\bar{K}^0\rangle = -\,|K^0\rangle\,.$$

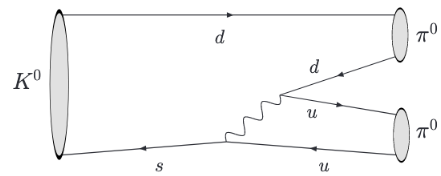
From these equations, it's clear we can construct CP eigenstates:

$$|K^0_+\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\bar{K}^0\rangle), \quad |K^0_-\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K}^0\rangle),$$

where $\hat{C}\hat{P}\,|K^0_+\rangle = |K^0_+\rangle$ and $\hat{C}\hat{P}\,|K^0_-\rangle = -\,|K^0_-\rangle$.

Consider the weak decays $K^0 \to \pi^0\pi^0$ and $K^0 \to \pi^+\pi^-$ given by the Feynman diagrams:





Starting in the kaon's rest frame, there is zero angular momentum before the decay, so by conservation of angular momentum, both $\pi^0\pi^0$ and $\pi^+\pi^-$ must have zero angular momentum. In particular, their relative orbital angular momentum L must be zero: $L=0$.

Therefore:

$$\hat{C}\hat{P}\,|\pi^+\pi^-\rangle = \hat{C}\underbrace{(-1)^2}_{\substack{\text{intrinisic}\\ \text{parity}}}(-1)^L\,|\pi^+\pi^-\rangle = |\pi^-\pi^+\rangle = |\pi^+\pi^-\rangle\,.$$

Similarly, $\hat{C}\hat{P}\,|\pi^0\pi^0\rangle = |\pi^0\pi^0\rangle$. So the final states are both CP-even. Thus if CP is conserved in this interaction, we expect $|K^0_+\rangle$ to be able to decay into 2 pions, but $|K^0_-\rangle$ should not be able to do this. However, $|K^0_-\rangle$ *could* decay into 3, or more, pions.

Thus we expect $|K^0_+\rangle$ to be short-lived (since there is a larger phase space to decay into), and $|K^0_-\rangle$ to be long-lived (smaller decay phase space available).

Experimentally, we find there are two species of neutral kaon: K^0_S, which is short-lived, and K^0_L, which is long-lived. We find that

$$\eta_{+-} = \frac{|\langle \pi^+\pi^-|H|K^0_L\rangle|}{|\langle \pi^+\pi^-|H|K^0_S\rangle|}, \quad \eta_{00} = \frac{|\langle \pi^0\pi^0|H|K^0_L\rangle|}{|\langle \pi^0\pi^0|H|K^0_S\rangle|},$$

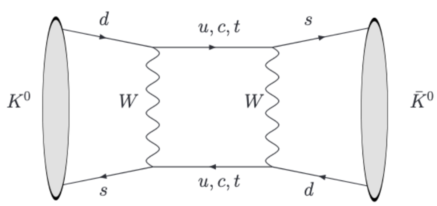
have experimental values $\eta_{+-} = \eta_{00} \approx 2.2\times 10^{-3} \neq 0$, and so we conclude that the weak interaction violates CP symmetry. In particular, K^0_L decays to *both* $\pi^+\pi^-$ and $\pi^0\pi^0$.

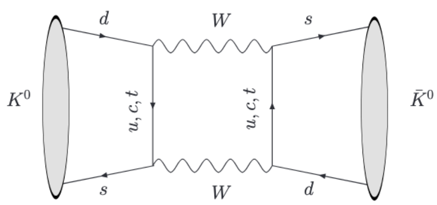
There are two ways in which this can occur here:

1. **Direct CP violation.** There is a complex phase in V_{CKM}, violating CP symmetry of the s, u interactions (see way earlier in the course).

2. **Indirect CP violation.** There is CP violation due to K^0, $\bar{K}^0$ *mixing* (this ultimately comes from a complex phase in V_{CKM} too). It turns out this effect is mainly responsible for the violation we see experimentally.

There is no tree-level mixing of K^0, $\bar{K}^0$. The mixing comes primarily from the loop diagrams:





Therefore, we get small quantum loop corrections to the short/long-lived states, which to a first approximation were CP eigenstates:

$$|K_S^0\rangle = \underbrace{\frac{1}{\sqrt{1+|\epsilon_1|^2}}}_{\text{for normalisation}} \left(|K_+^0\rangle + \epsilon_1 |K_-^0\rangle\right) \approx |K_+^0\rangle\,,$$

$$|K_L^0\rangle = \frac{1}{\sqrt{1+|\epsilon_2|^2}} \left(|K_-^0\rangle + \epsilon_2 |K_+^0\rangle\right) \approx |K_-^0\rangle\,.$$

In general it is hard to compute ϵ_1, ϵ_2 (see AQFT). Instead, here we make:

Definition: The *Wigner-Weisskopf approximation* assumes that

(i) $|K_S^0\rangle$ and $|K_L^0\rangle$ are linear combinations of $|K_+^0\rangle$ and $|K_-^0\rangle$ alone (and not of any excited states);

(ii) we can ignore details of the strong interaction in considering the mixing.

Then, we can assume that as they propagate the states have the form:

$$|K_{S/L}^0\rangle = a_{S/L}(t)\,|K^0\rangle + b_{S/L}(t)\,|\bar{K}^0\rangle\,.$$

Making this assumption, the Schrödinger equation gives us:

$$i\frac{d}{dt}\begin{pmatrix} a_i(t) \\ b_i(t)\end{pmatrix} = \underbrace{\begin{pmatrix} \langle K^0|H'|K^0\rangle & \langle K^0|H'|\bar{K}^0\rangle \\ \langle \bar{K}^0|H'|K^0\rangle & \langle \bar{K}^0|H'|\bar{K}^0\rangle \end{pmatrix}}_{R} \begin{pmatrix} a_i(t) \\ b_i(t)\end{pmatrix}.$$

Here, $i = S, L$ and H' is the next-to-leading order weak Hamiltonian. The off-diagonal elements in R are responsible for the mixing.

Since the kaons are decaying, the amplitudes $a_i(t)$, $b_i(t)$ are not conserved, and hence R is not Hermitian. Write

$$R = M - \frac{i}{2}\Gamma,$$

where M is the *mass matrix* and Γ is the *decay matrix*, both of which are Hermitian.

We can use the CPT Theorem to find relationships between the matrix elements, and to find ϵ_1, ϵ_2 in terms of the R matrix:

Theorem: $R_{11} = R_{22}$.

Proof: Let $\hat{\Theta} = \hat{C}\hat{P}\hat{T}$. Since CPT is a good symmetry, $\hat{\Theta}H'\hat{\Theta}^{-1} = H'^{\dagger}$ (since time-reversal acts to change direction of time in $U(t) = e^{iHt}$; this is achieved by taking the $\dagger$).

In the rest frame of the kaons, we must have $\hat{T}\,|K^0\rangle = |K^0\rangle$ and $\hat{T}\,|\bar{K}^0\rangle = |\bar{K}^0\rangle$, and so recalling their CP transformations, we have $\hat{\Theta}\,|K^0\rangle = -\,|\bar{K}^0\rangle$ and $\hat{\Theta}\,|\bar{K}^0\rangle = -\,|K^0\rangle$.

Therefore, we can write R_{11} as:

$$R_{11} = (K^0, H'K^0) = (\hat{\Theta}^{-1}\hat{\Theta}K^0, H'\hat{\Theta}^{-1}\hat{\Theta}K^0)$$
$$= (\hat{\Theta}^{-1}\bar{K}^0, H'\hat{\Theta}^{-1}\bar{K}^0) = (\bar{K}^0, \hat{\Theta}H'\hat{\Theta}^{-1}\bar{K}^0)^* = (\bar{K}^0, H'^{\dagger}\bar{K}^0),$$

where in the last step, we used the fact that $\hat{\Theta}$ is anti-unitary (since it contains two unitary operators and one anti-unitary operator). Now use conjugate symmetry of the inner product:

$$R_{11} = (H'^{\dagger}\bar{K}^0, \bar{K}^0) = (\bar{K}^0, H'\bar{K}^0) = R_{22},$$

and we're done. □

Theorem: If CP symmetry is respected, then $R_{12} = R_{21}$.

Proof: By CPT, T must be a good symmetry if CP is. Hence $\hat{T}H'\hat{T}^{-1} = H'^{\dagger}$.

Now do exactly the same calculation as above Theorem:

$$R_{12} = (K^0, H'\bar{K}^0) = (\hat{T}^{-1}K^0, H'\hat{T}^{-1}\bar{K}^0)$$
$$= (K^0, H'^{\dagger}\bar{K}^0)^* = (H'^{\dagger}\bar{K}^0, K^0) = (\bar{K}^0, H'K^0) = R_{21}. \;\square$$

Theorem: $\epsilon_1 = \epsilon_2 = \epsilon$, where

$$\epsilon = \frac{\sqrt{R_{12}} - \sqrt{R_{21}}}{\sqrt{R_{12}} + \sqrt{R_{21}}}.$$

Proof: The independent solutions of the Wigner-Weisskopf approximation equation are of the form $e^{-\lambda t}\,|v\rangle$, where $|v\rangle$

is an eigenvector of R. By construction though, at time $t = 0$, the solution of the S equation is $|K_S^0\rangle$ and the solution of the L equation is $|K_L^0\rangle$. Hence these must be eigenvectors of R.
Writing $|K_S^0\rangle$ and $|K_L^0\rangle$ in the $|K^0\rangle$ and $|\bar{K}^0\rangle$ basis, we have:

$$|K_S^0\rangle \propto \begin{pmatrix} 1+\epsilon_1 \\ -1+\epsilon_1 \end{pmatrix}, \quad |K_L^0\rangle \propto \begin{pmatrix} 1+\epsilon_2 \\ 1-\epsilon_2 \end{pmatrix}.$$

Hence these are eigenvectors of R.

Using this fact, we have

$$\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{11} \end{pmatrix} \begin{pmatrix} 1+\epsilon_1 \\ -1+\epsilon_1 \end{pmatrix} = \mu \begin{pmatrix} 1+\epsilon_1 \\ -1+\epsilon_1 \end{pmatrix},$$

where μ is some evalue. Separating this into two equations, and dividing to get rid of μ, we have an equation for ϵ_1. Recalling $\epsilon_1 < 1$, we can solve this equation to get

$$\epsilon_1 = \frac{\sqrt{R_{12}} - \sqrt{R_{21}}}{\sqrt{R_{12}} + \sqrt{R_{21}}},$$

as desired. In exactly the same way, we find ϵ_2, which is equal to ϵ_1. □

In particular, the final two Theorems show that if CP is respected, then $\epsilon_1 = \epsilon_2 = 0$, i.e. there is *no* mixing. This contradicts the experimental observations we noted earlier.

---

6 Quantum chromodynamics

6.1 The QCD Lagrangian

Definition: *Quantum chromodynamics* (QCD) is an $SU(3)$ gauge theory, where $SU(3)$ is the called *colour* symmetry. The gauge bosons of the theory are called *gluons*. The Lagrangian is given by:

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}F^{a\mu\nu}F_{a\mu\nu} + \sum_f \overline{f}(i\not{D} - m_f)f,$$

where f denotes all possible quark flavours: u, d, s, c, t and b, and m_f denotes their mass. The quarks here are in the fundamental representation of $SU(3)$, i.e.

$$f = \begin{pmatrix} f_{\text{red}} \\ f_{\text{green}} \\ f_{\text{blue}} \end{pmatrix}.$$

The covariant derivative here is $D_\mu = \partial_\mu + igA_\mu^a T^a$, where the iT^a are the generators of the Lie algebra of $SU(3)$. We may choose

$$T^a = \frac{1}{2}\lambda^a,$$

where the λ^a are called the *Gell-Mann matrices* (of which there are 8, and hence there are 8 gluon fields).

Note that $SU(3)$ is *not* spontaneously broken in the Standard Model by the Higgs mechanism. Gluons are therefore massless particles, just like photons.

Also note that the mass term for the quarks here comes from their coupling to the Higgs boson; we saw this in the electroweak theory.

6.2 Renormalisation of QCD

The parameters in the Lagrangian are *not* the physical parameters we observe in experiments.

Definition: We relate the Lagrangian's couplings to physical couplings by *renormalisation conditions*: $g_i^0 = G_i^0(\{g_i(\mu)\}, \mu)$, where μ is an energy scale.

The renormalised couplings depend on μ via a *beta function*:

$$\beta_j(\{g_j(\mu)\}, \mu) = \mu\frac{d}{d\mu}(g_j(\mu)).$$

Theorem: It can be shown that the beta function for the coupling in a non-Abelian gauge theory is:

$$\beta(g) = -\beta_0\frac{g^3}{16\pi^2} + O(g^5),$$

where

$$\beta_0 = \frac{11}{3}c - \frac{4}{3}\sum_f T_f.$$

Here, $c\delta^{ab} = f^{acd}f^{bcd}$, where the f^{abc}'s are the structure constants of the rep (note $c = N$ for $SU(N)$) and $T_f\delta^{ab} = \text{Tr}(t_f^a t_f^b)$, where the it_f^a are the generators of the Lie algebra rep for the particle f (for a fermion in the fundamental rep, $T_f = \frac{1}{2}$, and for a scalar particle in the trivial rep, $T_f = 0$).

Proof: Beyond scope of course. □

In our case, this expression gives us:

$$\beta_0 = 11 - \frac{2}{3}N_f,$$

where N_f is the number of active quark flavours. We note $\beta_0 > 0$ for $N_f \leq 16$ (depending on energy scale, more quarks become activated, so N_f can vary between 1 and 6 depending on the problem we are solving).

Definition: The standard coupling used in QCD is the *strong coupling*, given by $\alpha_s = g^2/4\pi$ in terms of the coupling g above for a generic non-Abelian gauge theory.

Theorem: We have:

$$\alpha_S(\mu) = \frac{2\pi}{\beta_0\log(\mu/\Lambda_{\text{QCD}})},$$

where Λ_{QCD} is the energy scale at which α_S diverges.

Proof: Using the beta function, we have

$$\mu \frac{d\alpha_S}{d\mu} = \frac{d\alpha_S}{d\log(\mu)} = -\frac{\beta_0}{2\pi}\alpha_S^2 + \cdots .$$

Therefore, integrating this equation we get:

$$\alpha_S(\mu) = \frac{2\pi}{\beta_0} \cdot \frac{1}{\log(\mu/\mu_0) + 2\pi/\beta_0 \alpha_S(\mu_0)}.$$

This diverges at Λ_{QCD}, hence:

$$\log(\Lambda_{\text{QCD}}) = \log(\mu_0) - \frac{2\pi}{\beta_0 \alpha_S(\mu_0)},$$

and the result follows. □

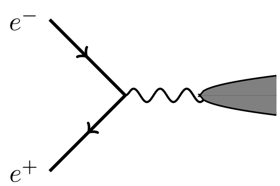
Notice that for β_0 the strong coupling *decreases* for increasing μ. This property is called *asymptotic freedom*. That is, the quarks are less strongly bound at higher energies, but are more strongly bound at lower energies. This means that we can't use perturbation theory at the lower energies we are interested in.

Free quarks are never seen at low energies, a property called *confinement*.

6.3 $e^+e^- \to$ hadrons

Example: Consider annihilation of an electron and positron to give quarks. We know that quarks can never be seen alone (at low energies), and hence they must turn into hadrons, a process called *hadronisation*.

The Feynman diagram for the process is:



Here, we get some non-perturbative mess after the electrons interact.

The first vertex is just a QED vertex so has an easy contribution. The second vertex is a coupling of quarks to a photon, so we must use the hadronic electromagnetic current from the electroweak theory, which recall has the form:

$$J_h^\mu = \sum_f q_f \overline{f} \gamma^\mu f,$$

where f sums over all *quark flavours*. If the final state is X, the amplitude is then (where $q = p_1 + p_2$):

$$iA = \frac{(-ie)^2 i}{q^2} \langle X | J_h^\mu | 0 \rangle \, \overline{v}_e(p_2) \gamma_\mu u_e(p_1).$$

The cross-section (inclusive of all possible hadrons X) is then:

$$\sigma(e^+e^- \to \text{hadrons}) = \frac{1}{8p_1^0 p_2^0} \frac{1}{4} \sum_X \sum_{\text{spins}} (2\pi)^4 \delta^4(q - p_X) |A|^2,$$

where the sum over X means:

$$\sum_X = \sum_{\substack{\text{possible} \\ \text{hadrons } X}} \int \frac{d^4 p_X}{2p_X^0},$$

where p_X is the momentum of the hadron species X.

In order to deal with $|A|^2$ non-perturbatively, we introduce the *hadronic spectral density function*:

Definition: The *hadronic spectral density* is defined by

$$\rho_h^{\mu\nu}(q) = (2\pi)^3 \sum_X \delta^4(q - p_X) \langle 0 | J_h^\mu | X \rangle \langle X | J_h^\nu | 0 \rangle .$$

Theorem: We can write $\rho_h^{\mu\nu}(q)$ in the form:

$$\rho_h^{\mu\nu}(q) = (-\eta^{\mu\nu} q^2 + q^\mu q^\nu) \theta(q^0) \rho_h(q^2),$$

where $\rho_h(q^2)$ is a scalar function and θ is the Heaviside step function.

Proof: Note $\rho_h^{\mu\nu}$ is symmetric on μ, ν, so must be a linear function of $\eta^{\mu\nu}$ and $q^\mu q^\nu$ be Lorentz covariance. Furthermore, we can use the *Ward identity* (proved in AQFT) which gives $q_\mu \rho_h^{\mu\nu} = q_\nu \rho_h^{\mu\nu} = 0$. Finally, since states labelled by X have positive energy, $\rho_h^{\mu\nu}$ should vanish for $q^0 < 0$. The result follows. □

Using this in the cross-section formula, we find that

$$\sigma(e^+e^- \to \text{hadrons}) = \frac{16\pi^3 \alpha^2}{q^2} \rho_h((p_1 + p_2)^2),$$

where α is the fine structure constant, $e^2/4\pi$.

Is it possible to actually calculate this function? No. However, we can come up with models that allow us to approximate it.

For example, we might consider X to actually be a quark, anti-quark or gluon (these are called *partons* in this context). Then we assume:

$$\sum_X = \sum_{X = f, \overline{f}, g}$$

This allows us to actually compute $\rho_h^{\mu\nu}$ and in turn ρ_h.

6.4 τ decay into hadrons

Example: Consider τ decay into hadrons and a τ neutrino via the weak hadronic current term:

$$\mathcal{L}_W = -\frac{G_F}{\sqrt{2}} J^{h\dagger}_\alpha \overline{\nu}_\tau \gamma^\alpha (1-\gamma_5)\tau.$$

Again, we can define spectral density functions via:

$$\sum_X (2\pi)^3 \delta^4(P_X - k) \langle 0|J^h_\alpha|X\rangle \langle X|J^{h\dagger}_\beta|0\rangle$$
$$= k_\alpha k_\beta \rho_0(k^2) + (-\eta_{\alpha\beta}k^2 + k_\alpha k_\beta)\rho_1(k^2).$$

We need two spectral density functions because we don't necessarily have the Ward identity for this current (it applies only to electromagnetic currents, coupled to the photon).

Computing the decay rate, we get (after a lengthy calculation):

$$\Gamma = \frac{G_F^2 m_\tau^3}{16\pi} \int_0^{m_\tau^2} d\sigma \left(1 - \frac{\sigma}{m_\tau^2}\right)^2 \left(\rho_0(\sigma) + \left(1 + \frac{2\sigma}{m_\tau^2}\rho_1(\sigma)\right)\right).$$

If we somehow *knew* that the final hadron would be a *pion*, we could find the spectral density functions. Recalling that $\langle 0|J^h_\alpha|\pi\rangle = V_{ud} i\sqrt{2}F_\pi p_\alpha$, where F_π is the pion decay constant, we can find that

$$\rho_0(\sigma) = 2F_\pi^2 \cos^2(\theta_C)\delta(\sigma - m_\pi^2), \quad \rho_1(\sigma) = 0,$$

where θ_C is the Cabibbo angle. Thus the final decay rate for τ to a pion is:

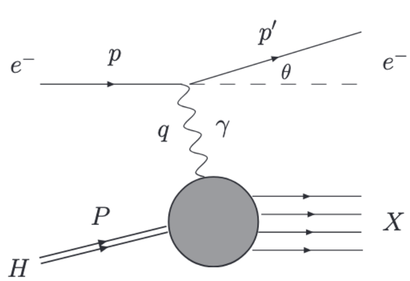
$$\Gamma = \frac{G_F^2 m_\tau^3}{8\pi} F_\pi^2 \cos^2(\theta_C)\left(1 - \frac{m_\pi^2}{m_\tau^2}\right)^2.$$

6.5 Deep inelastic scattering

Consider electron-proton scattering via exchange of a photon in the *deep inelastic scattering regime*.

Definition: *Deep* refers to a high energy process. *Inelastic* means that the electron's energy can change during the scattering process.

In general, a Feynman diagram for the process looks like:



Kinematics: First deal with kinematics of problem. We have:

Definition: Define the *scattering angle* by $\mathbf{p}\cdot\mathbf{p}' = |\mathbf{p}||\mathbf{p}'|\cos(\theta)$. Write also the electron energies as $E = p^0$ and $E' = p'^0$.

Since the electrons have mass negligible to the proton, treat them as massless.

Definition: Define $Q^2 = -q^2$, where q is the photon momentum. Define $\nu = p_H \cdot q$, where p_H is the proton momentum. Finally, define the dimensionless *Bjorken quantities* by

$$x = \frac{Q^2}{2\nu}, \qquad y = \frac{\nu}{p_H \cdot p}.$$

Theorem (Kinematics): We have (i) $Q^2 \geq 0$; (ii) $Q^2 \leq 2\nu$; (iii) $0 \leq x \leq 1$ and (iv) $0 \leq y \leq 1$.

Proof: (i) Note that $q = p - p'$ so that $q^2 = 2p\cdot p' = 2EE'(\cos(\theta) - 1) \leq 0$, and therefore $Q^2 \geq 0$.

(ii) In the rest frame of the proton, $\nu = M(E - E') \geq 0$, where M is the proton's mass. Now by conservation of momentum, the momentum of the final state X is:

$$p_X = p_H + q \Rightarrow M_X^2 = M^2 + 2\nu - Q^2. \qquad (*)$$

Now note that by 3-momentum conservation, $\mathbf{q} = \mathbf{p}_X$, and also by energy conservation, we have $\sqrt{M_X^2 + |\mathbf{p}_X|^2} = |\mathbf{q}| + M$, which on squaring gives: $M_X^2 - M^2 = |\mathbf{q}|^2 - |\mathbf{p}_X|^2 + 2M|\mathbf{q}| = 2M|\mathbf{q}| \geq 0$. Therefore $M_X^2 \geq M^2$ and (ii) then follows from $(*)$.

(iii) follows immediately by (i) and (ii).

(iv) In the proton's rest frame,

$$y = \frac{\nu}{ME}.$$

which is clearly greater than 0. Also, recall $\nu = M(E - E')$, so

$$y = 1 - \frac{E'}{E} \leq 1.$$

and we're done. $\square$

Definition: The *deep inelastic limit* of the problem is to take $Q^2 \to \infty$, $\nu \to \infty$ (i.e. high energies) but keep x and y finite.

Let us now go ahead and calculate the cross section. The amplitude is easily seen to be (by the Feynman rules):

$$i\mathcal{M} = (-ie)^2 \overline{u}_e(p')\gamma^\mu u_e(p)\left(\frac{-i\eta_{\mu\nu}}{q^2}\right)\langle X|J_h^\nu|H(p_H)\rangle .$$

Working in the *rest frame of the proton*, the flux factor is $|\mathbf{v}_e - \mathbf{v}_H| = 1$, since electrons are massless in this regime and the proton is stationary. The differential cross section is then:

$$d\sigma = \frac{1}{4ME}\frac{d^3\mathbf{p}'}{(2\pi)^3 2E'}\sum_X (2\pi)^4\delta^4(q+p_H-p_X)\cdot\frac{1}{2}\sum_{\text{spins}}|\mathcal{M}|^2.$$

Write

$$\frac{1}{2}\sum_{\text{spins}}|\mathcal{M}|^2 = \frac{e^4}{2q^4}L_{\mu\nu}\langle H|J_h^{\mu\dagger}|X\rangle\langle X|J_h^\nu|H\rangle ,$$

where $L_{\mu\nu}$ is the contribution from the electrons, given by

$$L_{\mu\nu} = \mathrm{Tr}(\not{p}\gamma_\mu\not{p}'\gamma_\nu) = 4(p_\mu p'_\nu + p'_\mu p_\nu - \eta_{\mu\nu}p\cdot p').$$

Notice that $L_{\mu\nu}$ is symmetric under $\mu \leftrightarrow \nu$.

Also write:

$$W_H^{\mu\nu} = \frac{1}{4\pi}\sum_X (2\pi)^4\delta^4(q+p_H-p_X)\langle H|J_h^{\mu\dagger}|X\rangle\langle X|J_h^\nu|H\rangle .$$

Then

$$\frac{d\sigma}{d^3\mathbf{p}'}\frac{1}{8(2\pi)^2}\cdot\frac{1}{mEE'}\cdot\frac{e^4}{q^4}L_{\mu\nu}W_H^{\mu\nu}.$$

As is now standard, we use an argument based on Lorentz invariance for the form of $W_H^{\mu\nu}$. Notice it is contracted with $L_{\mu\nu}$, so we may as well assume it is symmetric (any antisymmetric part will vanish). We also have the Ward identity $q_\mu W_H^{\mu\nu} = 0$ since it comes from an EM vertex. These restrictions are enough to force:

$$W_H^{\mu\nu} = \left(-\eta^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right)W_1(\nu, Q^2) + \left(p_H^\mu - \frac{p_H\cdot q}{q^2}q^\mu\right)\left(p_H^\nu - \frac{p_H\cdot q}{q^2}q^\nu\right)W_2(\nu, Q^2).$$

for some scalar functions W_1 and W_2.

Notice that $q^\mu L_{\mu\nu} = 0$ by the Ward identity too. Thus we're left with:

$$L_{\mu\nu}W_H^{\mu\nu} = 4Q^2W_1 + 2M^2(4EE' - Q^2)W_2.$$

It follows that

$$\frac{d\sigma}{d^3\mathbf{p}'} = \frac{e^4}{8(2\pi)^2MEE'q^4}(4Q^2W_1 + 2M^2(4EE'-Q^2)W_2).$$

Writing this in terms of the Bjorken x and y, and taking the DIS limit $Q^2\to\infty$ and $\nu\to\infty$, we see that

$$L_{\mu\nu}W_H^{\mu\nu} = 8EM\left(xyW_1 + \frac{1-y}{y}\nu W_2\right)$$

The measure also becomes:

$$d^3\mathbf{p}' = 2\pi(E')^2 d\cos(\theta)dE' = \pi E' dQ^2 dy = 2\pi E'\nu dx dy.$$

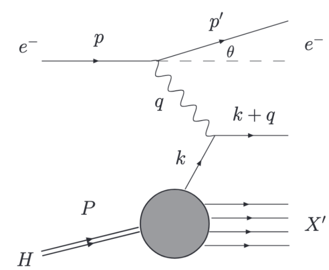
Therefore the final differential cross section is:

$$\frac{d\sigma}{dxdy} = \frac{8\pi\alpha^2 ME}{Q^4}\left(xy^2F_1 + (1-y)F_2\right),$$

where the functions $F_1 = W_1$ and $F_2 = \nu W_2$ are the dimensionless *structure functions* of the proton.

6.6 The parton model

The *parton model* assumes that the photon interacts with a *single* constituent of the proton. The leading order approximation in the parton model is then:



We assume we can write:

$$\sum_X = \sum_{X'}\sum_f \frac{1}{(2\pi)^3}\int d^4\tilde{k}\theta(\tilde{k}^0)\delta(\tilde{k}^2)\sum_{\text{parton spins}} .$$

Here, the Heaviside function forces the parton to have positive energy, and the delta function forces it to be massless.

Inserting this expression in $W_H^{\mu\nu}$, we have

$$W_H^{\mu\nu} = \sum_f\int d^4k\,\mathrm{Tr}(W_f^{\mu\nu}\Gamma_{H,f}(p_H,k) + \overline{W}_f^{\mu\nu}\overline{\Gamma}_{H,f}(p_H,k)),$$

where

$$\Gamma_{H,f}(p_H,k)_{\beta\alpha} = \sum_{X'}\delta^4(p_H-k-p_{X'})\langle H|f_\alpha|X'\rangle\langle X'|f_\beta|H\rangle ,$$

and $\overline{\Gamma}$ is the same but with antiquarks, $\overline{f}$. We can compute the proton structure functions as:

$$F_1(x,Q^2) = \frac{1}{2}\sum_f q_f^2(f(x)+\overline{f}(x)), \quad F_2(x,Q^2) = 2xF_1,$$

where q_f are the charges of the partons (assumed to be quarks here) and $f(x)$ and $\overline{f}(x)$ are dimensionless functions called *parton distribution functions*.