Masses and the Higgs Mechanism

Kevin Cahill for 524

February 21, 2016

1 Masses

Masses occur in the action density as the coefficients of terms quadratic in the fields. Thus, in the case of a neutral, spin-zero field ϕ the action density is

$$L = -\tfrac{1}{2}\partial_a\phi\partial^a\phi - \tfrac{1}{2}\mu^2\phi^2, \tag{1}$$

and the mass is μ. The equation of motion is

$$\left(\partial_a\partial^a - \mu^2\right)\phi(x) = 0. \tag{2}$$

The field obeying this equation is

$$\phi(x) = \int \left[a(k)e^{ikx} + a^\dagger(k)e^{-ikx}\right] \frac{d^3k}{\sqrt{(2\pi)^3 2k^0}}. \tag{3}$$

The charged spin-zero field is a complex linear combination of two equal-mass real fields

$$\phi = \frac{1}{\sqrt{2}}\left(\phi^{(1)} + i\phi^{(2)}\right). \tag{4}$$

Its action density is

$$L = -\partial_a \phi^* \partial^a \phi - \mu^2 |\phi|^2, \tag{5}$$

and its equation of motion is

$$\left(\partial_a \partial^a - \mu^2\right) \phi(x) = 0. \tag{6}$$

It is

$$\phi(x) = \int \left[a(k) e^{ikx} + b^\dagger(k) e^{-ikx}\right] \frac{d^3k}{\sqrt{(2\pi)^3 2k^0}}. \tag{7}$$

For a Dirac spin-one-half field, the action density is

$$L = -\overline{\psi}\left(\gamma^a \partial_a + m\right)\psi \equiv -\overline{\psi}\left(\partial\!\!\!/ + m\right)\psi \tag{8}$$

in which $\overline{\psi} = i\psi^\dagger\gamma^0 = \psi^\dagger\beta$. Weinberg's choice of gamma matrices is

$$\gamma^0 = -i\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \vec{\gamma} = -i\begin{pmatrix}0 & \vec{\sigma}\\ -\vec{\sigma} & 0\end{pmatrix}, \quad \gamma^5 = \gamma_5 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{9}$$

He also uses $\beta = i\gamma^0$ and $\overline{\psi} = \psi^\dagger\beta = i\psi^\dagger\gamma^0$. The Dirac equation of motion is

$$\left(\gamma^a \partial_a + m\right)\psi = \left(\partial\!\!\!/ + m\right)\psi = 0. \tag{10}$$

The field is

$$\begin{aligned}\psi_j(x) &= \frac{1}{\sqrt{2}}\left(\psi_j^{(1)}(x) + i\psi_j^{(2)}(x)\right)\\ &= \sum_{s=-}^{+}\int \left[u_j(\vec{p},s) b(p,s) e^{ipx} + v_j(\vec{p},s) c^\dagger(p,s) e^{-ipx}\right] \frac{d^3p}{(2\pi)^{3/2}}\end{aligned} \tag{11}$$

in which the fields ψ^1 and ψ^2 are the Majorana fields that make the Dirac field, and the Dirac index j runs from 1 to 4, $px = \vec{p}\cdot\vec{x} - p^0 t$, $p^0 = \sqrt{\vec{p}^2 + m^2}$,

$$b(p,s) = \frac{1}{\sqrt{2}}\left(a(p,s,1) + ia(p,s,2)\right) \tag{12}$$

$$c^\dagger(p,s) = \frac{1}{\sqrt{2}}\left(a^\dagger(p,s,1) + ia^\dagger(p,s,2)\right), \tag{13}$$

and the annihilation a_i and creation $a_j^\dagger$ operators satisfy the anticommutation relations

$$\begin{aligned} \{a(p,s,i), a(p',s',j)\} &\equiv a(p,s,i)\, a(p',s',j) + a(p',s',j)\, a(p,s,i) = 0 \\ \{a(p,s,i), a^\dagger(p',s',j)\} &= \delta_{i,j}\, \delta_{s,s'}\, \delta^{(3)}(\vec{p} - \vec{p}'). \end{aligned} \tag{14}$$

For a massive vector field, the action density is

$$L = -\frac{1}{4} F_{ab} F^{ab} - \frac{1}{2} m^2 A_a A^a \tag{15}$$

in which $F_{ab} = \partial_a A_b - \partial_b A_a$. The equation of motion is

$$\partial_a F^{ab}(x) = m^2 A^b(x). \tag{16}$$

This field contains a part that is spin zero. The spin-zero part is the divergence $\partial_b A^b$, and the spin-one part has zero divergence

$$\partial_b A^b = 0. \tag{17}$$

So the equation of motion of the spin-one part is

$$(\Box - m^2)\, A_b(x) = (\triangle - \partial_t^2 - m^2) A_b(x) = 0. \tag{18}$$

The spin-one field is

$$A_b(x) = \sum_{s=-1}^{1} \int \left[e_b(p,s) a(p,s) e^{ipx} + e_b^*(p,s) a^\dagger(p,s) e^{-ipx} \right] \frac{d^3p}{\sqrt{(2\pi)^3 2p^0}} \tag{19}$$

in which the sum is over $s = -1, 0, 1$,

$$p^a e_a(p,s) = 0, \tag{20}$$

and the spin sum is

$$\sum_{s=-1}^{1} e_a(p,s) e_b^*(p,s) = \eta_{ab} + \frac{p_a p_b}{m^2}. \tag{21}$$

Homework problem 1 of set 2: Use Lagrange's equations to derive the equation of motion (16) from the action density (15).

Homework problem 2 of set 2: Use the condition $\partial_b A^b = 0$ to convert the equation of motion (16) to its spin-one form (18).

Homework problem 3 of set 2: Show that the zero-divergence condition (17) implies the spin condition (20).

2 Spontaneous Symmetry Breaking

The usual ϕ^4 theory has action density

$$L = -\frac{1}{2}\partial_a\phi\,\partial^a\phi - \frac{1}{2}m^2\phi^2 - \frac{1}{4}\lambda\phi^4. \tag{22}$$

In the limit $\lambda \to 0$, this theory is that of equations (1– 3). If we flip the sign of the mass term in (23), then we have

$$L = -\frac{1}{2}\partial_a\phi\,\partial^a\phi + \frac{1}{2}m^2\phi^2 - \frac{1}{4}\lambda\phi^4. \tag{23}$$

Now both action densities are symmetric under the reflection $\phi(x) \to -\phi(x)$, which is a **discrete symmetry**.

To the extent that we understand such theories, the vacuum of the first theory has $\langle 0|\phi|0\rangle = 0$. This vacuum is invariant under the reflection $\phi(x) \to -\phi(x)$. But there are two vacua of the second theory. The potential energy of that theory is

$$V = -\frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda\,\phi^4, \tag{24}$$

and it has two minima at

$$\phi_\pm = \pm\frac{m}{\sqrt{\lambda}} \equiv \pm v. \tag{25}$$

These vacua are not invariant under the reflection $\phi(x) \to -\phi(x)$; they transform into each other $\phi_\pm \to \phi_\mp$. So if the states of the universe are clustered about ϕ_+, then in that universe, the mean value of the field in its vacuum is

$$\langle 0_+|\phi|0_+\rangle = \phi_+. \tag{26}$$

We say, "the vacuum spontaneously breaks the reflection symmetry."

If we set

$$\phi(x) = v + \sigma(x), \tag{27}$$

then the action density of the theory near ϕ_+ is

$$\begin{aligned} L = & -\frac{1}{2}\partial_a\sigma\,\partial^a\sigma + \frac{1}{2}m^2\,(v+\sigma)^2 - \frac{1}{4}\lambda(v+\sigma)^4 \\ = & -\frac{1}{2}\partial_a\sigma\,\partial^a\sigma + \frac{1}{2}m^2\,(v^2+2v\sigma+\sigma^2) - \frac{1}{4}\lambda(v^4+4v^3\sigma+6v^2\sigma^2+4v\sigma^3+\sigma^4) \\ = & -\frac{1}{2}\partial_a\sigma\,\partial^a\sigma + \frac{1}{2}m^2\left[\left(\frac{m}{\sqrt{\lambda}}\right)^2 + 2\frac{m}{\sqrt{\lambda}}\sigma + \sigma^2\right] \\ & -\frac{1}{4}\lambda\left[\left(\frac{m}{\sqrt{\lambda}}\right)^4 + 4\left(\frac{m}{\sqrt{\lambda}}\right)^3\sigma + 6\left(\frac{m}{\sqrt{\lambda}}\right)^2\sigma^2 + 4\frac{m}{\sqrt{\lambda}}\sigma^3 + \sigma^4\right] \\ = & -\frac{1}{2}\partial_a\sigma\,\partial^a\sigma - m^2\sigma^2 - \sqrt{\lambda}m\,\sigma^3 - \frac{1}{4}\lambda\,\sigma^4 + \frac{1}{4}\frac{m^4}{\lambda} \end{aligned} \tag{28}$$

in which the last term is a constant (and so is relevant only in gravitational theories where it might represent dark energy). In the limit $\lambda \to 0$, this theory is that of a particle of mass $\sqrt{2}\,m$.

If we replace the single field ϕ by an n-vector of fields ϕ_i, then we get the **linear sigma model** with action density

$$L = -\frac{1}{2}\sum_{i=1}^{n}\partial_a\phi_i\,\partial^a\phi_i + \frac{1}{2}m^2\sum_{i=1}^{n}\phi_i^2 - \frac{1}{4}\lambda\left(\sum_{i=1}^{n}\phi_i^2\right)^2 \tag{29}$$

or with $\phi^2 \equiv \phi_1^2 + \cdots + \phi_n^2$

$$L = -\frac{1}{2}\sum_{i=1}^{n}\partial_a\phi_i\,\partial^a\phi_i + \frac{1}{2}m^2\phi^2 - \frac{1}{4}\lambda\left(\phi^2\right)^2. \tag{30}$$

Again the mass term has the wrong sign. In what follows, we will not bother to indicate sums over a repeated index i from 1 to n.

This L is invariant when the fields change by

$$\phi_i' = O_{ik}\phi_k \tag{31}$$

in which O is an $n \times n$ orthogonal matrix. That is, the squared length

$$\phi_i'^2 = (O_{ik}\phi_k)^2 = \phi_k^2. \tag{32}$$

The action density is invariant under the nonabelian Lie group $O(n)$. This is a continuous symmetry.

What are the minima of the potential energy

$$V = -\frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda\phi^4 \tag{33}$$

in which $\phi^2 = \phi_i^2$? Now they are the points on the sphere

$$\phi^2 = \frac{m^2}{\lambda} \quad \text{of radius} \quad \phi = v = \frac{m}{\sqrt{\lambda}}. \tag{34}$$

Whereas in the discrete case there were two degenerate vacua, here there are infinitely many.

As in the discrete case, we pick one vacuum. We imagine that in the physical vacuum $|0\rangle$ the mean values of the n fields ϕ_i are

$$\langle 0|\phi_i|0\rangle = \frac{m}{\sqrt{\lambda}}\,\delta_{in} = v\,\delta_{in}. \tag{35}$$

Now we write call the components of the field

$$\phi_i = (\pi_1, \pi_2, \ldots, \pi_{n-1}, v + \sigma). \tag{36}$$

So now

$$\phi^2 = \phi_1^2 \ldots \phi_n^2 = \pi_1^2 + \cdots + \pi_{n-1}^2 + (v+\sigma)^2 \equiv \pi^2 + (v+\sigma)^2, \tag{37}$$

and the action density (30) is

$$\begin{aligned} L = & -\frac{1}{2}\partial_a\phi_i\,\partial^a\phi_i + \frac{1}{2}m^2\phi^2 - \frac{1}{4}\lambda\left(\phi^2\right)^2 \\ = & -\frac{1}{2}\partial_a\pi_i\,\partial^a\pi_i - \frac{1}{2}\partial_a\sigma\,\partial^a\sigma + \frac{1}{2}m^2\left[\pi^2 + (v+\sigma)^2\right] - \frac{1}{4}\lambda\left[\pi^2 + (v+\sigma)^2\right]^2 \\ = & -\frac{1}{2}\partial_a\pi_i\,\partial^a\pi_i - \frac{1}{2}\partial_a\sigma\,\partial^a\sigma + \frac{1}{2}m^2\left[\pi^2 + v^2 + 2v\sigma + \sigma^2\right] - \frac{1}{4}\lambda\left[\pi^2 + v^2 + 2v\sigma + \sigma^2\right]^2. \end{aligned} \tag{38}$$

In this expression, $m^2 = \lambda\, v^2$, and so the coefficient of π^2 vanishes while that of σ^2 is $-m^2$

$$L = -\frac{1}{2}\,\partial_a\pi_i\,\partial^a\pi_i - \frac{1}{2}\,\partial_a\sigma\,\partial^a\sigma - m^2\,\sigma^2 - m\sqrt{\lambda}\,\sigma\pi^2 - m\sqrt{\lambda}\sigma^3 - \frac{1}{4}\lambda\,\sigma^4 - \frac{1}{2}\lambda\,\pi^2\sigma^2 + \frac{1}{4}\lambda\,\pi^4 - \frac{1}{4}\frac{m^4}{\lambda}. \tag{39}$$

So the theory describes one field σ of mass $\sqrt{2}m$ and $n-1$ massless fields π_k. These massless fields are called **Goldstone bosons**.

3 Goldstone's Theorem

Let $V(\phi)$ be a potential that is bounded below and that depends upon n fields ϕ_j. Assume that the action density

$$L = -\frac{1}{2}\partial_a\phi_j\partial^a\phi_j - V(\phi) \tag{40}$$

is invariant under the global linear transformation

$$\phi'_j = O_{jk}\phi_k \tag{41}$$

in which the $n \times n$ unitary matrix O is a member of a representation of a continuous Lie group G. Since $V(\phi)$ is bounded below and invariant under the symmetry (41), it has several minima ϕ_0. At these minima, the first-order partial derivatives must vanish

$$\left.\frac{\partial V(\phi)}{\partial\phi_k}\right|_{\phi=\phi_0} = 0 \tag{42}$$

and the mixed second-order partial derivatives must be nonnegative

$$\left.\frac{\partial^2 V(\phi)}{\partial\phi_k \partial\phi_\ell}\right|_{\phi=\phi_0} \equiv m_{k\ell} \geq 0. \tag{43}$$

Near each minimum, the potential looks like

$$V(\phi) = V(\phi_0) + \frac{1}{2}(\phi_k - \phi_{k0})(\phi_\ell - \phi_{\ell 0})\, m^2_{k\ell}. \tag{44}$$

The matrix $m^2_{k\ell}$ is real and symmetric. So it can be diagonalized by an orthogonal transformation. Its eigenvalues are the squares of the masses of the scalar bosons of the theory.

Near the identity, the matrix O looks like

$$O = I + i\theta_r t^r \tag{45}$$

in which the generators t^r are hermitian, $t^\dagger_r = t^r$. Since $V(O\phi) = V(\phi)$, the derivatives with respect to the θ_r's must vanish

$$\frac{\partial V(\phi)}{\partial\theta_r} = \frac{\partial V(\phi)}{\partial\phi_j}\frac{\partial\phi_j}{\partial\theta_r} = 0. \tag{46}$$

Now

$$\phi_j(\theta) = \phi_j(0) + i\theta_r\, t^r_{jk}\, \phi_k(0) \tag{47}$$

so the derivative (46) is

$$\frac{\partial V(\phi)}{\partial\phi_j}\frac{\partial\phi_j}{\partial\theta_r} = \frac{\partial V(\phi)}{\partial\phi_j} i t^r_{jk}\phi_k = 0 \tag{48}$$

because of the symmetry. We now differentiate this *vanishing* quantity with respect to ϕ_ℓ

$$\frac{\partial^2 V(\phi)}{\partial\phi_j \partial\phi_\ell} i t^r_{jk}\phi_k + \frac{\partial V(\phi)}{\partial\phi_j} i t^r_{j\ell} = 0 \tag{49}$$

and then set all the fields equal to their values ϕ_{k0} at one of the vacua

$$\frac{\partial^2 V(\phi_0)}{\partial\phi_j\partial\phi_\ell} it^r_{jk}\phi_{k0} + \frac{\partial V(\phi_0)}{\partial\phi_j} it^r_{j\ell} = \frac{\partial^2 V(\phi_0)}{\partial_\ell\partial\phi_j} it^r_{jk}\phi_{k0} = 0 \tag{50}$$

in which we used the vanishing (42) of the first derivatives at ϕ_0. But the second derivatives are just the elements $m_{\ell j}$ of the mass matrix (43). So we have for each generator t^r an eigenvector $t^r_{jk}\phi_{0k}$ of the mass matrix with eigenvalue zero

$$m_{\ell j}\, t^r_{jk}\phi_{0k} = 0 \tag{51}$$

unless the generator annihilates the vector ϕ_0. That is, if $I + i\theta_r t^r$ leaves the vector ϕ_0 unchanged

$$O\phi_0 = \phi_0, \tag{52}$$

then

$$\theta_r t^r \phi_0 = 0 \tag{53}$$

and we don't get an eigenvector of the mass matrix $m_{\ell j}$ with eigenvalue zero. There are as many massless bosons as there are generators of the symmetry group that don't annihilate the vacuum vector ϕ_0. If the symmetry group G has a subgroup H that leaves ϕ_0 invariant, then one says that there is a massless boson for every generator of the quotient group G/H. But for every generator in the subgroup H that leaves ϕ_0 invariant, there may be a massive boson.

4 Gauge Invariance

The reason we could generalize our formulas for muon pair production to tau pair production is that all the charged leptons are coupled to the photon in the same way. Although electrodynamics is an abelian gauge theory, we might as well consider the general case of a nonabelian gauge theory.

The action density of a Yang-Mills theory is unchanged when a space-time dependent unitary matrix $U(x)$ changes a vector $\psi(x)$ of matter fields to $\psi'(x) = U(x)\psi(x)$. Terms like $\psi^\dagger\psi$ are invariant because $\psi^\dagger(x)U^\dagger(x)U(x)\psi(x) = \psi^\dagger(x)\psi(x)$, but how can kinetic terms like $\partial_i\psi^\dagger\,\partial^i\psi$ be made invariant? Yang and

Mills introduced matrices A_i of gauge fields, replaced ordinary derivatives ∂_i by **covariant derivatives** $D_i \equiv \partial_i + A_i$, and required that $D'_i\psi' = UD_i\psi$ or that

$$(\partial_i + A'_i)\,U = \partial_i U + U\partial_i + A'_i U = U\,(\partial_i + A_i)\,. \tag{54}$$

Their nonabelian gauge transformation is

$$\begin{aligned} \psi'(x) &= U(x)\psi(x) \\ A'_i(x) &= U(x)A_i(x)U^\dagger(x) - (\partial_i U(x))\,U^\dagger(x). \end{aligned} \tag{55}$$

One often writes the unitary matrix as $U(x) = \exp(-ig\,\theta_a(x)\,t_a)$ in which g is a coupling constant, the functions $\theta_a(x)$ parametrize the gauge transformation, and the generators t_a belong to the representation that acts on the vector $\psi(x)$ of matter fields.

In the case of electrodynamics, the unitary matrix is a member of the group $U(1)$; it is just a phase factor $U(x) = \exp(-ie\,\theta_a(x))$. The abelian gauge transformation is

$$\begin{aligned} \psi'(x) &= U(x)\psi(x) = e^{-ie\,\theta(x)}\psi(x) \\ A'_i(x) &= U(x)A_i(x)U^\dagger(x) - (\partial_i U(x))\,U^\dagger(x) = A_i(x) + ie\partial_i\theta(x). \end{aligned} \tag{56}$$

I have been using a notation in which A_i is antihermitian to simplify the algebra. So if $A_i = iA_b$, then the abelian gauge transformation is

$$\begin{aligned} \psi'(x) &= U(x)\psi(x) = e^{-ie\,\theta(x)}\psi(x) \\ A'_b(x) &= U(x)A_b(x)U^\dagger(x) + i\,(\partial_b U(x))\,U^\dagger(x) = A_b(x) + e\partial_b\theta(x). \end{aligned} \tag{57}$$

Similarly, with real gauge fields, $A_b = -iA_i$, the nonabelian gauge transformation is

$$\begin{aligned} \psi'(x) &= U(x)\psi(x) \\ A'_b(x) &= U(x)A_b(x)U^\dagger(x) + i\,(\partial_b U(x))\,U^\dagger(x). \end{aligned} \tag{58}$$

5 Abelian Higgs Mechanism

A theory with action density

$$L = -\frac{1}{4}F_{ab}F^{ab} - (D_a\phi)^* D^a\phi - m^2\phi^*\phi - \frac{1}{2}\lambda(\phi^*\phi)^2 \tag{59}$$

in which the complex field ϕ is

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \tag{60}$$

and it covariant derivative is

$$D_b\phi = (\partial_b + ieA_b)\phi \tag{61}$$

describes charged bosons of mass m interacting with themselves directly and through the massless electromagnetic field A_b. This theory has an abelian gauge symmetry. That is, the action density is invariant under the spacetime-dependent $U(1)$ transformation $U(x) = \exp(-ie\theta(x))$

$$\begin{aligned} \psi'(x) &= U(x)\psi(x) = e^{-ie\,\theta(x)}\psi(x) \\ A'_b(x) &= U(x)A_b(x)U^\dagger(x) + i\,(\partial_i U(x))\,U^\dagger(x) = A_b(x) + e\partial_i\theta(x). \end{aligned} \tag{62}$$

But if we flip the sign of the mass term from $-m^2\phi^*\phi$ to $m^2\phi^*\phi$

$$L = -\frac{1}{4}F_{ab}F^{ab} - (D_a\phi)^* D^a\phi + m^2\phi^*\phi - \frac{1}{2}\lambda(\phi^*\phi)^2 \tag{63}$$

then things get more interesting. The complex field ϕ now minimizes the energy of the ground state of the theory by assuming a mean value

$$|\phi| = \frac{m}{\sqrt{\lambda}}. \tag{64}$$

The various possible vacua lie on a circle of radius $m/\sqrt{\lambda}$ in the complex ϕ-plane. We choose the one in which $\langle 0|\phi|0\rangle \equiv \phi_0 = m/\sqrt{\lambda}$ is real and set

$$\phi = \phi_0 + \frac{1}{\sqrt{2}}(\sigma_1 + i\sigma_2) = \frac{m}{\sqrt{\lambda}} + \frac{1}{\sqrt{2}}(\sigma_1 + i\sigma_2) = \frac{m}{\sqrt{\lambda}} + \sigma. \tag{65}$$

Now the potential energy is

$$
\begin{aligned}
V =& - m^2 \phi^* \phi + \frac{1}{2} \lambda \left(\phi^* \phi \right)^2 \\
=& - m^2 \left[\left(\phi_0 + \sigma_1/\sqrt{2} \right)^2 + \sigma_2^2/2 \right] + \frac{1}{2} \lambda \left[\left(\phi_0 + \sigma_1/\sqrt{2} \right)^2 + \sigma_2^2/2 \right]^2 \\
=& - m^2 \left[\phi_0^2 + \sqrt{2} \phi_0 \sigma_1 + \sigma_1^2/2 + \sigma_2^2/2 \right] + \frac{1}{2} \lambda \left[\phi_0^2 + \sqrt{2} \phi_0 \sigma_1 + \sigma_1^2/2 + \sigma_2^2/2 \right]^2 \\
=& - m^2 \left[\phi_0^2 + \sqrt{2} \phi_0 \sigma_1 + \sigma_1^2/2 + \sigma_2^2/2 \right] + \frac{1}{2} \lambda \left[\phi_0^2 (\sigma_1^2 + \sigma_2^2) + 2 \phi_0^2 \sigma_1^2 + 2\sqrt{2} \phi_0^3 \sigma_1 \right] \\
& + \sqrt{2}\, \lambda \phi_0 \sigma_1 \sigma^2 + \frac{\lambda}{2} \left(\sigma^2 \right)^2 + \frac{1}{2} \lambda \phi_0^4 .
\end{aligned}
\tag{66}
$$

The terms linear in σ_1 cancel, as do the terms quadratic in σ_2. We then have

$$
V = m^2 \, \sigma_1^2 + \sqrt{2\lambda} \, m \, \sigma_1 \sigma^2 + \frac{\lambda}{2} \left(\sigma^2 \right)^2 - \frac{m^4}{2\lambda} . \tag{67}
$$

So the theory seems to have a spinless boson σ_1 of mass $\sqrt{2}m$ and a massless spinless boson σ_2.

But wait. What about the kinetic action of the scalar fields? It is

$$
\begin{aligned}
L_{\phi,K} =& - \left(D_a \phi \right)^* D^a \phi = - \left(\partial_a - ie A_a \right) \left(\phi_0 + \sigma^* \right) \left(\partial^a + ie A^a \right) \left(\phi_0 + \sigma \right) \\
=& - \left(-ie \phi_0 A_a + D_a^* \sigma^* \right) \left(ie \phi_0 A^a + D^a \sigma \right) \\
=& - \left[-ie A_a \phi_0 + \left(\partial_a - ie A_a \right) \left(\frac{\sigma_1}{\sqrt{2}} - i \frac{\sigma_2}{\sqrt{2}} \right) \right] \left[ie A^a \phi_0 + \left(\partial^a + ie A^a \right) \left(\frac{\sigma_1}{\sqrt{2}} + i \frac{\sigma_2}{\sqrt{2}} \right) \right] \\
=& - e^2 \phi_0^2 A_a A^a - \left(D_a \sigma \right)^* D^a \sigma + ie A_a \phi_0 \, D^a \sigma - ie A^a \phi_0 \, \left(D_a \sigma \right)^* \\
=& - e^2 \phi_0^2 A_a A^a - \left(D_a \sigma \right)^* D^a \sigma - \sqrt{2} e^2 A_a A^a \phi_0 \sigma_1 - \sqrt{2} e A^a \phi_0 \partial_a \sigma_2
\end{aligned}
\tag{68}
$$

in which

$$
\sigma = \frac{1}{\sqrt{2}} \left(\sigma_1 + i \sigma_2 \right) . \tag{69}
$$

The gauge field A_a has acquired a mass

$$M = \sqrt{2}\, e\phi_0 = \sqrt{2}\, \frac{em}{\sqrt{\lambda}}. \tag{70}$$

It makes sense to change the name of this field to

$$B_a \equiv A_a + \frac{1}{M}\, \partial_a \sigma_2 = A_a + \frac{1}{\sqrt{2}\, e\phi_0}\, \partial_a \sigma_2 = A_a + \frac{\sqrt{\lambda}}{\sqrt{2}\, em}\, \partial_a \sigma_2. \tag{71}$$

Note that the extra gradient of σ_2 does not change the Faraday tensor

$$F_{ab} = \partial_a A_b - \partial_b A_a = \partial_a B_b - \partial_b B_a. \tag{72}$$

In these terms, the action density is

$$L = -\frac{1}{4} F_{ab} F^{ab} - \frac{1}{2} M^2 B_a B^a - (D_a \sigma)^* D^a \sigma + \text{a constant plus cubic and quartic terms} \tag{73}$$

in which

$$-(D_a \sigma)^* D^a \sigma = -\frac{1}{2}\left(\partial_a \sigma_1 \partial^a \sigma_1 + \partial_a \sigma_2 \partial^a \sigma_2\right) + eA^a \left(\sigma_1 \partial_a \sigma_2 - \sigma_2 \partial_a \sigma_1\right) - \frac{1}{2} e^2 A_a A^a \left(\sigma_1^2 + \sigma_2^2\right). \tag{74}$$

But it is more straightforward to write the action as

$$\begin{aligned} L = & -\frac{1}{4} F_{ab} F^{ab} - \frac{1}{2} M^2 A_a A^a - (D_a \sigma)^* D^a \sigma - m^2 \sigma_1^2 \\ & - \frac{\sqrt{2}}{\sqrt{\lambda}} e^2 m\, A_a A^a \sigma_1 - \frac{\sqrt{2}}{\sqrt{\lambda}}\, em\, A^a \partial_a \sigma_2 - \sqrt{2\lambda}\, m\, \sigma_1 \sigma^2 - \frac{1}{2} \lambda \left(\sigma^2\right)^2 + \frac{m^4}{2\lambda}. \end{aligned} \tag{75}$$

Apart from the quadratic term $-\sqrt{2}\, em\, A^a \partial_a \sigma_2 / \sqrt{\lambda}$, this theory would describe a vector boson A_b of mass $M = em\sqrt{2/\lambda}$ interacting with a scalar boson σ_1 of mass $\sqrt{2}\, m$, at least at low energies and low temperatures.

But this is a $U(1)$ gauge theory, so we can rotate the complex field ϕ at every point of space-time so as to make it real. Then, instead of

$$L = -\frac{1}{4}F_{ab}F^{ab} - (D_a\phi)^* D^a\phi + m^2\phi^*\phi - \frac{1}{2}\lambda\left(\phi^*\phi\right)^2, \tag{76}$$

we have

$$\begin{aligned} L &= -\frac{1}{4}F_{ab}F^{ab} - (D_a\phi)^* D^a\phi + m^2\phi^2 - \frac{1}{2}\lambda\phi^4 \\ &= -\frac{1}{4}F_{ab}F^{ab} - (\partial_a - ieA_a)\,\phi\,(\partial^a + ieA^a)\,\phi + m^2\phi^2 - \frac{1}{2}\lambda\phi^4 \\ &= -\frac{1}{4}F_{ab}F^{ab} - \partial_a\phi\,\partial^a\phi - e^2\,A_aA^a\,\phi^2 + m^2\phi^2 - \frac{1}{2}\lambda\phi^4. \end{aligned} \tag{77}$$

Now, we have the simplest kind of spontaneous symmetry breaking in which the real field ϕ assumes a mean value ϕ_0 whose square is m^2/λ. We choose

$$\langle 0|\phi|0\rangle = \phi_0 = \frac{m}{\sqrt{\lambda}} \tag{78}$$

and set

$$\phi = \frac{m}{\sqrt{\lambda}} + \frac{\sigma}{\sqrt{2}} \tag{79}$$

where now both ϕ and σ are real. In these terms, L is

$$\begin{aligned} L &= -\frac{1}{4}F_{ab}F^{ab} - \partial_a\phi\,\partial^a\phi - e^2\,A_aA^a\,\phi^2 + m^2\phi^2 - \frac{1}{2}\lambda\phi^4 \\ &= -\frac{1}{4}F_{ab}F^{ab} - \frac{1}{2}\partial_a\sigma\,\partial^a\sigma - e^2\,A_aA^a\left(\frac{m}{\sqrt{\lambda}} + \frac{\sigma}{\sqrt{2}}\right)^2 + m^2\left(\frac{m}{\sqrt{\lambda}} + \frac{\sigma}{\sqrt{2}}\right)^2 - \frac{1}{2}\lambda\left(\frac{m}{\sqrt{\lambda}} + \frac{\sigma}{\sqrt{2}}\right)^4 \\ &= -\frac{1}{4}F_{ab}F^{ab} - \frac{e^2m^2}{\lambda}\,A_aA^a - \frac{1}{2}\partial_a\sigma\,\partial^a\sigma - m^2\sigma^2 \\ &\quad - \sqrt{\frac{2}{\lambda}}\,me^2\,A_aA^a\,\sigma - \frac{1}{2}e^2\,A_aA^a\,\sigma^2 - \sqrt{\frac{\lambda}{2}}\,m\,\sigma^3 - \frac{\lambda}{8}\,\sigma^4 + \frac{m^4}{\lambda}. \end{aligned} \tag{80}$$

In this **unitary gauge**, the theory has a real scalar boson of mass $\sqrt{2}\,m$ interacting with a massive vector boson A_a of mass $M = em\sqrt{2/\lambda}$. In the quadratic part of L, there are no terms coupling σ to A_b. If both e and λ are small, then perturbation theory should describe σ interacting with A_b and with itself through the cubic and quartic terms in the second line of the last form of this equation. This is the abelian **Higgs mechanism**.

One may wonder whether one can transform to the unitary gauge even when the mass term $-m^2|\phi|^2$ has the "right" sign so that the $U(1)$ symmetry is unbroken. The phase $e\theta = \text{atan}(\phi_2/\phi_1)$ of the required gauge transformation is not defined where it is most needed, namely in the vicinity of the vacuum where $\langle 0|\phi|0\rangle = 0$, and has derivatives that are singular there.

6 *O(n)* Nonabelian Higgs Mechanism

Let's start with our $O(n)$ theory (30)

$$L = -\frac{1}{2}\sum_{i=1}^{n} \partial_a\phi_i\,\partial^a\phi_i + \frac{1}{2}m^2\phi^2 - \frac{1}{4}\lambda\left(\phi^2\right)^2. \tag{81}$$

in which the sign of the mass term induces spontaneous symmetry breaking and

$$\phi^2 = \sum_{i=1}^{n} \phi_i\,\phi_i. \tag{82}$$

We can make this **global** $O(n)$ symmetry **local** by introducing $n(n-1)$ gauge fields A_b^f, one for each generator t^f of the group $O(n)$. The antihermitian gauge-field matrix is

$$A_b(x) = ie\sum_{f=1}^{n(n-1)} t^f A_b^f(x) \tag{83}$$

in which the generators obey the commutation relations

$$[t^f, t^g] = if_{fgk}t^k \tag{84}$$

with totally antisymmetric **structure constants** f_{fgk}. The generators are orthogonal but not normalized

$$\mathrm{Tr}(t^{f\dagger} t^g) = k \delta_{fg} \tag{85}$$

in which the positive constant k depends upon the representation to which the generators belong. In the defining representation of $O(n)$, the generators are $n \times n$ imaginary antisymmetric matrices. The covariant derivative is

$$D_b = \partial_b + A_b \tag{86}$$

The nonabelian Faraday tensor

$$F_{ik} = [D_i, D_k] = \partial_i A_k - \partial_k A_i + [A_i, A_k] \tag{87}$$

transforms covariantly

$$F'_{ik} = U F_{ik} U^{-1}. \tag{88}$$

The action of the nonabelian Faraday tensor is invariant

$$\mathrm{Tr}\left(U F_{ik} U^{-1} U F^{ik} U^{-1}\right) = \mathrm{Tr}\left(U F_{ik} F^{ik} U^{-1}\right) = \mathrm{Tr}\left(F_{ik} F^{ik}\right). \tag{89}$$

The action density of this theory is

$$L = \frac{1}{4}\mathrm{Tr}(F_{ik}F^{ik}) - \frac{1}{2}(D_a\phi)^{\mathsf{T}}\,(D^a\phi) + \frac{1}{2}m^2\phi^2 - \frac{1}{4}\lambda\left(\phi^2\right)^2 \tag{90}$$

in which the sign of the trace is due to the i in $A_b = i\, e\, t^f\, A^f_b$ (summed over f).

Once again, we have spontaneous symmetry breaking. In the vacuum state $|0\rangle$, the field ϕ, which is a real n-vector assumes a value on the sphere of radius $|\phi| = \phi_0 = m/\sqrt{\lambda}$. As before, we write

$$\phi = (\pi_1, \pi_2, \ldots, \pi_{n-1}, \phi_0 + \sigma). \tag{91}$$

We have one scalar field of mass $m_\sigma = \sqrt{2}\, m$ and, at least apparently, $n-1$ massless scalar fields π_i, one for each generator of the quotient group G/H in which H is the subgroup of rotations that leave the vector

$$\langle 0|\phi_i|0\rangle = (0, 0, \ldots, 0, \phi_0) \tag{92}$$

invariant. This subgroup H is $O(n-1)$. So there are

$$\tfrac{1}{2}n(n-1) - \tfrac{1}{2}(n-1)(n-2) = n-1 \tag{93}$$

apparently massless scalar fields π_i

Look now at the kinetic action of the fields ϕ_i

$$L_{K\phi} = -\frac{1}{2}(D_a\phi)^{\mathsf{T}}(D^a\phi) = -\frac{1}{2}\sum_{i,k,\ell=1}^{n}\left(\delta_{ik}\partial_a + \sum_{f=1}^{n(n-1)/2} iet^f_{ik}A^f_a\right)\phi_k\left(\delta_{i\ell}\partial^a + \sum_{g=1}^{n(n-1)/2} iet^g_{i\ell}A^{ga}\right)\phi_\ell. \tag{94}$$

Since $\langle 0|\pi_i|0\rangle = \phi_0\delta_{in}$, the part of this that involves ϕ_0 quadratically is

$$L_{K\phi_0} = -\frac{1}{2}ie\,t^f_{ik}\,A^f_a\phi_{0k}\,ie\,t^g_{i\ell}\,A^{ga}\phi_{0\ell} = \frac{1}{2}\sum_{i=1}^{n}e^2\,t^f_{in}\,t^g_{in}\,\phi_0^2\,A^f_a\,A^{ga} \tag{95}$$

in which we sum over i but not n. The $n\times n$ generators t^f are hermitian and imaginary so that $it^f\phi$ is real; they are imaginary and antisymmetric. Thus

$$L_{K\phi_0} = \frac{1}{2}e^2\,\phi_0^2\,t^f_{in}t^g_{in}\,A^f_a\,A^{ga} \equiv -\frac{1}{2}e^2\,\phi_0^2\,it^f_{in}it^g_{in}\,A^f_a\,A^{ga} = -\frac{1}{2}M^2_{fg}\,A^f_a\,A^{ga} \tag{96}$$

in which the $n(n-1)/2 \times n(n-1)/2$ matrix

$$M^2_{fg} = e^2\phi_0^2 it^f_{in}it^g_{in} \tag{97}$$

is real and symmetric. The mass-squared matrix M^2_{fg} also is nonnegative since for any real vector V_g, the vector

$$R_i = it^f_{in}\,V_f \tag{98}$$

is real, and we have

$$V_f M^2_{fg} V_g = e^2\phi_0^2 V_f it^f_{in}it^g_{in}V_g = e^2\phi_0^2\,R_i^2 \geq 0. \tag{99}$$

In the present theory, this matrix does have zero eigenvalues because the $(n-1)(n-2)/2$ generators t^h of the subgroup H leave the vector

$$\langle 0|\phi_i|0\rangle = (0, 0, \ldots, 0, \phi_0) \tag{100}$$

invariant, and so

$$t^h_{ik}\langle 0|\phi_k|0\rangle = t^h_{in}\langle 0|\phi_n|0\rangle = 0 \quad \text{or} \quad t^h_{in} = 0. \tag{101}$$

That is,

$$M^2_{fh} = M^2_{hg} = M^2_{hh'} = 0. \tag{102}$$

So only the gauge bosons of the $(n-1)(n-2)/2$ generators t^h of the subgroup H that leave the vector

$$\langle 0|\phi_i|0\rangle = (0, 0, \ldots, 0, \phi_0) \tag{103}$$

invariant remain massless. The gauge bosons of the $n(n-1)/2 - (n-1)(n-2)/2 = n-1$ other generators get masses from the diagonal part of the $(n-1) \times (n-1)$ block of the mass-squared matrix M^2_{fg} that can have positive eigenvalues. These $n-1$ eigenvalues are all positive because if we choose the generators so that $t^f_{in} = i\delta_{fi}\sqrt{k/2}$ for $1 \le f, g \le n-1$, then

$$M^2_{fg} = e^2\phi_0^2 it^f_{in} it^g_{in} = e^2\phi_0^2\, \delta_{fg}\, k/2 \tag{104}$$

for $1 \le f, g \le n-1$, in which k is the positive constant that appears in the orthogonality relation (85). So there are $n-1$ massive vector bosons. These are the ones that absorb the $n-1$ apparently massless scalar bosons π_i. They all get the same mass $e\phi_0\sqrt{k}$, which depends upon the parameter k as well as upon e and ϕ_0.

As in the theory in which a nonzero mean value in the vacuum of a scalar field broke a local $U(1)$ symmetry, we can here transform to the unitary gauge in which the scalar field points in the nth direction at every point of spacetime

$$\phi = (0, 0, \ldots, 0, \phi_0 + \sigma). \tag{105}$$

In this gauge, the kinetic action of the fields ϕ_i

$$L_{K\phi} = -\frac{1}{2}(D_a\phi)^{\mathsf{T}}(D^a\phi) = -\frac{1}{2}\sum_{i,k,\ell=1}^{n}\left(\delta_{ik}\partial_a + \sum_{f=1}^{n(n-1)/2} iet^f_{ik}A^f_a\right)\phi_k\left(\delta_{i\ell}\partial^a + \sum_{g=1}^{n(n-1)/2} iet^g_{i\ell}A^{ga}\right)\phi_\ell \tag{106}$$

reduces to

$$L_{K\phi ug} = -\frac{1}{2}\sum_{i=1}^{n}\left(\delta_{in}\partial_a + \sum_{f=1}^{(n-1)} iet^f_{in}A^f_a\right)\phi_n\left(\delta_{in}\partial^a + \sum_{g=1}^{(n-1)} iet^g_{in}A^{ga}\right)\phi_n \tag{107}$$

in which the sum is only over the $n-1$ generators for which $t^f_{in} \neq 0$ for some i. Thus, summing over repeated indices, we have

$$\begin{aligned} L_{K\phi ug} &= -\frac{1}{2}\left(\delta_{in}\partial_a + iet^f_{in}A^f_a\right)(\phi_0+\sigma)\left(\delta_{in}\partial^a + iet^g_{in}A^{ga}\right)(\phi_0+\sigma) \\ &= -\frac{1}{2}\left(\delta_{in}\partial_a + iet^f_{in}A^f_a\right)\sigma\left(\delta_{in}\partial^a + iet^g_{in}A^{ga}\right)\sigma - ie\phi_0 t^f_{nn}A^f_a\partial^a\sigma - \frac{1}{2}e^2\, it^f_{in}\, it^g_{in}\, \phi_0^2\, A^f_a\, A^{ga} \\ &= -\frac{1}{2}\left(\delta_{in}\partial_a\sigma + iet^f_{in}A^f_a\sigma\right)\left(\delta_{in}\partial^a\sigma + iet^g_{in}A^{ga}\sigma\right) - \frac{1}{2}M^2_{fg}\, A^f_a\, A^{ga} \end{aligned} \tag{108}$$

because $t^f_{nn} = 0$ since the generators of $O(n)$ are antisymmetric.

7 *SU(2)* Higgs Mechanism

Let's now consider a theory of a 2-component complex scalar field ϕ with action density

$$L = -(D_a\phi)^\dagger D^a\phi + \frac{1}{4k}\mathrm{Tr}(F_{ab}F^{ab}) + m^2|\phi|^2 - \lambda|\phi|^4 \tag{109}$$

in which $|\phi|^2 = |\phi_1|^2 + |\phi_2|^2$. Here the antihermitian gauge-field matrix is

$$A_b = ie\frac{1}{2}\boldsymbol{\sigma}\cdot\boldsymbol{A}_b \tag{110}$$

in which the $\boldsymbol{\sigma}$ are the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{111}$$

The covariant derivative is

$$(D_b\phi)_i = \partial_b\phi_i + (A_b)_{ij}\phi_j = \partial_b\phi_i + (ie\frac{1}{2}\boldsymbol{\sigma}\cdot\boldsymbol{A}_b)_{ij}\phi_j = \partial_b\phi_i + ie\,\frac{1}{2}(\sigma^k)_{ij}\,A_b^k\,\phi_j. \tag{112}$$

This action density is invariant under $SU(2)$ transformations

$$\begin{aligned}\phi'(x) &= U(x)\phi(x) = \exp(-i\boldsymbol{\theta}(x)\cdot\boldsymbol{\sigma}/2)\phi(x)\\ A_b'(x) &= U(x)A_b(x)U^\dagger(x) + i\,(\partial_b U(x))\,U^\dagger(x)\\ &= \exp(-i\boldsymbol{\theta}(x)\cdot\boldsymbol{\sigma}/2)A_b(x)\exp(i\boldsymbol{\theta}(x)\cdot\boldsymbol{\sigma}/2) + i\,(\partial_b\exp(-i\boldsymbol{\theta}(x)\cdot\boldsymbol{\sigma}/2))\exp(i\boldsymbol{\theta}(x)\cdot\boldsymbol{\sigma}/2)\end{aligned} \tag{113}$$

that depend upon the space-time point x. The trace relation for the generators $\boldsymbol{\sigma}/2$ is

$$\mathrm{Tr}(\tfrac{1}{2}\sigma_i\tfrac{1}{2}\sigma_j) = \frac{1}{2}\delta_{ij}, \tag{114}$$

so the constant $k = 1/2$.

Going again to the unitary gauge, we rotate the field ϕ to

$$\phi = \begin{pmatrix}0\\ v+s\end{pmatrix} \quad \text{with} \quad \phi_0 = \begin{pmatrix}0\\ v\end{pmatrix} \tag{115}$$

in which $v = m\sqrt{2/\lambda}$ is real. The kinetic action of the Higgs field ϕ has a term quadratic in ϕ

$$\left[ie\,\frac{1}{2}(\sigma^k)_{ij}\,A_b^k\,\phi_j\right]^\dagger ie\,\frac{1}{2}(\sigma^m)_{i\ell}\,A^{mb}\,\phi_\ell \tag{116}$$

which is simpler in matrix notation

$$\frac{e^2}{4}\,\phi_0^\dagger\sigma^k\sigma^m\phi_0\,A_b^k\,A^{mb}. \tag{117}$$

Symmetrizing and using the relation

$$\sigma^k\sigma^m = \delta_{km}I + i\epsilon_{km\ell}\sigma^\ell, \tag{118}$$

we find at $\phi = \phi_0$

$$\frac{e^2}{8}\,\phi_0^\dagger\,\{\sigma^k,\sigma^m\}\,\phi_0\,A_b^k\,A^{mb} = \frac{e^2}{4}\,\phi_0^\dagger\,\delta_{km} I\,\phi_0\,A_b^k\,A^{mb} = \frac{e^2 v^2}{4}\,A_b^k\,A^{kb} \equiv \frac{1}{2} M^2\,A_b^k\,A^{kb}. \tag{119}$$

So all three gauge bosons get the same mass $M = ev/\sqrt{2}$. Why are none left massless? Because the $SU(2)$ symmetry is completely broken; H is just the trivial group consisting of an identity element.

What if we had put ϕ in the adjoint representation of $SU(2)$? In this case, its mean value would have been a real three vector of some length pointing in some direction. The unbroken subgroup H would have consisted of rotations about that direction. So one gauge boson would have remained massless. In fact, this is a model we already have studied: it is just the $O(3)$ gauge theory with the Higgs in the defining representation.

8 *SU(2)* with the Higgs in the adjoint representation

If the Higgs field h is a matrix that transforms in the adjoint representation, so that $h' = UhU^\dagger$, then its covariant derivative is

$$D_b h = \partial_b h + [A_b, h] \tag{120}$$

in which the anti hermitian gauge field is

$$A_b = ie\frac{1}{2}\boldsymbol{\sigma}\cdot\boldsymbol{A}_b \tag{121}$$

which is (110). The mass term for the gauge fields arises from the action of the Higgs field

$$-\operatorname{Tr}\left[(D_b h)^\dagger D^b h\right] = -\operatorname{Tr}\left[(\partial_b h + [A_b, h])^\dagger(\partial_b h + [A_b, h])\right]. \tag{122}$$

If h_0 is the mean value of the Higgs field in the vacuum, then this kinetic action makes the mass term

$$-\operatorname{Tr}\left([A_b, h_0]^\dagger [A_b, h_0]\right) = -\operatorname{Tr}\left([h_0^\dagger, A_b^\dagger][A_b, h_0]\right) = -\operatorname{Tr}\left([A_b, h_0^\dagger][A_b, h_0]\right) \tag{123}$$

since A is antihermitian. When h_0 is a multiple of the identity matrix, the commutator $[A_b, h_0]$ vanishes, and all the gauge bosons remain massless.

9 Which gauge fields are left massless?

Suppose the Higgs field ϕ has a mean value ϕ^0 in the vacuum. Suppose the generator $c_b t^b$ sends ϕ^0 to zero

$$c_b\, t^b_{ij}\, \phi^0_j = 0. \tag{124}$$

The mass-squared term is

$$\frac{1}{2} A^a_\mu\, M^2_{ab}\, A^{b\mu} = A^a_\mu\, \phi^{0\dagger}_i t^a_{ij}\, t^b_{jk} \phi^0_k\, A^{b\mu} \tag{125}$$

and so the vector $(c_1, c_2, \dots)$ is an eigenvector of the matrix M^2_{ab} with eigenvalue zero

$$\frac{1}{2} M^2_{ab}\, c_b = \phi^{0\dagger}_i t^a_{ij}\, (t^b_{jk} \phi^0_k\, c_b) = 0. \tag{126}$$

Thus the diagonal form of this matrix is

$$M^2_{ab} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} 0 \begin{pmatrix} c_1 & c_2 & \dots \end{pmatrix} + \begin{pmatrix} d_1 \\ d_2 \\ \vdots \end{pmatrix} m_d \begin{pmatrix} d_1 & d_2 & \dots \end{pmatrix} + \dots, \tag{127}$$

and so coefficient or mass of the (unnormalized) gauge field

$$A_\mu = c_b A^b_\mu \tag{128}$$

in the mass-squared term (125) is zero:

$$\frac{1}{2} A^a_\mu M^2_{ab} A^{b\mu} = \frac{1}{2}(c_a A^a_\mu)\, 0\, (c_b A^{b\mu}) + \frac{1}{2}(d_a A^a_\mu)\, m_d\, (d_b A^{b\mu}) + \dots. \tag{129}$$

and so a linear combination of gauge fields

$$A_\mu = \frac{c_b A^b_\mu}{\sqrt{c_1^2 + c_2^2 + \dots}} \tag{130}$$

remains massless if the corresponding linear combination of generators sends the mean value ϕ^0 of the Higgs field to zero (124) or equivalently if the unitary transformation $U = \exp(ic_b\, t^b)$ leaves that mean value invariant

$$U\phi_0 = e^{ic_b\, t^b}\, \phi_0 = \phi_0. \tag{131}$$

The corresponding charge $c_b T^b$ leaves the vacuum invariant

$$U|\phi_0\rangle = e^{ic_b\, T^b}\, |\phi_0\rangle = |\phi_0\rangle \tag{132}$$

and so generates an unbroken symmetry.

10 *SU(3)* Pure Gauge Theory

The Gell-Mann matrices are

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\
\lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \text{and} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
\end{aligned} \tag{133}$$

The generators t_a of the 3×3 defining representation of $SU(3)$ are these Gell-Mann matrices divided by 2

$$t_a = \lambda_a/2 = t^a = \lambda^a/2 \tag{134}$$

(Murray Gell-Mann, 1929–).

The eight generators t_a are orthogonal with $k = 1/2$

$$\mathrm{Tr}\,(t_a t_b) = \frac{1}{2}\delta_{ab} \tag{135}$$

and satisfy the commutation relation

$$[t_a, t_b] = i f_{abc}\, t_c. \tag{136}$$

A trace formula gives us the ***SU*(3) structure constants** as

$$f_{abc} = (-i/k)\mathrm{Tr}([t_a, t_b]t_c) = -2i\mathrm{Tr}\,([t_a, t_b]t_c)\,. \tag{137}$$

They are real and totally antisymmetric with $f_{123} = 1$, $f_{458} = f_{678} = \sqrt{3}/2$, and $f_{147} = -f_{156} = f_{246} = f_{257} = f_{345} = -f_{367} = 1/2$.

While no two generators of $SU(2)$ commute, two generators of $SU(3)$ do. In the representation (133,134), t_3 and t_8 are diagonal and so commute

$$[t_3, t_8] = 0. \tag{138}$$

They generate the **Cartan subalgebra** of $SU(3)$.

The gauge-field matrix is

$$A_\mu(x) = ig \sum_{b=1}^{8} t^b A^b_\mu(x) \tag{139}$$

in the defining representation. The covariant derivative in that representation is

$$D_\mu = I\partial_\mu + A_\mu(x) = I\partial_\mu + ig \sum_{b=1}^{8} t^b A^b_\mu(x). \tag{140}$$

The Faraday matrix is

$$F_{\mu\nu} = [D_\mu, D_\nu] = [I\partial_\mu + A_\mu(x), I\partial_\nu + A_\nu(x)] = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \tag{141}$$

in matrix notation. With more indices exposed, it is

$$
\begin{aligned}
(F_{\mu\nu})_{cd} &= (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu])_{cd} \\
&= ig \sum_{b=1}^{8} t^b_{cd} \left(\partial_\mu A^b_\nu - \partial_\nu A^b_\mu\right) + \left([ig \sum_{b=1}^{8} t^b A^b_\mu, ig \sum_{e=1}^{8} t^e A^e_\nu] \right)_{cd} .
\end{aligned}
\tag{142}
$$

To avoid the sum signs, we sum over repeated indices from 1 to 8

$$
\begin{aligned}
(F_{\mu\nu})_{cd} &= igt^b_{cd} \left(\partial_\mu A^b_\nu - \partial_\nu A^b_\mu\right) - g^2 A^b_\mu A^e_\nu \left([t^b, t^e]\right)_{cd} = igt^b_{cd} \left(\partial_\mu A^b_\nu - \partial_\nu A^b_\mu\right) - g^2 A^b_\mu A^e_\nu i f_{bef} t^f_{cd} \\
&= igt^b_{cd} \left(\partial_\mu A^b_\nu - \partial_\nu A^b_\mu\right) - ig^2 A^b_\mu A^e_\nu f_{bef} t^f_{cd} = igt^b_{cd} \left(\partial_\mu A^b_\nu - \partial_\nu A^b_\mu\right) - ig^2 A^f_\mu A^e_\nu f_{feb} t^b_{cd} \\
&= igt^b_{cd} \left(\partial_\mu A^b_\nu - \partial_\nu A^b_\mu - g A^f_\mu A^e_\nu f_{feb}\right) = igt^b_{cd} \left(\partial_\mu A^b_\nu - \partial_\nu A^b_\mu - g f_{bfe} \, A^f_\mu A^e_\nu\right) \\
&= igt^b_{cd} \, F^b_{\mu\nu}
\end{aligned}
\tag{143}
$$

where

$$
F^b_{\mu\nu} = \partial_\mu A^b_\nu - \partial_\nu A^b_\mu - g f_{bfe} \, A^f_\mu A^e_\nu \tag{144}
$$

is the Faraday tensor.

The action density of this tensor is

$$
L_F = \, - \frac{1}{4} F^b_{\mu\nu} F^{\mu\nu}_b \tag{145}
$$

in which raising and lowering an index of a compact group is of cosmetic, not cosmic, significance. The trace of the square of the Faraday matrix is

$$
\mathrm{Tr}\left[F_{\mu\nu} F^{\mu\nu}\right] = \mathrm{Tr}\left[igt^b \, F^b_{\mu\nu} \, igt^c \, F^{\mu\nu}_c\right] = -g^2 F^b_{\mu\nu} \, F^{\mu\nu}_c \, \mathrm{Tr}(t^b t^c) = -g^2 F^b_{\mu\nu} \, F^{\mu\nu}_c \, k\delta_{bc} = -kg^2 \, F^b_{\mu\nu} \, F^{\mu\nu}_b . \tag{146}
$$

So the Faraday action density is

$$
L_F = \, - \frac{1}{4} F^b_{\mu\nu} F^{\mu\nu}_b = \frac{1}{4kg^2} \, \mathrm{Tr}\left[F_{\mu\nu} F^{\mu\nu}\right] = \frac{1}{2g^2} \, \mathrm{Tr}\left[F_{\mu\nu} F^{\mu\nu}\right] . \tag{147}
$$

The theory described by this action density, without scalar or spinor fields, is called **pure** gauge theory.

11 Quantum Chromodynamics

If we add massless quarks in the fundamental or defining representation, then we gets theory of the strong interactions called **Quantum Chromodynamics**. Thus, let ψ be a complex 3-vector of Dirac fields

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix} \tag{148}$$

(so 12 fields in all). This complex 12-vector could represent u or "up" quarks. We use the covariant derivative

$$D_\mu = I\partial_\mu + A_\mu(x) = I\partial_\mu + ig\sum_{b=1}^{8} t^b A^b_\mu(x). \tag{149}$$

The action density then is

$$L = \frac{1}{2g^2}\,\mathrm{Tr}\left[F_{\mu\nu}F^{\mu\nu}\right] - \overline{\psi}\left(\gamma^\mu D_\mu + m\right)\psi. \tag{150}$$

Nonperturbative effects are supposed to "confine" the quarks and massless gluons. There are 6 known "flavors" of quarks—u, d, c, s, t, b.

12 *SU(3)* Higgs Mechanism

Let's now add a triplet of complex scalar fields that transform according to the defining representation

$$\phi'_b(x) = U_{bc}(x)\,\phi_c(x) = \left[e^{-i\theta^a(x)t^a}\right]_{bc}\phi_c(x). \tag{151}$$

The $SU(3)$ gauge fields will transform as

$$A'_\mu(x) = U(x)A_\mu(x)U^\dagger(x) + i(\partial_\mu U(x))U^\dagger(x) = e^{-i\theta^a(x)t^a}\,A_\mu(x)\,e^{i\theta^a(x)t^a} + i(\partial_\mu e^{-i\theta^a(x)t^a})e^{i\theta^a(x)t^a} \tag{152}$$

in which the gauge-field matrix is

$$A_\mu(x) = ig \sum_{b=1}^{8} t^b A^b_\mu(x) \tag{153}$$

in the defining representation.

Suppose the action density is

$$L = \frac{1}{4kg^2} \operatorname{Tr}\left[F_{\mu\nu}F^{\mu\nu}\right] - (D_\mu\phi)^\dagger D^\mu\phi + m^2\,|\phi|^2 - \frac{1}{2}\lambda^2\,|\phi|^4. \tag{154}$$

I have fiddled with the coefficients so as to avoid extra factors and roots of 2. Once again, we have spontaneous breaking of the local $SU(3)$ symmetry as the vacuum arranges itself so as to give the scalar field a mean value

$$\phi_0 \equiv \langle 0|\phi_a(x)|0\rangle = v\delta_{a3} = \frac{m}{\lambda}\,\delta_{a3} \tag{155}$$

so that

$$\phi_3 = \frac{1}{\sqrt{2}}(\sqrt{2}v + s + i\phi_{3i}) = v + \frac{1}{\sqrt{2}}(s + i\phi_{3i}) \tag{156}$$

in which the choice of the third direction was arbitrary. The complex fields ϕ_1 and ϕ_2, and the imaginary part of ϕ_3 remain massless, but the real part of ϕ_3 acquires the mass $\sqrt{2}\,m$.

Now instead of (119), we have

$$\frac{e^2}{8}\,\phi_0^\dagger\,\{\lambda^k, \lambda^m\}\,\phi_0\,A^k_b\,A^{mb} \equiv \frac{1}{2}M^2\,A^k_b\,A^{kb}. \tag{157}$$

The gauge fields that don't move ϕ_0, that is, the ones that have

$$\lambda^m\phi_0 = 0 \tag{158}$$

remain massless. So A^1, A^2, and A^3 remain massless. The other five gauge bosons, $A^4 \ldots A^8$ absorb the five massless scalar fields and acquire masses.

Homework set 3: Find those masses.

Let's put the scalar fields in the adjoint representation of $SU(3)$. Now there are 8 real scalar fields, and we can write them as an 8-vector ϕ or as a 3×3 matrix

$$\Phi = \sum_{a=1}^{8} t^a \phi^a. \tag{159}$$

The covariant derivative now is

$$D_\mu \phi = (\partial_\mu + igA_\mu)\phi = (\partial_\mu + igA^b_\mu T^b)\phi \tag{160}$$

where the generators in the adjoint representation are

$$T^b_{ac} = i f_{abc} \tag{161}$$

in which the structure constants f_{abc} are real and totally antisymmetric. Thus, we have

$$(D_\mu \phi)_a = (\delta_{ac}\partial_\mu + igA_{\mu ac})\phi_c = \partial_\mu \phi_a - gA^b_\mu f_{abc}\phi_c. \tag{162}$$

We also can write this as

$$\begin{aligned} D_\mu \Phi &= \partial_\mu \Phi + g[A_\mu, \Phi] = t^a \partial_\mu \phi_a + ig[t^b, t^c] A^b_\mu \phi^c = t^a \partial_\mu \phi_a + igi f_{bca} t^a A^b_\mu \phi^c \\ &= t^a \partial_\mu \phi_a - g t^a A^b_\mu f_{abc} \phi_c = t^a \left(\partial_\mu \phi_a - g A^b_\mu f_{abc} \phi_c \right) = t^a (D_\mu \phi)_a. \end{aligned} \tag{163}$$

Now the gauge-boson mass term inside $\frac{1}{2}(D_\mu \phi)^a (D^\mu \phi)_a$ is the proportional to the trace

$$g^2 \mathrm{Tr} \left([t^a, \Phi][t^b, \Phi] \right) A^a_\mu A^\mu_b. \tag{164}$$

So is the vacuum gives Φ the mean value

$$\Phi_0 = \langle 0|\Phi|0\rangle, \tag{165}$$

then the gauge-boson mass term is proportional to the trace

$$g^2 \mathrm{Tr} \left([t^a, \Phi_0][t^b, \Phi_0] \right) A^a_\mu A^\mu_b. \tag{166}$$

So those linear combinations of gauge fields times generators that commute with Φ_0 remain massless.

For instance, if

$$\Phi_0 \propto \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \tag{167}$$

then the gauge bosons A^1, A^2, A^3 and A^8 remain massless, while A^4, A^5, A^6, and A^7 become massive. Interestingly, the $SU(3)$ symmetry is broken to $SU(2) \times U(1)$.

On the other hand, if

$$\Phi_0 \propto \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{168}$$

then only A^3 and A^8 remain massless, and the unbroken symmetry is just $U(1) \times U(1)$.

13 The GSW Electroweak Model

The local gauge group is $SU(2)_\ell \times U(1)$. What's weird is that the $SU(2)_\ell$ symmetry applies only to the "left-handed" quarks and leptons and to a complex doublet (or 2-vector) of scalar fields H, the Higgs.

13.1 The Higgs Mechanism

Let's leave out the fermions for the moment, and focus just on the Higgs and the gauge fields. The gauge transformation is

$$\begin{aligned} H'(x) &= U(x)H(x) \\ A'_\mu(x) &= U(x)A_\mu(x)U^\dagger(x) + (\partial_\mu U(x))U^\dagger(x) \end{aligned} \tag{169}$$

in which the 2×2 unitary matrix $U(x)$ is

$$U(x) = \exp\left[i\, g\, \frac{\sigma^a}{2}\, \alpha^a(x) + i\, g'\, \frac{Y\, I}{2}\, \beta(x)\right]. \tag{170}$$

The generators here are the 3 Pauli matrices and the matrix $I/2$, where I is the 2×2 identity matrix.

The action density of the theory (without the fermions) is

$$L = -(D_\mu H)^\dagger D^\mu H + \frac{1}{4k}\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}) + m^2|H|^2 - \lambda|H|^4 \tag{171}$$

in which the covariant derivative for the Higgs doublet is

$$D_\mu H = \left(I\partial_\mu + i\, g\, \frac{\sigma^a}{2}\, A_\mu^a + i\, g'\, \frac{Y\, I}{2}\, B_\mu \right) H. \tag{172}$$

The minimum of the Higgs potential is where

$$0 = \frac{\partial V}{\partial |H|^2} = 2\lambda|H|^2 - m^2. \tag{173}$$

So

$$\langle 0||H||0\rangle = \frac{m}{\sqrt{2\lambda}} = \frac{v}{\sqrt{2}}. \tag{174}$$

Going to unitary gauge, we transform this mean value to

$$H_0 = \langle 0|H(x)|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{175}$$

In unitary gauge, the Higgs potential is

$$V(v) = -\frac{1}{2}m^2 v^2 + \frac{1}{4}\lambda v^4, \tag{176}$$

and its second derivative is

$$V''(v) = 3\lambda v^2 - m^2 = 2m^2 = m_H^2. \tag{177}$$

The mass of the Higgs then is

$$m_H = \sqrt{2}\, m = \sqrt{2}\, v\, \lambda. \tag{178}$$

Experiments at LEP2 (see below) revealed value of v to be

$$v = 246 \quad \text{GeV}. \tag{179}$$

In 2012, experiments at the LHC showed the Higgs's mass to be

$$m_H = 126 \quad \text{GeV}. \tag{180}$$

The self coupling λ therefore is

$$\lambda = \frac{m_H}{\sqrt{2}\,v} = \frac{126}{\sqrt{2}\,246} = 0.36. \tag{181}$$

In unitary gauge and after spontaneous symmetry breaking, the mass terms for the gauge bosons that emerge from $-(D_\mu H)^\dagger D^\mu H$ are

$$\begin{aligned} L_M = & -\frac{1}{2}(0,v)\left(g\,\frac{\sigma^a}{2}\,A^a_\mu + g'\,\frac{I}{2}\,B_\mu\right)\left(g\,\frac{\sigma^a}{2}\,A^\mu_a + g'\,\frac{I}{2}\,B^\mu\right)\begin{pmatrix}0\\ v\end{pmatrix} \\ = & -\frac{1}{8}(0,v)\begin{pmatrix} gA^3_\mu + g'B_\mu & g(A^1_\mu - iA^2_\mu) \\ g(A^1_\mu + iA^2_\mu) & -gA^3_\mu + g'B_\mu \end{pmatrix}\begin{pmatrix} gA^\mu_3 + g'B^\mu & g(A^\mu_1 - iA^\mu_2) \\ g(A^\mu_1 + iA^\mu_2) & -gA^\mu_3 + g'B^\mu \end{pmatrix}\begin{pmatrix}0\\ v\end{pmatrix} \\ = & -\frac{v^2}{8}\left(g(A^1_\mu + iA^2_\mu), \quad -gA^3_\mu + g'B_\mu\right)\begin{pmatrix} g(A^\mu_1 - iA^\mu_2) \\ -gA^\mu_3 + g'B^\mu \end{pmatrix} \\ = & -\frac{v^2}{8}\left[g^2\left(A^1_\mu A^\mu_1 + A^2_\mu A^\mu_2\right) + \left(-gA^3_\mu + g'B_\mu\right)\left(-gA^\mu_3 + g'B^\mu\right)\right]. \end{aligned} \tag{182}$$

The normalized complex, charged gauge bosons are

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(A^1_\mu \mp iA^2_\mu\right) \tag{183}$$

and the normalized neutral one is

$$Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}}\left(gA^3_\mu - g'B_\mu\right). \tag{184}$$

In terms of these properly normalized fields, the mass terms are

$$L_M = -\frac{g^2 v^2}{4} W^-_\mu W^{+\mu} - \frac{(g^2 + g'^2) v^2}{8} Z_\mu Z^\mu. \tag{185}$$

So the W^+ and the W^- get the same mass

$$M_W = g\,\frac{v}{2} = 80.385 \text{ GeV}/c^2. \tag{186}$$

while the Z (also called the Z^0) has mass

$$M_Z = \sqrt{g^2 + g'^2}\,\frac{v}{2} = 91.1876 \text{ GeV}/c^2. \tag{187}$$

The orthogonal, normalized gauge boson

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}} \left(g' A^3_\mu + g\, B_\mu \right) \tag{188}$$

remains massless. It is the photon. The fine-structure constant is

$$\alpha = \frac{e^2}{4\pi\hbar c} = 1/137.035\,999\,074(44) \approx 1/137.036. \tag{189}$$

Why do three gauge bosons become massive? Because there are three Goldstone bosons corresponding to three ways of moving $\langle 0|H|0\rangle$ without changing the Higgs potential. Why does one gauge boson stay massless? Because one linear combination of the generators of $SU_L(2) \otimes U(1)$ maps $\langle 0|H|0\rangle$ to zero, and so does not make an eigenstate of the gauge-boson mass matrix with eigenvalue zero.

In terms of these mass eigenstates, the original gauge bosons are

$$\begin{aligned}
A^1_\mu &= \frac{1}{\sqrt{2}} \left(W^+_\mu + W^-_\mu \right) \\
A^2_\mu &= \frac{1}{i\sqrt{2}} \left(W^-_\mu - W^+_\mu \right) \\
A^3_\mu &= \frac{1}{\sqrt{g^2 + g'^2}} \left(g' A_\mu + g Z_\mu \right) \\
B_\mu &= \frac{1}{\sqrt{g^2 + g'^2}} \left(g A_\mu - g' Z_\mu \right).
\end{aligned} \tag{190}$$

Thus, the covariant derivative for a fermion of $U(1)$ charge Y and coupling g to the $SU(2)_\ell$ gauge fields is

$$\begin{aligned}
D_\mu &= I \partial_\mu + i\, g \, \frac{\sigma^a}{2} \, A^a_\mu + i\, g'\, Y\, I\, B_\mu \\
&= I \partial_\mu + i\, g \, \left[\frac{\sigma_1}{2} \frac{1}{\sqrt{2}} \left(W^+_\mu + W^-_\mu \right) + \frac{\sigma_2}{2} \frac{1}{i\sqrt{2}} \left(W^-_\mu - W^+_\mu \right) + \frac{\sigma_3}{2} \frac{1}{\sqrt{g^2 + g'^2}} \left(g' A_\mu + g Z_\mu \right) \right] \\
&\quad + i\, g'\, Y \, \frac{I}{\sqrt{g^2 + g'^2}} \left(g A_\mu - g' Z_\mu \right) \\
&= I \partial_\mu + i\, \frac{g}{\sqrt{2}} \, \left(W^+_\mu T^+ + W^-_\mu T^- \right) + i \frac{1}{\sqrt{g^2 + g'^2}} Z_\mu \left(g^2 T^3 - g'^2 Y I \right) + i\, \frac{g g'}{\sqrt{g^2 + g'^2}} \, A_\mu \left(T^3 + Y I \right)
\end{aligned} \tag{191}$$

in which

$$T^\pm = \frac{1}{2} \left(\sigma_1 \pm i \sigma_2 \right). \tag{192}$$

We call the interaction strength or coupling constant of the photon A_μ

$$0 < e = \frac{g g'}{\sqrt{g^2 + g'^2}}, \tag{193}$$

and set the charge operator equal to

$$Q = T^3 + YI. \tag{194}$$

When acting on the doublet

$$E_\ell = \begin{pmatrix} \nu_e \\ e \end{pmatrix} \tag{195}$$

to which we assign $Y = -1/2$, the charge Q gives 0 as the charge of the neutrino and -1 as the charge of the electron. The photon-lepton term then is

$$\frac{gg'}{\sqrt{g^2 + g'^2}} A_\mu \left(T^3 + YI\right) E_\ell = e\, A_\mu \, \left(T^3 + YI\right) E_\ell = e\, A_\mu \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \nu_e \\ e \end{pmatrix} = -e\, A_\mu\, e \tag{196}$$

in which the first e is the absolute value of the charge (193) of the electron and the second is the field of the electron.

The **weak mixing angle** θ_w is defined by

$$\begin{pmatrix} Z \\ A \end{pmatrix} = \begin{pmatrix} \cos\theta_w & -\sin\theta_w \\ \sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} A^3 \\ B \end{pmatrix}. \tag{197}$$

Our equations (184 and 188) identify thse trigonometric values as

$$\cos\theta_w = \frac{g}{\sqrt{g^2 + g'^2}} \quad \text{and} \quad \sin\theta_w = \frac{g'}{\sqrt{g^2 + g'^2}}. \tag{198}$$

Since the charge is $Q = T^3 + YI$, we can use $Q - T^3$ instead of YI, so that the coupling to the Z is

$$\frac{1}{\sqrt{g^2 + g'^2}} Z_\mu \left[g^2 T^3 - g'^2 Y\right] = \frac{1}{\sqrt{g^2 + g'^2}} Z_\mu \left[(g^2 + g'^2) T^3 - g'^2 Q\right] \tag{199}$$

and

$$\frac{gg'}{\sqrt{g^2 + g'^2}} A_\mu \left(T^3 + Y\right) = \frac{gg'}{\sqrt{g^2 + g'^2}} A_\mu\, Q. \tag{200}$$

We also have

$$\frac{g^2+g'^2}{\sqrt{g^2+g'^2}} = \sqrt{g^2+g'^2} = \frac{g}{\cos\theta_w} \quad \text{and} \quad \frac{g'^2}{\sqrt{g^2+g'^2}} = \sqrt{g^2+g'^2}\,\frac{g'^2}{g^2+g'^2} = \frac{g}{\cos\theta_w}\,\sin^2\theta_w. \tag{201}$$

So the coupling to the Z is

$$\frac{1}{\sqrt{g^2+g'^2}} Z_\mu \left[(g^2+g'^2)T^3 - g'^2 Q\right] = \frac{g}{\cos\theta_w} Z_\mu \left(T^3 - \sin^2\theta_w\, Q\right) \tag{202}$$

and, if we set

$$e = g\,\sin\theta_w, \tag{203}$$

then the coupling to the photon A is

$$\frac{gg'}{\sqrt{g^2+g'^2}}\, A_\mu\, Q = g\,\sin\theta_w\, A_\mu\, Q = e\, A_\mu\, Q. \tag{204}$$

In these terms, the covariant derivative (191) is

$$\begin{aligned} D_\mu &= I\partial_\mu + i\,\frac{g}{\sqrt{2}}\,\left(W_\mu^+ T^+ + W_\mu^- T^-\right) + i\frac{1}{\sqrt{g^2+g'^2}} Z_\mu \left(g^2 T^3 - g'^2 Y\right) + i\,\frac{gg'}{\sqrt{g^2+g'^2}}\, A_\mu \left(T^3 + Y\right) \\ &= I\partial_\mu + i\,\frac{g}{\sqrt{2}}\,\left(W_\mu^+ T^+ + W_\mu^- T^-\right) + i\frac{g}{\cos\theta_w} Z_\mu \left(T^3 - \sin^2\theta_w\, Q\right) + ie\, A_\mu\, Q \end{aligned} \tag{205}$$

in which the generators $T^\pm$ and T^3 are those of the representation to which the fields they act on belong. When acting on **left-handed** fermions, they are half the Pauli matrices, $\boldsymbol{T} = \frac{1}{2}\boldsymbol{\sigma}$. When acting on **right-handed** fermions, they are zero, $\boldsymbol{T} = \boldsymbol{0}$. Since $g = e/\sin\theta_w$, the couplings involve one new parameter θ_w.

Our mass formulas (186 and 187) for the W and the Z show that their masses are related by

$$M_W = g\,\frac{v}{2} = \frac{g}{\sqrt{g^2+g'^2}}\,\sqrt{g^2+g'^2}\,\frac{v}{2} = \cos\theta_w\,\sqrt{g^2+g'^2}\,\frac{v}{2} = \cos\theta_w\, M_Z. \tag{206}$$

Experiments have determined the masses and shown that

$$\sin^2\theta_w = 0.231 \quad \text{or} \quad \theta_w = 0.233 \tag{207}$$

and that

$$v = 246.22 \text{ GeV}. \tag{208}$$

13.2 Quark and Lepton Interactions

The right-handed fermions u_r, d_r, e_r, and $\nu_{e,r}$ are singlets under $SU_L(2) \otimes U_Y(1)$. So they have $T^3 = 0$. The definition (194) of the charge Q

$$Q = T^3 + Y I \tag{209}$$

then implies that

$$Y_r = Q_r. \tag{210}$$

That is, $Y_{\nu_{e,r}} = 0$, $Y_{e_r} = -1$, $Y_{u_r} = 2/3$, and $Y_{d_r} = -1/3$.

The left-handed fermions are in doublets

$$E_\ell = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix} \quad \text{and} \quad Q_\ell = \begin{pmatrix} u \\ d \end{pmatrix} \tag{211}$$

with $T^3 = \pm 1/2$. So the choices $Y_E = -1/2$ and $Y_Q = 1/6$ and the definition (194) of the charge Q give the right charges:

$$Q E_\ell = Q \begin{pmatrix} \nu_e \\ e^- \end{pmatrix} = \begin{pmatrix} 0 \\ -1 e^- \end{pmatrix} \quad \text{and} \quad Q Q_\ell = Q \begin{pmatrix} u \\ d \end{pmatrix} = \begin{pmatrix} 2u/3 \\ -d/3 \end{pmatrix}. \tag{212}$$

Fermion-gauge-boson interactions are due to the covariant derivative (205) acting on either the left- or right-handed fields. On right-handed fermions, the covariant derivative is just

$$\begin{aligned} D^r_\mu &= I\partial_\mu + i\frac{g}{\cos\theta_w} Z_\mu \left(-\sin^2\theta_w\, Q\right) + ie\, A_\mu\, Q = I\partial_\mu - ig\,\frac{\sin^2\theta_w}{\cos\theta_w}\, Z_\mu\, Q + ie\, A_\mu\, Q \\ &= I\partial_\mu - ie\,\frac{\sin\theta_w}{\cos\theta_w}\, Z_\mu\, Q + ie\, A_\mu\, Q = I\partial_\mu - ie\, \tan\theta_w\, Z_\mu\, Q + ie\, A_\mu\, Q. \end{aligned} \tag{213}$$

So the covariant derivative of a neutral right-handed fermion is just the ordinary derivative.

On left-handed fermions, the covariant derivative is

$$\begin{aligned} D^\ell_\mu &= I\partial_\mu + i\,\frac{g}{\sqrt{2}}\,\left(W^+_\mu T^+ + W^-_\mu T^-\right) + i\frac{g}{\cos\theta_w} Z_\mu \left(T^3 - \sin^2\theta_w\, Q\right) + ie\,A_\mu\, Q \\ &= I\partial_\mu + i\,\frac{e}{\sqrt{2}\ \sin\theta_w}\,\left(W^+_\mu T^+ + W^-_\mu T^-\right) + i\frac{e}{\cos\theta_w} Z_\mu \left(\frac{T^3}{\sin\theta_w} - \sin\theta_w\, Q\right) + ie\,A_\mu\, Q. \end{aligned} \tag{214}$$

For the first **family** or **generation** of quarks and leptons, the kinetic action density is

$$L_k = \, - \overline{E}_\ell \not{D}^\ell E_\ell - \overline{E}_r \not{D}^r E_r - \overline{Q}_\ell \not{D}^\ell Q_\ell - \overline{Q}_r \not{D}^r Q_r \tag{215}$$

in which $\not{D} \equiv \gamma^\mu D_\mu$. The 4×4 matrix $\gamma_5 = \gamma^5$ plays the role of a fifth (spatial) gamma matrix $\gamma^4 = \gamma_5$ in 5-dimensional space-time in the sense that the anticommutator

$$\{\gamma^a, \gamma^b\} = 2\eta^{ab} \tag{216}$$

in which η is the 5×5 diagonal matrix with $\eta^{00} = -1$ and $\eta^{aa} = 1$ for $a = 1, 2, 3, 4$. In Weinberg's notation, γ_5 is

$$\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{217}$$

The combinations

$$P_\ell = \frac{1}{2}\,(1 + \gamma_5) \quad \text{and} \quad P_r = \frac{1}{2}\,(1 - \gamma_5) \tag{218}$$

are projection operators onto the left- and right-handed fields. That is,

$$P_\ell Q = Q_\ell \quad \text{and} \quad P_\ell Q_\ell = Q_\ell \tag{219}$$

with a similar equation for P_r. We can write L_k as

$$\begin{aligned} L_k =& \, - \overline{E}\not{D}^\ell \tfrac{1}{2}(1+\gamma_5)E - \overline{E}\not{D}^r \tfrac{1}{2}(1-\gamma_5)E - \overline{Q}\not{D}^\ell \tfrac{1}{2}(1+\gamma_5)Q - \overline{Q}\not{D}^r \tfrac{1}{2}(1-\gamma_5)Q \\ =& \, - \frac{1}{2}\left[\overline{E}\not{D}^\ell(1+\gamma_5)E + \overline{E}\not{D}^r(1-\gamma_5)E + \overline{Q}\not{D}^\ell(1+\gamma_5)Q + \overline{Q}\not{D}^r(1-\gamma_5)Q\right]. \end{aligned} \tag{220}$$

Homework set 4, problem 1: Show that

$$\overline{E}_\ell \not{D}^\ell E_\ell = [\tfrac{1}{2}(1+\gamma_5)E]^\dagger i\gamma^0 \not{D}^\ell \tfrac{1}{2}(1+\gamma_5)E = \overline{E} \not{D}^\ell \tfrac{1}{2}(1+\gamma_5)E. \tag{221}$$

Recall that in Weinberg's notation

$$\overline{\psi} = \psi^\dagger i\gamma^0 = \psi^\dagger \beta = \psi^\dagger \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \tag{222}$$

in which I is the 2×2 identity matrix.

13.3 Quark and Lepton Masses

The Higgs mechanism also gives masses to the fermions, but somewhat arbitrarily. Dirac's action density (8) has as its mass term

$$-m\,\overline{\psi}\psi = -\,im\,\psi^\dagger\gamma^0\psi = -im\,\psi^\dagger\gamma^0(P_\ell + P_r)\psi = -im\,\psi^\dagger\gamma^0(P_\ell^2 + P_r^2)\psi. \tag{223}$$

Since $\{\gamma^0, \gamma^5\} = 0$, this mass term is

$$\begin{aligned} -m\,\overline{\psi}\psi &= -\,im\,\psi^\dagger P_r\gamma^0 P_\ell\psi - im\,\psi^\dagger P_\ell\gamma^0 P_r\psi = -im\,(P_r\psi)^\dagger\gamma^0 P_\ell\psi - im\,(P_\ell\psi)^\dagger\gamma^0 P_r\psi \\ &= -\,im\,\psi_r^\dagger\gamma^0\psi_\ell - im\,\psi_\ell^\dagger\gamma^0\psi_r = -m\,\overline{\psi}_r\psi_\ell - m\,\overline{\psi}_\ell\psi_r. \end{aligned} \tag{224}$$

Incidentally, because the fields ψ_ℓ and ψ_r are independent, we can redefine them

$$\psi'_\ell = e^{i\theta}\,\psi_\ell \tag{225}$$

$$\psi'_r = e^{i\phi}\,\psi_r \tag{226}$$

at will. Such a redefinition changes the mass term to

$$-m'\,\overline{\psi}_r\psi_\ell - m'^*\,\overline{\psi}_\ell\psi_r = -m\,e^{i(\theta-\phi)}\,\overline{\psi}_r\psi_\ell - m\,e^{-i(\theta-\phi)}\,\overline{\psi}_\ell\psi_r. \tag{227}$$

So the phase of a Dirac mass term has no significance.

The definition (222) of $\overline{\psi}$ shows that the Dirac mass term is

$$-m\overline{\psi}\psi = -m\,\psi^\dagger\beta\psi = -m\left(\psi_\ell^\dagger, \psi_r^\dagger\right)\begin{pmatrix}0 & I\\ I & 0\end{pmatrix}\begin{pmatrix}\psi_\ell\\ \psi_r\end{pmatrix} = -m\left(\psi_\ell^\dagger\psi_r + \psi_r^\dagger\psi_\ell\right). \tag{228}$$

These mass terms are invariant under the Lorentz transformations

$$\begin{aligned}\psi_\ell' &= \exp(-\boldsymbol{z}\cdot\boldsymbol{\sigma})\,\psi_\ell\\ \psi_r' &= \exp(\boldsymbol{z}^*\cdot\boldsymbol{\sigma})\,\psi_r\end{aligned} \tag{229}$$

because

$$\begin{aligned}\psi_\ell'^\dagger\,\psi_r' &= \psi_\ell^\dagger\exp(-\boldsymbol{z}^*\cdot\boldsymbol{\sigma})\exp(\boldsymbol{z}^*\cdot\boldsymbol{\sigma})\,\psi_r = \psi_\ell^\dagger\,\psi_r\\ \psi_r'^\dagger\,\psi_\ell' &= \psi_r^\dagger\exp(\boldsymbol{z}\cdot\boldsymbol{\sigma})\exp(-\boldsymbol{z}\cdot\boldsymbol{\sigma})\,\psi_\ell = \psi_r^\dagger\,\psi_\ell.\end{aligned} \tag{230}$$

They are *not* invariant under rigid, let alone local, $SU(2)_\ell$ transformations. But we can make them invariant by using the Higgs field $H(x)$. For instance, the quantity $\overline{Q}_\ell H d_r$ is invariant under local $SU(2)_\ell$ transformations. In unitary gauge, its mean value in the vacuum is

$$\langle 0|\overline{Q}_\ell H d_r|0\rangle = \frac{1}{\sqrt{2}}\,v\,\langle 0|d_\ell^\dagger d_r|0\rangle. \tag{231}$$

So the term

$$-\,c_d\,\overline{Q}_\ell H d_r - c_d^*\,\overline{d}_r H^\dagger Q_\ell \tag{232}$$

is invariant, and in the vacuum it is

$$\langle 0| - c_d\,\overline{Q}_\ell H d_r - c_d^*\,\overline{d}_r H^\dagger Q_\ell|0\rangle = \frac{1}{\sqrt{2}}\,v\,\langle 0| - c_d\,d_\ell^\dagger d_r - c_d^*\,d_r^\dagger d_\ell|0\rangle \tag{233}$$

which gives to the d quark the mass

$$m_d = \frac{|c_d|}{\sqrt{2}}\,v. \tag{234}$$

Note that we must add one new parameter c_d to get one new mass m_d. This parameter c_d is complex in general, but the mass m_d depends only upon the absolute value and not upon its phase of c_d.

Similarly, the term

$$- c_e \, \overline{E}_\ell H e_r - c_e^* \, \overline{e}_r H^\dagger E_\ell \tag{235}$$

is invariant, and in the vacuum it is

$$\langle 0| - c_e \, \overline{E}_\ell H e_r - c_e^* \, \overline{e}_r H^\dagger Q_\ell |0\rangle = \frac{1}{\sqrt{2}} \, v \, \langle 0| - c_e \, e_\ell^\dagger e_r - c_e^* \, e_r^\dagger e_\ell |0\rangle \tag{236}$$

which gives to the electron the mass

$$m_e = \frac{|c_e|}{\sqrt{2}} \, v. \tag{237}$$

Again, we must add one new (complex) parameter c_e to get one new mass m_e.

The mass of the up quark requires a new trick. The Higgs field H transforms under $SU(2)_\ell \times U(1)$ as

$$H'(x) = \exp \left[i \, g \, \frac{\sigma^a}{2} \, \alpha^a(x) + i \, g' \, \frac{I}{2} \, \beta(x) \right] \, H(x). \tag{238}$$

If for clarity, we leave aside the $U(1)$ part for the moment, then the Higgs field H transforms under $SU(2)_\ell$ as

$$H'(x) = \exp \left[i \, g \, \frac{\sigma^a}{2} \, \alpha^a(x) \right] \, H(x). \tag{239}$$

Let us use H^* to be the complex column vector whose components are $H_1^\dagger$ and $H_2^\dagger$. How does $\sigma_2 H^*$ transform under $SU(2)_\ell$? Suppressing our explicit mention of the space-time dependence and using the asterisk to mean hermitian conjugation when applied to operators but complex conjugation when applied to matrices and vectors, we have, since σ_2 is imaginary with $\sigma_2^2 = I$ while σ_1 and σ_3 are real,

$$\begin{aligned} (\sigma_2 H^*)' &= \sigma_2 \left[\exp \left(i \, g \, \frac{\sigma^a}{2} \, \alpha^a \right) \, H \right]^* = \sigma_2 \exp \left(-i \, g \, \frac{\sigma_a^*}{2} \, \alpha^a \right) \, H^* \\ &= \sigma_2 \exp \left(-i \, g \, \frac{\sigma_a^*}{2} \, \alpha^a \right) \, \sigma_2 \sigma_2 \, H^* = \exp \left(i \, g \, \frac{\sigma^a}{2} \, \alpha^a \right) \, \sigma_2 \, H^*. \end{aligned} \tag{240}$$

Thus, the term

$$-c_u\overline{Q}_\ell\sigma_2 H^* u_r - c_u^*\overline{u}_r H^\mathsf{T}\sigma_2 Q_\ell \tag{241}$$

is invariant under $SU(2)_\ell$. In the vacuum of the unitary gauge, it is

$$\langle 0| - c_u\overline{Q}_\ell\sigma_2 H^* u_r - c_u^*\overline{u}_r H^\mathsf{T}\sigma_2 Q_\ell|0\rangle = \frac{1}{\sqrt{2}}\,v\langle 0|ic_u\overline{u}_\ell u_r - ic_u^*\overline{u}_r u_\ell|0\rangle = \frac{1}{\sqrt{2}}\,v\langle 0|ic_u u_\ell^\dagger u_r - ic_u^* u_r^\dagger u_\ell|0\rangle \tag{242}$$

which gives the up quark the mass

$$m_u = \frac{|c_u|}{\sqrt{2}}\,v. \tag{243}$$

But there are three families of generations of quarks and leptons on which the gauge fields act simply:

$$F_1 = \begin{pmatrix} u \\ d \\ \nu_e \\ e \end{pmatrix}', \quad F_2 = \begin{pmatrix} c \\ s \\ \nu_\mu \\ \mu \end{pmatrix}', \quad \text{and} \quad F_3 = \begin{pmatrix} t \\ b \\ \nu_\tau \\ \tau \end{pmatrix}'. \tag{244}$$

The quark and lepton flavor families are

$$\begin{aligned} Q_1' &= \begin{pmatrix} u \\ d \end{pmatrix}', \quad Q_2' = \begin{pmatrix} c \\ s \end{pmatrix}', \quad \text{and} \quad Q_3' = \begin{pmatrix} t \\ b \end{pmatrix}'; \\ E_1' &= \begin{pmatrix} \nu_e \\ e \end{pmatrix}', \quad E_2' = \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}', \quad \text{and} \quad E_3' = \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}'. \end{aligned} \tag{245}$$

These are called the **flavor** eigenstates or more properly flavor eigenfields, designated here with primes. They are the ones on which the $W^\pm$ act simply. The weak interactions use $W_\mu^- T^-$ to map the flavor *up* fields $u_1' = u'$, $u_2' = c'$, $u_3' = t'$ into the flavor *down* fields $d_1' = d'$, $d_2' = s'$, $d_3' = b'$, and $W_\mu^+ T^+$ to map the flavor *down* fields d_i' into the flavor *up* fields u_i'.

The action density

$$\sum_{i,j=1}^{3} -c_{dij}\,\overline{Q}_{\ell i}' H d_{rj}' - c_{dij}^*\,\overline{d}_{rj}' H^\dagger Q_{\ell i}' \tag{246}$$

gives for the d', s', and b' quarks the mixed mass terms

$$\frac{v}{\sqrt{2}} \sum_{i,j=1}^{3} -c_{dij}\, \overline{d}'_{\ell i} d'_{rj} - c^*_{dij}\, \overline{d}'_{rj} d'_{\ell i} = \frac{v}{\sqrt{2}} \sum_{i,j=1}^{3} -c_{dij}\, d'^{\dagger}_{\ell i} d'_{rj} - c^*_{dij}\, d'^{\dagger}_{rj} d'_{\ell i}. \tag{247}$$

The 3×3 mass matrix M_d with entries

$$[M_d]_{ij} = \frac{v}{\sqrt{2}}\, c_{dij} \tag{248}$$

need have no special properties. It need not be hermitian because for each i and j, the term (247) is hermitian. But every 3×3 complex matrix has a **singular-value decomposition**

$$M_d = L_d \Sigma_d R_d^{\dagger} \tag{249}$$

in which L_d and R_d are 3×3 unitary matrices, and Σ_d is a 3×3 diagonal matrix with nonincreasing positive singular values on its main diagonal.

The singular value decomposition works for any $N \times M$ (real or) complex matrix. Every complex $M \times N$ rectangular matrix A is the product of an $M \times M$ unitary matrix U, an $M \times N$ rectangular matrix Σ that is zero except on its main diagonal which consists of its nonnegative singular values S_k, and an $N \times N$ unitary matrix $V^{\dagger}$

$$A = U \, \Sigma \, V^{\dagger}. \tag{250}$$

This singular-value decomposition is a key theorem of matrix algebra. One can use the *Matlab* command "[U,S,V] = svd(A)" to perform the svd $A = USV^{\dagger}$.

The singular values of Σ_d are the masses m_b, m_s, and m_d:

$$\Sigma_d = \begin{pmatrix} m_b & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_d \end{pmatrix}. \tag{251}$$

So the mass eigenfields of the left and right down-quark fields are

$$d_{ri} = R^{\dagger}_{dij} d'_{rj} \quad \text{and} \quad d^{\dagger}_{\ell i} = d'^{\dagger}_{\ell j} L_{dji} \quad \text{or} \quad d_{\ell i} = L^{\dagger}_{dij} d_{\ell j}. \tag{252}$$

The inverse relations are

$$d'_{ri} = R_{dij} d_{rj} \quad \text{and} \quad d'^{\dagger}_{\ell i} = d^{\dagger}_{\ell j} L^{\dagger}_{dji} \quad \text{or} \quad d'_{\ell i} = L_{dij} d_{\ell j} \tag{253}$$

or in matrix notation

$$\boldsymbol{d}'_r = R_d\, \boldsymbol{d}_r, \quad \boldsymbol{d}'^{\dagger}_r = \boldsymbol{d}^{\dagger}_r\, R^{\dagger}_d, \quad \boldsymbol{d}'^{\dagger}_\ell = \boldsymbol{d}^{\dagger}_\ell L^{\dagger}_d, \quad \text{and} \quad \boldsymbol{d}'_\ell = L_d\, \boldsymbol{d}_\ell \tag{254}$$

in which

$$\boldsymbol{d}_\ell = \begin{pmatrix} b \\ s \\ d \end{pmatrix}_\ell \tag{255}$$

are the down-quark fields of definite masses.

Similarly, the *up* quark action density

$$\sum_{i,j=1}^{3} -c_{uij} \overline{Q}'_{\ell i} \sigma_2 H^* u'_{rj} - c^*_{uij} \overline{u}'_{rj} H^{\mathsf{T}} \sigma_2 Q'_{\ell i} \tag{256}$$

gives for the three known families the mixed mass terms

$$\frac{iv}{\sqrt{2}} \sum_{i,j=1}^{3} c_{uij} \overline{u}'_{\ell i}\, u'_{rj} - c^*_{uij} \overline{u}'_{rj}\, u'_{\ell i} = \frac{iv}{\sqrt{2}} \sum_{i,j=1}^{3} c_{uij} u'^{\dagger}_{\ell i}\, u'_{rj} - c^*_{uij} u'^{\dagger}_{rj}\, u'_{\ell i}. \tag{257}$$

The 3×3 mass matrix M_u with entries

$$[M_u]_{ij} = \frac{iv}{\sqrt{2}}\, c_{uij} \tag{258}$$

need have no special properties. It need not be hermitian because for each i and j, the term (257) is hermitian. But every 3×3 complex matrix M_u has a singular value decomposition

$$M_u = L_u \Sigma_u R^{\dagger}_u \tag{259}$$

in which L_u and R_u are 3×3 unitary matrices, and Σ_u is a 3×3 diagonal matrix with nonincreasing positive singular values on its main diagonal. These singular values are the masses m_t, m_c, and m_u:

$$\Sigma_u = \begin{pmatrix} m_t & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_u \end{pmatrix}. \tag{260}$$

So the mass eigenfields of the left and right up-quark fields are

$$u_{ri} = R^\dagger_{uij} u'_{rj} \quad \text{and} \quad u^\dagger_{\ell i} = u'^\dagger_{\ell j} L_{uji} \quad \text{or} \quad u_{\ell i} = L^\dagger_{uij} u'_{\ell j}. \tag{261}$$

The inverse relations are

$$u'_{ri} = R_{uij} u_{rj} \quad \text{and} \quad u'^\dagger_{\ell i} = u^\dagger_{\ell j} L^\dagger_{uji} \quad \text{or} \quad u'_{\ell i} = L_{uij} u_{\ell j} \tag{262}$$

or in matrix notation

$$\boldsymbol{u}'_r = R_u \, \boldsymbol{u}_r, \quad \boldsymbol{u}'^\dagger_r = \boldsymbol{u}^\dagger_r \, R^\dagger_u \quad \boldsymbol{u}'^\dagger_\ell = \boldsymbol{u}^\dagger_\ell L^\dagger_u, \quad \text{and} \quad \boldsymbol{u}'_\ell = L_u \boldsymbol{u}_\ell \tag{263}$$

in which

$$\boldsymbol{u}_\ell = \begin{pmatrix} t \\ c \\ u \end{pmatrix}_\ell \tag{264}$$

are the up-quark fields of definite masses.

13.4 The CKM Matrix

We will use the labels u, c, t and d, s, b for the states that are eigenstates of the quadratic part of the hamiltonian after the Higgs mechanism has given a mean value to the real part of the neutral Higgs boson in the unitary gauge. The u, c, t quarks have the same charge $2e/3 > 0$ and the same $T^3 = 1/2$, so they all have the same electroweak interactions. Similarly, the d, s, b quarks have the same charge $-e/3 < 0$ and the same $T^3 = -1/2$, so they also all have the same electroweak interactions.

The right-handed covariant derivative (213)

$$D^r_\mu = I\partial_\mu - ie\,\tan\theta_w\,Z_\mu\,Q + ie\,A_\mu\,Q \tag{265}$$

just sends the fields of these mass eigenstates into themselves multiplied by their charge and either a Z or a photon. That is,

$$\begin{aligned}\overline{\boldsymbol{u}}'_r D^r_\mu \boldsymbol{u}'_r &= \overline{\boldsymbol{u}}_r R^\dagger_u D^r_\mu R_u \boldsymbol{u}_r = \overline{\boldsymbol{u}}_r D^r_\mu \boldsymbol{u}_r \\ \overline{\boldsymbol{d}}'_r D^r_\mu \boldsymbol{d}'_r &= \overline{\boldsymbol{d}}_r R^\dagger_d D^r_\mu R_d \boldsymbol{d}_r = \overline{\boldsymbol{d}}_r D^r_\mu \boldsymbol{d}_r\end{aligned} \tag{266}$$

In these terms, the interactions of the Z and the photon with the right-handed fields are diagonal both in mass and in flavor.

But the left-handed covariant derivative (214) is

$$D^\ell_\mu = I\partial_\mu + i\,\frac{e}{\sqrt{2}\,\sin\theta_w}\,\left(W^+_\mu T^+ + W^-_\mu T^-\right) + i\frac{e}{\cos\theta_w} Z_\mu \left(\frac{T^3}{\sin\theta_w} - \sin\theta_w\,Q\right) + ie\,A_\mu\,Q. \tag{267}$$

So we have

$$\begin{pmatrix}\overline{\boldsymbol{u}}'_\ell & \overline{\boldsymbol{d}}'_\ell\end{pmatrix} D^\ell_\mu \begin{pmatrix}\boldsymbol{u}'_\ell \\ \boldsymbol{d}'_\ell\end{pmatrix} = \begin{pmatrix}\overline{\boldsymbol{u}}_\ell L^\dagger_u & \overline{\boldsymbol{d}}_\ell L^\dagger_d\end{pmatrix} D^\ell_\mu \begin{pmatrix}L_u\,\boldsymbol{u}_\ell \\ L_d\,\boldsymbol{d}_\ell\end{pmatrix}. \tag{268}$$

Some of the unitary matrices just give unity, $L^\dagger_u L_u = I$ and $L^\dagger_d L_d = I$ like $R^\dagger_u R_u = I$ and $R^\dagger_d R_d = I$ in the right-handed covariant derivatives (266). Thus the interactions of the Z and the photon with the both the right-handed fields and with the left-handed fields are diagonal both in mass and in flavor. The Z and the photon do not mediate top-to-charm or charm-to-up or $\mu^- \to e^- + \gamma$ decays. **There are no flavor-changing neutral-currents.**

The only changes are in the nonzero parts of $T^\pm$ which become

$$T^+_{\text{CKM}} = \begin{pmatrix}0 & L^\dagger_u L_d \\ 0 & 0\end{pmatrix} \equiv \begin{pmatrix}0 & V \\ 0 & 0\end{pmatrix} \quad \text{and} \quad T^-_{\text{CKM}} = \begin{pmatrix}0 & 0 \\ L^\dagger_d L_u & 0\end{pmatrix} \equiv \begin{pmatrix}0 & 0 \\ V^\dagger & 0\end{pmatrix} \tag{269}$$

in which the unitary matrix $V = L_u^\dagger L_d$ is the CKM matrix (Nicola Cabibbo, Makoto Kobayashi, and Toshihide Maskawa). The left-handed covariant derivative on the mass eigenfields then is

$$D_\mu^\ell = I\partial_\mu + i\,\frac{e}{\sqrt{2}\,\sin\theta_w}\,\left(W_\mu^+ T_{\text{CKM}}^+ + W_\mu^- T_{\text{CKM}}^-\right) + i\frac{e}{\cos\theta_w} Z_\mu \left(\frac{T^3}{\sin\theta_w} - \sin\theta_w\, Q\right) + ie\,A_\mu\, Q. \tag{270}$$

It has a second part that acts more or less like the right-handed covariant derivative, but the first part uses $W_\mu^- T^-$ to map the *up* fields u, c, t into linear combinations of the *down* fields d, s, b and $W_\mu^+ T^+$ to map the *down* fields into linear combinations of the *up* fields. The $W_\mu^\pm$ terms are sensitive to the CKM matrix $V = L_u^\dagger L_d$. We write them suggestively as

$$\begin{pmatrix} u & c & t & d & s & b \end{pmatrix}^\dagger \begin{pmatrix} 0 & VW_\mu^+ \\ V^\dagger W_\mu^- & 0 \end{pmatrix} \begin{pmatrix} u \\ c \\ t \\ d \\ s \\ b \end{pmatrix} = \left(\begin{matrix} u & c & t & V \begin{pmatrix} d \\ s \\ b \end{pmatrix} \end{matrix} \right)^\dagger \begin{pmatrix} 0 & W_\mu^+ \\ W_\mu^- & 0 \end{pmatrix} \begin{pmatrix} u \\ c \\ t \\ V \begin{pmatrix} d \\ s \\ b \end{pmatrix} \end{pmatrix}. \tag{271}$$

By choosing the phases of the six quark fields, that is, $u(x) \to e^{i\theta_u} u(x) \ldots b(x) \to e^{i\theta_b} b(x)$, one may make the CKM matrix $L_u^\dagger L_d$ real apart from a single phase. The existence of that phase probably is the cause of most of the breakdown of CP invariance that Fitch and Cronin and others have observed since 1964. The magnitudes of the elements of the CKM matrix V are

$$V = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97427 & 0.22536 & 0.00355 \\ 0.22522 & 0.97343 & 0.0414 \\ 0.00886 & 0.0405 & 0.99914 \end{pmatrix}. \tag{272}$$

Although there is only one phase $\exp(i\delta)$ in the CKM matrix V, the experimental constraints on this phase often are expressed in terms of the angles α, β, and γ defined as

$$\begin{aligned} \alpha &= \arg\left[-V_{td}V_{tb}^*/\left(V_{ud}V_{ub}^*\right)\right] \\ \beta &= \arg\left[-V_{cd}V_{cb}^*/\left(V_{td}V_{tb}^*\right)\right] \\ \gamma &= \arg\left[-V_{ud}V_{ub}^*/\left(V_{cd}V_{cb}^*\right)\right]. \end{aligned} \tag{273}$$

If V is unitary, then $\alpha + \beta + \gamma = 180°$. From $B \to \pi\pi$, $\rho\pi$, and $\rho\rho$ decays, the limits on the angle α are roughly

$$\alpha = (85.4 \pm 4)°. \tag{274}$$

From $B^\pm \to DK^\pm$ decays, the limits on the angle γ are roughly

$$\gamma = (68.0 \pm 8)°. \tag{275}$$

So the angle β is about 26.6°.

One of the quark-Higgs interactions is

$$\begin{aligned} -c_{dij}\overline{Q}'_{\ell i} H d'_{rj} &= -\frac{\sqrt{2}}{v}\overline{\boldsymbol{Q}}'_\ell M_d \boldsymbol{d}'_r H = -\frac{\sqrt{2}}{v}\overline{\boldsymbol{Q}}'_\ell L_d \Sigma_d R_d^\dagger \boldsymbol{d}'_r H \\ &= -\frac{\sqrt{2}}{v}\overline{\boldsymbol{Q}}_\ell \Sigma_d \boldsymbol{d}_r H = -\frac{\sqrt{2}}{v}\overline{\boldsymbol{Q}}_\ell \begin{pmatrix} 0 \\ (v+h)/\sqrt{2} \end{pmatrix} \Sigma_d \boldsymbol{d}_r \\ &= -\overline{\boldsymbol{d}}_\ell \left(1 + \frac{h}{v}\right) \Sigma_d \boldsymbol{d}_r = -m_{di}\overline{d}_{\ell i}\left(1 + \frac{h}{v}\right) d_{ri}. \end{aligned} \tag{276}$$

A similar term describes the coupling of the *up* quarks to the Higgs

$$-m_{ui}\overline{u}_{\ell i}\left(1 + \frac{h}{v}\right) u_{ri}. \tag{277}$$

Thus, the rate of quark-antiquark to Higgs is proportional to the mass of the quark in the standard model.

14 Lepton Masses

We can treat the leptons just like the quarks. The *up* leptons are the flavor neutrinos ν'_e, ν'_μ, and ν'_τ, and the *down* leptons are the flavor charged leptons e', μ', and τ'. The action density

$$\sum_{i,j=1}^{3} -c_{eij}\,\overline{E}'_{\ell i} H e'_{rj} - c^*_{eij}\,\overline{e}'_{rj} H^\dagger E'_{\ell i} \tag{278}$$

gives for the e', μ', and τ' the mixed mass terms

$$\frac{v}{\sqrt{2}} \sum_{i,j=1}^{3} -c_{eij}\, \overline{e}'_{\ell i} e'_{rj} - c^*_{eij}\, \overline{e}'_{rj} e'_{\ell i}. \tag{279}$$

The 3×3 mass matrix M_e with entries

$$[M_e]_{ij} = \frac{v}{\sqrt{2}}\, c_{eij} \tag{280}$$

has a singular value decomposition

$$M_e = L_e \Sigma_e R_e^\dagger \tag{281}$$

in which L_e and R_e are 3×3 unitary matrices, and Σ_e is a 3×3 diagonal matrix with nonincreasing positive singular values m_τ, m_μ, and m_e on its main diagonal.

14.1 Before and After Symmetry Breaking

Before spontaneous symmetry breaking, all the fields of the standard model are massless, and the local symmetry under $SU(2)_\ell \otimes U(1)$ is exact. Under these gauge transformations, the left-handed electron and neutrino fields are rotated among themselves. If e'_ℓ is a linear combination of itself and of $\nu'_{e\ell}$, then these two fields, e'_ℓ and $\nu'_{e\ell}$, must be of the same kind. The left-handed electron field is a Dirac field. Thus, the left-handed neutrino field also must be a Dirac field. This makes sense because before symmetry breaking, all the fields are massless, and so there is no problem combining two Majorana fields of the same mass, namely zero, into one Dirac field. Thus, there are three flavor Dirac neutrino fields $\nu'_{e\ell}$, $\nu'_{\mu\ell}$, and $\nu'_{\tau\ell}$.

A massless left-handed neutrino field ν'_ℓ satisfies the two-component Dirac equation

$$\left(\partial_0 I - \boldsymbol{\nabla} \cdot \boldsymbol{\sigma}\right) \nu'_\ell(x) = 0 \tag{282}$$

which in momentum space is

$$\left(E + \boldsymbol{p} \cdot \boldsymbol{\sigma}\right) \nu'_\ell(p) = 0. \tag{283}$$

Since the angular momentum is $\boldsymbol{J} = \boldsymbol{\sigma}/2$, and $E = |\boldsymbol{p}|$, we have

$$\hat{\boldsymbol{p}} \cdot \boldsymbol{J}\, \nu'_\ell(p) = -\frac{1}{2}\, \nu'_\ell(p). \tag{284}$$

The left-handed neutrino field ν'_ℓ annihilates neutrinos of negative helicity and creates antineutrinos of positive helicity.

Since the neutrinos are massive, there may be right-handed neutrino fields. As for the *up* quarks, we can use them to make an action density

$$\sum_{i,j=1}^{3} -c_{\nu ij} \overline{E}'_{\ell i} \sigma_2 H^* \nu'_{rj} - c^*_{\nu ij} \overline{\nu}'_{rj} H^\mathsf{T} \sigma_2 E'_{\ell i} \tag{285}$$

that is invariant under $SU(2)_\ell \otimes U(1)$ and that gives for the neutrinos the mixed mass terms

$$\sum_{i,j=1}^{3} \frac{iv}{\sqrt{2}} \left(c_{\nu ij} \overline{\nu}'_{\ell i} \nu'_{rj} - c^*_{\nu ij} \overline{\nu}'_{rj} \nu'_{\ell i} \right). \tag{286}$$

The 3×3 mass matrix M_ν with entries

$$[M_\nu]_{ij} = \frac{iv}{\sqrt{2}}\, c_{\nu ij} \tag{287}$$

has a singular value decomposition

$$M_\nu = L_\nu \Sigma_\nu R^\dagger_\nu \tag{288}$$

in which L_ν and R_ν are 3×3 unitary matrices, and Σ_ν is a 3×3 diagonal matrix with nonincreasing positive singular values m_{ν_τ}, m_{ν_μ}, and m_{ν_e} on its main diagonal (here, I have assumed that the neutrino masses mimic those of the charged leptons and quarks, rising with family number). The neutrino CKM matrix then would be $L^\dagger_\nu L_e$, but since we are accustomed to treating the charged leptons as flavor and mass eigenfields, we apply the neutrino CKM matrix to the neutrinos rather than to the charged leptons. Thus the neutrino CKM matrix is

$$V_\nu = L^\dagger_e L_\nu. \tag{289}$$

By choosing the phases of the six lepton fields, we can make the neutrino CKM matrix real except for CP-breaking phases. If the neutrinos are Dirac fields, then there is one such phase; if not, there are three.

So far, I have assumed that the mass terms for the neutrinos are the usual Dirac mass terms. However, the right-handed Majorana neutrino fields $\boldsymbol{\nu}'_r$ are not affected by the $SU(2)_\ell \otimes U(1)$.

Note that a gauge transformation between e and ν_e rotates the operators $a(p,s,e)$ and $a(p,s,\nu_e)$ into each other. This rotation makes sense only when the two particles have the same mass. In the standard model, such a gauge transformation makes sense only before symmetry breaking when all the particles are massless. Moreover, only when the particles are massless can one say that they are left- or right-handed. While the particles are massless, the operator $a(p,-)$ annihilates a particle of negative helicity and occurs only in a left-handed field, while the operator $a(p,+)$ annihilates a particle of positive helicity and occurs only in a right-handed field. But when the particles are massive, the operator $a(p,\frac{1}{2})$ annihilates a particle that is spin up and occurs in both left-handed and right-handed fields. So a symmetry transformation that acted on the operator $a(p,\frac{1}{2})$ would change both left-handed and right-handed fields.

The left-handed fields of the neutrino and electron are

$$\nu_{e,\ell}(x) = \int u(p,-)\frac{a_1(p,-,\nu_e)+ia_2(p,-,\nu_e)}{\sqrt{2}}e^{ipx} + v(p,+)\frac{a_1^\dagger(p,+,\nu_e)+ia_2^\dagger(p,+,\nu_e)}{\sqrt{2}}e^{-ipx}\frac{d^3p}{(2\pi)^{3/2}} \tag{290}$$

$$e_\ell(x) = \int u(p,-)\frac{a_1(p,-,e)+ia_2(p,-,e)}{\sqrt{2}}e^{ipx} + v(p,+)\frac{a_1^\dagger(p,+,e)+ia_2^\dagger(p,+,e)}{\sqrt{2}}e^{-ipx}\frac{d^3p}{(2\pi)^{3/2}} \tag{291}$$

where $(p,-,\nu_e)$ means momentum p, spin down, and electron flavor, and $(p,+,\nu_e)$ means momentum p, spin up, and electron flavor. These fields satisfy equations like (282–284) apart from their interactions with other fields. Since a gauge transformation maps the fields $\nu_{e,\ell}(x)$ and $e_\ell(x)$ into each other, we know that when all the fields are massless, before symmetry breaking, there are (for each momentum) at least two neutrino and antineutrino states

$$\frac{1}{\sqrt{2}}\left[a_1^\dagger(p,-,\nu_e) - ia_2^\dagger(p,-,\nu_e)\right]|0\rangle \tag{292}$$

$$\frac{1}{\sqrt{2}}\left[a_1^\dagger(p,+,\nu_e) + ia_2^\dagger(p,+,\nu_e)\right]|0\rangle \tag{293}$$

for each of the three flavors, $f = e, \mu, \tau$. So there are at least six neutrino (and antineutrino) states.

The right-handed electron field exists and interacts with gauge bosons and other fields. So there are 12 electron states $a_i^\dagger(p, \pm, e_f)|0\rangle$ for $i = 1$ and 2 and for the three flavors, $f = e, \mu, \tau$. We don't know yet whether a right-handed neutrino field exists or interacts with other fields. So there may be only 6 neutrino states or as many as 12.

Neutrino oscillations tell us that neutrinos have masses. If there are 12 neutrino states, then there can be three massive Dirac neutrinos analagous to the e, μ, and τ or six massive Majorana neutrinos or some intermediate combination. If there are only 6, then there can be 3 massive Majorana neutrinos.

The Majorana mass terms for the right-handed neutrino fields are

$$\sum_{ij=1}^{6} \frac{1}{2} \left[i m_{ij} {\nu'^{\mathsf{T}}}_{ri}\, \sigma_2\, \nu'_{rj} + (i m_{ij} {\nu'^{\mathsf{T}}}_{ri}\, \sigma_2\, \nu'_{rj})^\dagger \right] . \tag{294}$$

They are Lorentz invariant because under the Lorentz transformations (229)

$$\begin{aligned} {\nu''^{\mathsf{T}}}_r \sigma_2 \nu''_r &= {\nu'^{\mathsf{T}}}_r \exp(\boldsymbol{z}^* \cdot \boldsymbol{\sigma}^{\mathsf{T}}) \sigma_2 \exp(\boldsymbol{z}^* \cdot \boldsymbol{\sigma})\, \nu'_r \\ &= {\nu'^{\mathsf{T}}}_r\, \sigma_2 \exp(-\boldsymbol{z}^* \cdot \boldsymbol{\sigma}) \exp(\boldsymbol{z}^* \cdot \boldsymbol{\sigma})\, \nu'_r = {\nu'^{\mathsf{T}}}_r\, \sigma_2\, \nu'_r . \end{aligned} \tag{295}$$

The Majorana mass terms (294) are unrelated to the scale v of the Higgs field's mean value. One can show that the complex matrix m_{ij} is symmetric. One then must combine the mass matrix in (294) with the mass matrix M_ν in (287). The resulting mass matrix will have a singular-value decomposition with six singular values that would be the masses of the "physical" neutrinos. If these six masses are equal in pairs, then the three pairs would form three Dirac neutrinos.

Whether or not there are right-handed neutrinos, we can make Majorana mass terms like ${\nu_\ell}^{\mathsf{T}} \sigma_2 \nu_\ell$, which are Lorentz invariant but not invariant under $SU_\ell(2)$ or $U_Y(1)$. We can make them gauge invariant by using a triplet $\vec{\phi} = \sigma_i \phi_i$ of Higgs fields that transforms as $\vec{\phi}' \cdot \vec{\sigma} = g(\vec{\phi} \cdot \vec{\sigma}) g^\dagger$ for $g \in SU_\ell(2)$ and that carries a value of $Y = -1$. Then if σ_2 has Lorentz indices and σ'_2 has $SU_\ell(2)$ indices, the term

$$E_\ell{}^{\mathsf{T}} \sigma_2 \sigma'_2 (\vec{\phi} \cdot \vec{\sigma}) E_\ell \tag{296}$$

is both Lorentz invariant and gauge invariant. If the potential $V(\vec{\phi})$ has minima at $\vec{\phi} \neq 0$, then this term violates lepton number and gives a Majorana mass to the neutrino. Another possibility is to say that at higher energies a theory with new fields of very high mass Λ plays a role, and that when one path-integrates over these heavy fields, one is left with an effective, nonrenormalizable term in the action

$$-\frac{(H^\dagger E_\ell)^2}{\Lambda} \tag{297}$$

which gives a Majorana mass to ν_ℓ.

Models with both right-handed and left-handed neutrinos are easier to think about, but only experiments can tell us whether right-handed neutrinos exist.

What is known experimentally is that there are at least three masses that satisfy

$$\begin{aligned} \left|\Delta m_{21}^2\right| &\equiv \left|m_2^2 - m_1^2\right| = (7.53 \pm 0.18) \times 10^{-5}\,\mathrm{eV}^2 \\ \left|\Delta m_{32}^2\right| &\equiv \left|m_3^2 - m_2^2\right| = (2.44 \pm 0.06) \times 10^{-3}\,\mathrm{eV}^2 \quad \text{normal mass hierarchy} \\ \left|\Delta m_{32}^2\right| &\equiv \left|m_3^2 - m_2^2\right| = (2.52 \pm 0.07) \times 10^{-3}\,\mathrm{eV}^2 \quad \text{inverted mass hierarchy.} \end{aligned} \tag{298}$$

If the neutrinos are Dirac particles, then they have a CKM matrix like that of the quarks with one CP-violating phase. But whereas one chooses to make the mass and flavor eigenfields the same for the up quarks u, c, t, for the leptons one makes the mass and flavor eigenfields the same for the down or charged leptons e, μ, τ. So the neutrino CKM matrix actually is $V = L_e^\dagger L_\nu$. If they are three Majorana particles, then their CKM matrix has two extra CP-violating phases α_{12} and α_{31}. A common convention for the neutrino CKM matrix is

$$V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \begin{pmatrix} \cos\theta_{13} & 0 & \sin\theta_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -\sin\theta_{13} e^{i\delta} & 0 & \cos\theta_{13} \end{pmatrix} \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\alpha_{12}/2} & 0 \\ 0 & 0 & e^{i\alpha_{31}/2} \end{pmatrix}. \tag{299}$$

This convention without the last 3×3 matrix also is used for the quark CKM matrix. The current estimates are

$$\sin^2(2\theta_{12}) = 0.846 \pm 0.021 \tag{300}$$

$$\sin^2(2\theta_{23}) = 0.999 \begin{array}{l} +0.001 \\ -0.018 \end{array} \quad \text{normal mass hierarchy} \tag{301}$$

$$\sin^2(2\theta_{23}) = 1.000 \begin{array}{l} +0.000 \\ -0.017 \end{array} \quad \text{inverted mass hierarchy} \tag{302}$$

$$\sin^2(2\theta_{13}) = 0.093 \pm 0.008. \tag{303}$$

Two of these are big angles: $2\theta_{12} \approx 2\theta_{23} = \pi/2 \pm n\pi$. In the normal hierarchy, the lightest neutrino is about 2/3 electron, 1/6 muon, and 1/6 tau; the very slightly heavier neutrino is about 1/3 electron, 1/3 muon, and 1/3 tau; and the much heavier heavier neutrino is about 1/6 electron, 5/12 muon, and 5/12 tau.

14.2 The Seesaw Mechanism

Why are the neutrino masses so light? Suppose we wish to find the eigenvalues of the real, symmetric mass matrix

$$\mathcal{M} = \begin{pmatrix} 0 & m \\ m & M \end{pmatrix} \tag{304}$$

in which m is an ordinary mass and M is a huge mass. The eigenvalues μ of this hermitian mass matrix satisfy $\det(\mathcal{M} - \mu I) = \mu(\mu - M) - m^2 = 0$ with solutions $\mu_\pm = \left(M \pm \sqrt{M^2 + 4m^2}\right)/2$. The larger mass $\mu_+ \approx M + m^2/M$ is approximately the huge mass M and the smaller mass $\mu_- \approx -m^2/M$ is tiny. The physical mass of a fermion is the absolute value of its mass parameter, here m^2/M.

The product of the two eigenvalues is the constant $\mu_+\mu_- = \det \mathcal{M} = -m^2$ so as μ_- goes down, μ_+ must go up. In 1975, Gell-Mann, Ramond, Slansky, and Jerry Stephenson invented this "**seesaw**" mechanism as an explanation of why neutrinos have such small masses, less than 1 eV$/c^2$. If $mc^2 = 10$ MeV, and $\mu_- c^2 \approx 0.01$ eV, which is a plausible light-neutrino mass, then the rest energy of the huge mass would be $Mc^2 = 10^7$ GeV. This huge mass would be one of the six neutrino masses and would point at new physics, beyond the standard model. Yet the small masses of the neutrinos may be related to the weakness of their interactions.

Before leaving the subject of fermion masses, let's look more closely at Dirac and Majorana mass terms. A Dirac field is a linear combination of two Majorana fields of the same mass

$$\psi = \frac{1}{\sqrt{2}} \begin{pmatrix} L + i\ell \\ R + ir \end{pmatrix} \tag{305}$$

in which L and ℓ are two-component left-handed spinors, and R and r are two-component right-handed spinors. The Dirac mass term

$$\begin{aligned} m\,\overline{\psi}\psi &= i\,m\,\psi^\dagger\gamma^0\psi = m\,\psi^\dagger \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \psi \\ &= m\frac{1}{2}\left(L^\dagger - i\ell^\dagger, \quad R^\dagger - ir^\dagger\right) \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \begin{pmatrix} L + i\ell \\ R + ir \end{pmatrix} \\ &= m\frac{1}{2}\left(L^\dagger - i\ell^\dagger, \quad R^\dagger - ir^\dagger\right) \begin{pmatrix} R + ir \\ L + i\ell \end{pmatrix} \\ &= m\frac{1}{2}\left[\left(L^\dagger - i\ell^\dagger\right)(R + ir) + \left(R^\dagger - ir^\dagger\right)(L + i\ell)\right] \\ &= m\,\frac{1}{2}\left(R^\dagger - ir^\dagger\right)(L + i\ell) + \text{h.c.}, \end{aligned} \tag{306}$$

in which h.c. means hermitian conjugate, gives mass m to the particle and antiparticle of the Dirac field ψ.

We may set

$$R = i\sigma_2\, L^* \iff L = -i\sigma_2\, R^* \tag{307}$$

$$r = i\sigma_2\, \ell^* \iff \ell = -i\sigma_2\, r^* \tag{308}$$

$$\tag{309}$$

or equivalently

$$\begin{pmatrix} R_1 \\ R_2 \end{pmatrix} = \begin{pmatrix} L_2^\dagger \\ -L_1^\dagger \end{pmatrix} \iff \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} = \begin{pmatrix} -R_2^\dagger \\ R_1^\dagger \end{pmatrix} \tag{310}$$

$$\begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} \ell_2^\dagger \\ -\ell_1^\dagger \end{pmatrix} \iff \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix} = \begin{pmatrix} -r_2^\dagger \\ r_1^\dagger \end{pmatrix} \tag{311}$$

$$\tag{312}$$

which are the Majorana conditions. Since $R^\dagger = -iL^\mathsf{T}\sigma_2$, we can write the Dirac mass term (306) in terms of left-handed fields as

$$\begin{aligned} m\,\overline{\psi}\psi &= \frac{1}{2}\,m\,\left(-i\,L^\mathsf{T} - \ell^\mathsf{T}\right)\sigma_2\,(L + i\ell) + \text{h.c.} && (313) \\ &= \frac{1}{2}\,m\,\left(L^\mathsf{T} - i\ell^\mathsf{T}\right)(-i\,\sigma_2)\,(L + i\ell) + \text{h.c.} && (314) \\ &= \frac{1}{2}\,m\,\left(L_1 - i\ell_1, \quad L_2 - i\ell_2\right)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} L_1 + i\ell_1 \\ L_2 + i\ell_2 \end{pmatrix} + \text{h.c.} && (315) \\ &= \frac{1}{2}\,m\,\left(L_1 - i\ell_1, \quad L_2 - i\ell_2\right)\begin{pmatrix} -L_2 - i\ell_2 \\ L_1 + i\ell_1 \end{pmatrix} + \text{h.c.} && (316) \\ &= \frac{1}{2}\,m\,(L_1 - i\ell_1)\,(-L_2 - i\ell_2) + (L_2 - i\ell_2)\,(L_1 + i\ell_1) + \text{h.c.} && (317) \end{aligned}$$

The fermion fields anticommute, so the Dirac mass term is

$$m\,\overline{\psi}\psi = \frac{1}{2}\,m\,(-2L_1L_2 - 2\ell_1\ell_2) + \text{h.c.} = -m\,(L_1L_2 + \ell_1\ell_2) + \text{h.c.}, \tag{318}$$

and it says that the fields L and ℓ have the same mass m, as they must if they are to form a Dirac field.

Since $L^\dagger = iR^\mathsf{T}\,\sigma_2$, we also can write the Dirac mass term in terms of the right-handed fields as

$$m\,\overline{\psi}\psi = \frac{1}{2}\,m\,\left(R^\mathsf{T} - ir^\mathsf{T}\right)\,i\sigma_2\,(R + ir) + \text{h.c.} \tag{319}$$

$$= \frac{1}{2}\,m\,\left(R_1 - ir_1, \quad R_2 - ir_2\right)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} R_1 + ir_1 \\ R_2 + ir_2 \end{pmatrix} + \text{h.c.} \tag{320}$$

$$= m\,(R_1R_2 + r_1r_2) + \text{h.c.} \tag{321}$$

So the fields R and r have the same mass m, as they must if they are to form a Dirac field.

The Majorana mass term for a right-handed field r of mass m evidently is

$$m\,r_1\,r_2 + \text{h.c.} \tag{322}$$

14.3 Neutrino Oscillations

The phase difference $\Delta\phi$ between two highly relativistic neutrinos of momentum p going a distance L in a time $t \approx L$ varies with their masses m_1 and m_2 as

$$\Delta\phi = t\,\Delta E = \frac{LE}{p}\Delta E = \frac{LE}{p}\left(\sqrt{p^2 + m_1^2} - \sqrt{p^2 + m_2^2}\right) \tag{323}$$

in natural units. We can approximate this phase by using the first two terms of the binomial expansion of the square roots with $y = 1$ and $x = m_i^2/p^2$

$$\Delta\phi = LE\left(\sqrt{1 + m_1^2/p^2} - \sqrt{1 + m_2^2/p^2}\right) \approx \frac{LE\Delta m^2}{p^2} \approx \frac{L\Delta m^2}{E} \tag{324}$$

or in ordinary units $\Delta\phi \approx L\Delta m^2 c^3/(\hbar E)$.

15 Grand Unification

The success of the electroweak unification of the standard model led physicists in the 1970s to propose what they called grand unification. Their goal was to unify the electroweak and the strong interactions.

Howard Georgi and Sheldon Glashow made the first attempt in 1974. They chose the group $SU(5)$ which with 24 generators is big enough to house $SU_c(3) \otimes SU_\ell(2) \otimes U_Y(1)$ and its 12 generators.

Compact internal-symmetry groups can't rotate left-handed fields into right-handed fields. So the first problem they overcame was how to combine transformations $SU_\ell(2)$ that act only on left-handed fields with ones $SU_c(3)$ that act on both left- and right-handed fields. They solved that problem by writing **all** fields as left-handed fields. Recall for instance that if u_r is a right-handed up-quark field, that is if it transforms like

$$u_r' = \exp(\vec{z} \cdot \vec{\sigma})\, u_r \tag{325}$$

then

$$u_\ell^c = \sigma_2\, u_r^* \tag{326}$$

is left-handed, that is, it transforms as

$$\begin{aligned} (u_\ell^c)' &= \sigma_2\, (u_r')^* = \sigma_2\, [\exp(\vec{z} \cdot \vec{\sigma})\, u_r]^* & (327)\\ &= \sigma_2\, \exp(\vec{z}^* \cdot \vec{\sigma}^*)\, u_r^* = \exp(-\vec{z}^* \cdot \vec{\sigma})\, \sigma_2\, u_r^* & (328)\\ &= \exp(-\vec{z}^* \cdot \vec{\sigma})\, u_\ell^c & (329) \end{aligned}$$

because

$$\sigma_2\, \vec{z}^* \cdot \vec{\sigma}^* = -\vec{z}^* \cdot \vec{\sigma}\, \sigma_2. \tag{330}$$

So they wrote all the fields as left-handed fields. They had 15 left-handed Fermi fields in each generation; at that time, only two generations were known. Since all 15 fields are left handed, we may drop the subscript ℓ. They had u_r, u_g, u_b, d_r, d_g, d_b, e, $\nu = \nu_e$, ${u_r}^c$, ${u_g}^c$, ${u_b}^c$, ${d_r}^c$, ${d_g}^c$, ${d_b}^c$, and e^c. They left out ν^c which is a right-handed neutrino and took the neutrinos to be massless. (The physics community had not yet accepted Ray Davis's late-1960s discovery of neutrino oscillations.)

Georgi and Glashow put the 15 left-handed quark and lepton fields into the 5^*

$$\mathbf{5}^* = \begin{pmatrix} {d_r}^c \\ {d_g}^c \\ {d_b}^c \\ e \\ \nu \end{pmatrix} \tag{331}$$

and a 10

$$\mathbf{10} = \begin{pmatrix} 0 & {u_b}^c & -{u_g}^c & -u_r & -d_r \\ -{u_b}^c & 0 & {u_r}^c & -u_g & -d_g \\ {u_g}^c & -{u_r}^c & 0 & -u_b & -d_b \\ u_r & u_g & u_b & 0 & -e^c \\ d_r & d_g & u_b & e^c & 0 \end{pmatrix} \tag{332}$$

and introduced 13 new gauge bosons ${Y_r}^\mu$, ${Y_g}^\mu$, ${Y_b}^\mu$; ${Y_r}^{c\mu}$, ${Y_g}^{c\mu}$, ${Y_b}^{c\mu}$; ${X_r}^\mu$, ${X_g}^\mu$, ${X_b}^\mu$; ${X_r}^{c\mu}$, ${X_g}^{c\mu}$, and ${X_b}^{c\mu}$; and A^μ.

All the generators of $SU(5)$ are traceless matrices. Thus the diagonalized charge operator Q is traceless. In the $\mathbf{5}^*$ representation, the sum of its its diagonal elements must vanish:

$$q({d_r}^c) + q({d_g}^c) + q({d_b}^c) + q(e) + q(\nu) = 0. \tag{333}$$

The neutrino is neutral, and the charges of the antidown quarks are color independent. Thus

$$3\,q(d^c) = \, -\, q(e) \tag{334}$$

or $q(d^c) = |e|/3$.

But gauge theories with quarks and antiquarks, leptons and antileptons in the same multiplet have gauge bosons that mediate changes of quark and lepton number. Putting quarks and antiquarks into the same multiplet means that nucleons are unstable. The proton is a colorless s-state of u_r, u_g, and u_b. The processes $u_g + d_b \to {Y_r}^c$ and ${Y_r}^c \to {u_r}^c + e^c$ lead to proton decay:

$$p = u_r + u_g + d_b \to u_r + {Y_r}^c \tag{335}$$

$$u_r + {Y_r}^c \to u_r + {u_r}^c + e^c \tag{336}$$

$$u_r + {u_r}^c + e^c \to \pi^0 + e^+. \tag{337}$$

Other processes lead to other modes of proton decay and to the decay of neutrons in otherwise stable nuclei. The lifetime τ_p of the proton is proportional to the fourth power of the mass of the Y

$$\tau_p \propto \frac{M_Y^4}{\alpha^2 m_p^5} \tag{338}$$

in which α is the fine structure constant of $SU(5)$. The lower bound on the lifetime of the proton due to this decay mode is 8.3×10^{33} years. Putting charge-conjugated right-handed fields into the same multiplet as left-handed fields changes the focus of physics from accessible energies to the GUT scale or $M_Y > 10^{16}$ GeV. This seems premature.

Harald Fritzsch, Peter Minkowski, Howard Georgi, and Edward Witten put the left-handed fields of a single family into a 16-dimensional multiplet of $SO(10)$, which to save paper I represent as a row vector

$$\mathbf{16}^{\mathsf{T}} = \begin{pmatrix} \nu^c & e^c & u_r & d_r & u_g & d_g & u_b & d_b & {u_r}^c & {u_g}^c & {u_b}^c & {d_r}^c & {d_g}^c & {d_b}^c & \nu & e \end{pmatrix} \tag{339}$$

This theory of grand unification is more symmetrical and has room for a right-handed neutrino, which appears as ν^c. This theory also produces proton decay unless the masses of the gauge bosons exceed about 10^{16} GeV. Unfortunately, $SO(10)$ does not seem to explain the charges as directly as $SU(5)$ because the sum of the charges of the particles of the **16** vanishes no matter what they are as long as $q + q^c = 0$. This may be why Georgi and Glashow opted for $SU(5)$, which Georgi discovered only a few hours after figuring out $SO(10)$. But if one fits the 16 particles of the **16** into $SU(5)$ multiplets, then one recovers the $SU(5)$ version of charge quantization.

The gauge group of the standard model is $SU_c(3) \otimes SU_\ell(2) \otimes U_Y(1)$ with three coupling constants g_s, g, and g' which have nothing to do with each other. Grand unification puts these three groups into a simple group with a single coupling constant and traceless generators T^a that are related to one another by the structure constants f_{abc}

$$[T^a, T^b] = i f_{abc} T^c \tag{340}$$

which are real and totally antisymmteric, and the same for every representation whether reducible or irreducible. A **simple** group G is one that has no nontrivial invariant subgroup S; that is, if

$$g^{-1} s g = s' \in S \text{ for all } s \in S \text{ and all } g \in G \tag{341}$$

then either $S = G$ or S consists of the identity element of G. The group of the standard model $SU_c(3) \otimes SU_\ell(2) \otimes U_Y(1)$ is not simple (or semi-simple). Its structure constants don't relate the $SU_c(3)$ generators to the $SU_\ell(2)$ generators or to the $U_Y(1)$ generator.

The generators of any representation whether reducible or irreducible of a group may be taken to be orthogonal with a normalization N_D that depends upon the representation D

$$\mathrm{Tr}\, T^a T^b = N_D\, \delta_{ab}. \tag{342}$$