Tutorial 4: The BEH mechanism
With answers

Dr. M Flowerdew

December 2, 2015

In the last tutorial, we finished by noting that gauge symmetries require all gauge bosons to be massless, and also that fermion masses are forbidden in the Standard Model. Observationally, we know that the W and Z bosons, as well as the fermions, *do* have mass, which should apparently be forbidden by the $\mathrm{SU}(2)_{\mathrm{L}}$ gauge symmetry of the weak nuclear force. We will now explore how this issue is resolved within the Standard Model.

1 The Higgs field

The solution, proposed by Robert Brout, François Englert, Peter Higgs and others, is not to abandon the symmetry principles behind the description of forces, but rather to change the vacuum state itself. In everything considered so far, it has been implicitly assumed that in the vacuum all expectation values $\langle\psi\rangle = \langle 0|\psi|0\rangle$ are zero. This arises naturally if the potential for a particle has a minumum at zero, as in Figure 1 (left).[1] Adding more energy to the field[2] increases the maximum possible value for $|\psi|$.

The Higgs field ϕ, on the other hand, has a quartic potential, illustrated in Figure 1 (right). At very high energies (the upper dotted line), this is difficult to distinguish from the quadratic case, but at low energies (lower dotted line), it is clear that the "bump" at $\phi \sim 0$ will affect the ground state significantly. In this simple one-dimensional example, there will be two degenerate vacuum states, each centered on one of the two minima illustrated, with the same average magnitude $\langle|\phi|\rangle = v$.

The physical Higgs field is a complex $\mathrm{SU}(2)_{\mathrm{L}}$ doublet, with four real components. Despite this, the vacuum expectation value (vev, parameterised by a constant v), and real fluctuations around it, $H(x)$, can be written with full generality as

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

[1]Recall from the first tutorial that a (quadratic) mass term in the Lagrangian density can be regarded as a kind of potential energy.

[2]This means adding quanta to the particular mode described by $\psi(E, \boldsymbol{p})$.

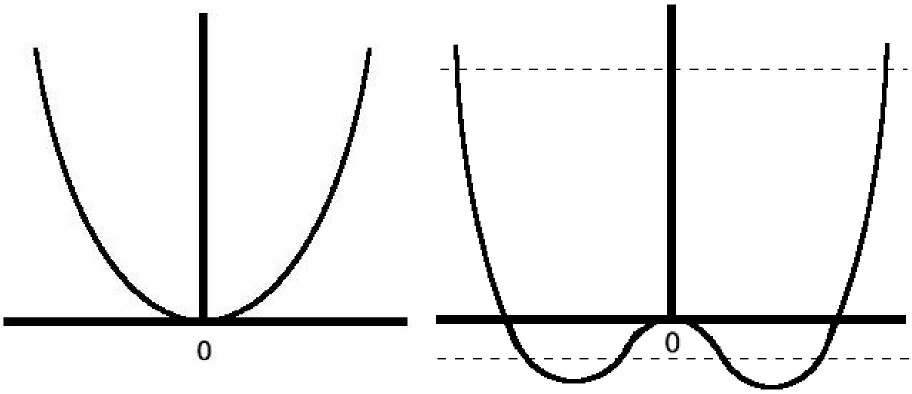

Figure 1: Left: A quadratic potential function for a massive particle. Right: The Higgs field potential, with quadratic and quartic terms. In both cases, the x-axis corresponds loosely to the field magnitude.

where $v > 0$ is a real constant. This amounts to a specific choice of (spacetime-dependent) $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_Y$ gauge. After making this choice, the underlying symmetry is no longer apparent; it has been *broken*.

Exercise 1: Using results from tutorial 2, generic $\mathrm{SU}(2)_\mathrm{L}$ and $\mathrm{U}(1)_Y$ gauge transformations may be written as

$$G_{\mathrm{SU}(2)_\mathrm{L}} = \begin{pmatrix} \omega & \zeta \\ -\zeta^* & \omega^* \end{pmatrix} \text{ and } G_{\mathrm{U}(1)_Y} = \epsilon \tag{2}$$
$$\text{with } |\omega|^2 + |\zeta|^2 = 1 \text{ and } |\epsilon|^2 = 1.$$

How many *real* free parameters does each transformation have?

Use these results to compute the full $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_Y$ transformation (simply the product of the individual transformations) applied to an arbitrary Higgs field:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} v_a \\ v_b \end{pmatrix}.$$

Find the specific rotation required to express the Higgs field in the form of Equation (1). How many free real parameters remain?

Answer: The $\mathrm{SU}(2)_\mathrm{L}$ transformation has two complex parameters (i.e. four real numbers) and one constraint, so three real free parameters in total. The $\mathrm{U}(1)_Y$ transformation has one real free parameter (one complex number and one constraint).

The combined $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_Y$ transformation (which we will write as G) can be written as the product of the individual transformations:

$$\begin{aligned} G &= G_{\mathrm{U(1)}_Y} G_{\mathrm{SU(2)_L}} \\ &= \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix} \begin{pmatrix} \omega & \zeta \\ -\zeta^* & \omega^* \end{pmatrix} \\ &= \begin{pmatrix} \epsilon\omega & \epsilon\zeta \\ -\epsilon\zeta^* & \epsilon\omega^* \end{pmatrix}. \end{aligned}$$

Applying this to the generic Higgs field given in the question, we obtain the transformed field

$$\begin{aligned} \phi' = G\phi &= \frac{1}{\sqrt{2}} \begin{pmatrix} \epsilon\omega & \epsilon\zeta \\ -\epsilon\zeta^* & \epsilon\omega^* \end{pmatrix} \begin{pmatrix} v_a \\ v_b \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} \epsilon\omega v_a + \epsilon\zeta v_b \\ -\epsilon\zeta^* v_a + \epsilon\omega^* v_b \end{pmatrix}. \end{aligned}$$

We wish this to be equal to the expression in Equation (1) (with $H(x) = 0$), which is achieved with the following constraints:

$$\omega = \frac{v_b}{v}\epsilon; \qquad\qquad \zeta = -\frac{v_a}{v}\epsilon,$$

where $v^2 = v_a^2 + v_b^2$. With these two constraints, only one free parameter remains. We could, for example, consider the phase of ϵ to represent this free parameter, and use the above equation to derive the necessary values of ω and ζ to fix the gauge.

In this new vacuum, previously massless particles now appear to have mass. It is easiest to see how this might work for fermions. Consider the Lagrangian density interaction term $y\overline{\psi_{\mathrm{L}}}\psi_{\mathrm{R}}\phi$, where y is an (at this point) arbitrary coupling constant between ψ and ϕ. This, unlike the fermion mass term $m\bar{\psi}\psi$, is a gauge-invariant scalar quantity, and is thus allowed in the Lagrangian density. With the specific choice of Equation (1), and supposing for the moment that we are concerned with electron-like fields $\psi_{\mathrm{L}} = \begin{pmatrix} \nu \\ e_{\mathrm{L}} \end{pmatrix}$ and $\psi_{\mathrm{R}} = e_{\mathrm{R}}$, we have

$$\begin{aligned} y\overline{\psi_{\mathrm{L}}}\psi_{\mathrm{R}}\phi &= \frac{y}{\sqrt{2}} \begin{pmatrix} \overline{\nu} & \overline{e_{\mathrm{L}}} \end{pmatrix} e_{\mathrm{R}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \\ &= \frac{y}{\sqrt{2}} v \overline{e_{\mathrm{L}}} e_{\mathrm{R}} + \frac{y}{\sqrt{2}} H(x) \overline{e_{\mathrm{L}}} e_{\mathrm{R}}. \end{aligned} \tag{3}$$

Now, the first term has the quadratic structure of a mass term, while the second term describes an interaction between the electron field and excitations of the Higgs field $H(x)$. Adding on the Hermitian conjugate $y^*\phi^\dagger\overline{\psi_{\mathrm{R}}}\psi_{\mathrm{L}}$ gives a term proportional to $\overline{e_{\mathrm{R}}}e_{\mathrm{L}}$. The two together therefore give the appearance that the fermion field $e = e_{\mathrm{L}} + e_{\mathrm{R}}$ has mass.

We will return to the question of gauge boson mass after a short diversion.

2 The Standard Model Lagrangian density

We are now ready to write down the complete Lagrangian density for the Standard Model. The gauge symmetry groups are:

- A $\mathrm{U}(1)_Y$ symmetry, acting on hypercharge, with an associated vector field B^μ.
- An $\mathrm{SU}(2)_\mathrm{L}$ symmetry acting on left-handed fermions and the Higgs field, with three vector fields $\boldsymbol{W}^\mu$.
- An $\mathrm{SU}(3)_C$ symmetry acting on colour, with eight vector field $\boldsymbol{G}^\mu$.

The matter fields are the following (for one generation - the other two generations are identically structured):

- A left-handed lepton $\mathrm{SU}(2)_\mathrm{L}$ doublet, $\ell_\mathrm{L} = \begin{pmatrix} \nu_\mathrm{L} \\ e_\mathrm{L} \end{pmatrix}$.
- A right-handed electron $\mathrm{SU}(2)_\mathrm{L}$ singlet, e_R.[3]
- A left-handed quark $\mathrm{SU}(2)_\mathrm{L}$ doublet, $q_\mathrm{L} = \begin{pmatrix} u_\mathrm{L} \\ d_\mathrm{L} \end{pmatrix}$.
- Two right-handed quark $\mathrm{SU}(2)_\mathrm{L}$ singlets, u_R and d_R.

Finally, there is the Higgs field ϕ, a complex scalar $\mathrm{SU}(2)_\mathrm{L}$ doublet. The hypercharges and gauge group representations of all these fields are summarised in Table 1.

[3]Theoretically, we could hypothesise a right-handed neutrino field ν_R and give neutrinos mass. However, there are other possbile mechanisms for this and there is as yet no solid evidence for ν_R fields specifically, so this is not considered to be part of the Standard Model. We will return to the topic of neutrino masses later in the course.

Particle	Y_f	$\mathrm{SU}(2)_\mathrm{L}$plet	$\mathrm{SU}(3)_C$-plet
ℓ_L	-1	$\mathbf{2}$	$\mathbf{1}$
e_R	-2	$\mathbf{1}$	$\mathbf{1}$
q_L	$\frac{1}{3}$	$\mathbf{2}$	$\mathbf{3}$
u_R	$\frac{4}{3}$	$\mathbf{1}$	$\mathbf{3}$
d_R	$-\frac{2}{3}$	$\mathbf{1}$	$\mathbf{3}$
ϕ	1	$\mathbf{2}$	$\mathbf{1}$
B	0	$\mathbf{1}$	$\mathbf{1}$
$\boldsymbol{W}$	0	$\mathbf{3}$	$\mathbf{1}$
$\boldsymbol{G}$	0	$\mathbf{1}$	$\mathbf{8}$

Table 1: Hypercharge values and gauge representations for the Standard Model fields.

Using these fields, the SM Lagrangian density *for one fermion generation* may be written as follows:

$$\mathcal{L} = \sum_f i\overline{\psi_f}\gamma^\mu \mathcal{D}^f_\mu \psi_f \tag{4a}$$

$$- \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}\boldsymbol{W}_{\mu\nu}\cdot\boldsymbol{W}^{\mu\nu} - \frac{1}{4}\boldsymbol{G}_{\mu\nu}\cdot\boldsymbol{G}^{\mu\nu} \tag{4b}$$

$$+ (\mathcal{D}^\phi_\mu\phi)^\dagger \mathcal{D}^{\phi\mu}\phi - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \tag{4c}$$

$$- \left[\overline{\ell_\mathrm{L}}\phi y_e e_\mathrm{R} + \text{h.c.}\right. \tag{4d}$$

$$+ \overline{q_\mathrm{L}}\phi y_d d_\mathrm{R} + \text{h.c.} \tag{4e}$$

$$\left. + \begin{pmatrix} -\overline{d_\mathrm{L}} & \overline{u_\mathrm{L}} \end{pmatrix} \phi^* y_u u_\mathrm{R} + \text{h.c.}\right] \tag{4f}$$

$$+ \theta\frac{\alpha_s}{8\pi}\epsilon^{\mu\nu\rho\sigma}\boldsymbol{G}_{\mu\nu}\cdot\boldsymbol{G}_{\rho\sigma}\,. \tag{4g}$$

Let us examine these lines one by one.

The first line, (4a), contains kinetic and gauge interaction terms for all of the fermion fields, indicated by the sum over f. The covariant derivative is now also labeled by f, as its form changes depending on the properties of the associated fermions. The most complicated covariant derivative is the one for the left-handed quarks, which has the following form:

$$\begin{aligned} \mathcal{D}^{q_\mathrm{L}}_\mu &= \partial_\mu + i\frac{g'}{2}Y_{q_\mathrm{L}}B_\mu + ig\boldsymbol{T}\cdot\boldsymbol{W}_\mu + ig_s\boldsymbol{T_s}\cdot\boldsymbol{G}_\mu \\ &= \partial_\mu + i\frac{g'}{2}Y_{q_\mathrm{L}}B_\mu + i\frac{g}{2}\boldsymbol{\tau}\cdot\boldsymbol{W}_\mu + i\frac{g_s}{2}\boldsymbol{\lambda}\cdot\boldsymbol{G}_\mu. \end{aligned} \tag{5}$$

The first line uses the notation from the previous tutorials, and is very convenient for whole matrix operations due to the consistent normalisation of $\mathrm{Tr}(T_a^2) = \frac{1}{2}$. In contrast, the $\boldsymbol{\tau}$ and $\boldsymbol{\lambda}$ matrices tend to be more convenient in calculations that involve the individual matrix elements. The factor of $\frac{1}{2}$ in the U(1) hypercharge term is purely conventional, but related to the above usage.

In (4b), we have the kinetic terms for the gauge bosons. The field tensors are related to the commutator of the relevant covariant derivative, as shown in the last tutorial. For example, $B_{\mu\nu} = -\frac{i}{g'}[\mathcal{D}_\mu, \mathcal{D}_\nu] = \partial_\mu B_\nu - \partial_\nu B_\mu$.

Lines (4c) to (4f) contain terms involving the Higgs field ϕ. The potential terms on line (4c) allow its vev to be non-zero. The remaining lines allow the fermion fields to acquire mass once this is done, parameterised by the Yukawa coupling constants y_e, y_u and y_d. For this single-generation example, the Yukawa constants are simply numbers, given by $y_f = m_f\frac{\sqrt{2}}{v}$. The mass terms are not themselves Hermitian, and so their Hermitian conjugates ("h.c.") must also be added to the Lagrangian density. For example, the "h.c." term in (4d), giving mass to the electron field, is $-\overline{e_\mathrm{R}} y_e^* \phi^\dagger \ell_\mathrm{L}$.

Exercise 2: Why are the left-handed quark fields reversed in Equation (4f)? *Hint:* What terms would result if they were replaced with $\overline{q_\mathrm{L}}$? Why are these terms forbidden?

Answer: Let's start by expanding the terms in Equation (4f) using the Higgs field given in Equation (1) with $H(x) = 0$:

$$-\begin{pmatrix}-\overline{d_\mathrm{L}} & \overline{u_\mathrm{L}}\end{pmatrix}\begin{pmatrix}0\\ v\end{pmatrix}\frac{y_u}{\sqrt{2}}u_\mathrm{R} - \overline{u_\mathrm{R}}\frac{y_u^*}{\sqrt{2}}\begin{pmatrix}0 & v\end{pmatrix}\begin{pmatrix}-d_\mathrm{L}\\ u_\mathrm{L}\end{pmatrix}$$
$$= -\frac{1}{\sqrt{2}}\left(y_u v\overline{u_\mathrm{L}}u_\mathrm{R} + y_u^* v\overline{u_\mathrm{R}}u_\mathrm{L}\right).$$

The terms proportional to v have the correct form for an up-quark mass term.

Now, let us consider the alternative suggested in the question:

$$-\overline{q_\mathrm{L}}\phi^* y_u u_\mathrm{R} - \overline{u_\mathrm{R}} y_u^* \phi q_\mathrm{L} = -\begin{pmatrix}\overline{u_\mathrm{L}} & \overline{d_\mathrm{L}}\end{pmatrix}\begin{pmatrix}0\\ v\end{pmatrix}\frac{y_u}{\sqrt{2}}u_\mathrm{R} - \overline{u_\mathrm{R}}\frac{y_u^*}{\sqrt{2}}\begin{pmatrix}0 & v\end{pmatrix}\begin{pmatrix}u_\mathrm{L}\\ d_\mathrm{L}\end{pmatrix}$$
$$= -\frac{1}{\sqrt{2}}\left(y_u v\overline{d_\mathrm{L}}u_\mathrm{R} + y_u^* v\overline{u_\mathrm{R}}d_\mathrm{L}\right).$$

These terms mix up- and down-type quarks, violating electric charge (amongst other things).

Finally, we have the term in (4g). This is the so-called "QCD θ" term, and is in principle allowed by all of the symmetries of the SM. However, the presence of the fully antisymmetric tensor $\epsilon^{\mu\nu\rho\sigma}$ means that this term violates CP, an effect which is not observed in strong interactions. Experimental evidence suggests that the parameter θ is less than about 10^{-9}, although there is no clear reason as to why it should be so small.

3 Electroweak symmetry breaking and the boson sector

3.1 The Higgs boson

In the stable vacuum at low energy, the Higgs field has a non-zero vacuum expectation value. In actual fact, an infinite number of equivalent vacua exist, which in turn has important consequences for the B^μ and $\boldsymbol{W}^\mu$ fields.

The potential, taken from Equation (4c), may be rewritten (up to a constant term that we may neglect) in terms of two new real parameters v and m_h:

$$\begin{aligned} V(\phi) &= \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \\ &= \frac{m_h^2}{2v^2}\left(\phi^\dagger\phi - \frac{v^2}{2}\right)^2 + \text{const.} \end{aligned} \tag{6}$$

In the second form, there is a ring of minima (as long as v is real), all satisfying $\phi^\dagger\phi = \frac{v^2}{2}$, or $|\phi| = v/\sqrt{2}$. The values of μ and λ in terms of m_h and v are easily found

$$\begin{aligned} \lambda &= \frac{m_h^2}{2v^2}, \\ \mu^2 &= \frac{m_h^2}{2v^2}\cdot 2\cdot\left(-\frac{v^2}{2}\right) = -\frac{1}{2}m_h^2. \end{aligned} \tag{7}$$

Now we shall consider the effect of a small perturbation from some point with $|\phi| = v/\sqrt{2}$. For the moment, we shall neglect the fact that ϕ is an $\mathrm{SU(2)_L}$ doublet, and just assume it is a complex field. Without loss of generality within this assumption, we can consider starting from the point $\phi = v/\sqrt{2}$. We may write the perturbation in terms of two real parameters, χ and ψ:

$$\phi = \frac{1}{\sqrt{2}}\left(v + \chi(x) + i\psi(x)\right). \tag{8}$$

Now we place this expression into the Higgs kinetic and self-interaction part of the Standard Model Lagrangian density, for now neglecting gauge interactions, so that $\mathcal{D}_\mu^\phi = \partial_\mu$:

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} &= (\partial_\mu\phi)^\dagger\partial^\mu\phi - V(\phi) \\ &= \frac{1}{2}\partial_\mu(\chi - i\psi)\partial^\mu(\chi + i\psi) - \frac{m_h^2}{2v^2}\cdot\frac{1}{4}\left\{(v+\chi-i\psi)(v+\chi+i\psi) - v^2\right\}^2 \\ &= \frac{1}{2}\partial_\mu\chi\partial^\mu\chi + \frac{1}{2}\partial_\mu\psi\partial^\mu\psi - \frac{m_h^2}{8v^2}\left\{2v\chi + \chi^2 + \psi^2\right\}^2 \\ &= \frac{1}{2}\partial_\mu\chi\partial^\mu\chi + \frac{1}{2}\partial_\mu\psi\partial^\mu\psi - \frac{m_h^2}{8v^2}\left[4v^2\chi^2 + 4v\chi(\chi^2+\psi^2) + (\chi^2+\psi^2)^2\right] \\ &= \frac{1}{2}\partial_\mu\chi\partial^\mu\chi - \frac{1}{2}m_h^2\chi^2 + \frac{1}{2}\partial_\mu\psi\partial^\mu\psi - \mathcal{O}(\chi^3, \chi\psi^2). \end{aligned} \tag{9}$$

Here, we assume that χ and ψ are small, so that cubic and higher terms can be neglected. Recalling tutorial 1, we recognise from the form of Equation (9) that the field χ has a mass m_h, while ψ is massless. In the Standard Model, χ represents an observable Higgs boson.

What about ψ? This is called a *Goldstone boson*, and its appearance is a general feature of spontaneous symmetry breaking. It is, however, not directly observable, thanks to gauge symmetry. In this simple example, the ψ part of the transformation Equation (8) can be removed by a local infinitesimal U(1) rotation $G = e^{-i\psi(x)/\sqrt{2}}$ (see also Exercise 1).

3.2 Electroweak boson mass

Now we consider the gauge boson fields after $\mathrm{U}(1)_Y \times \mathrm{SU}(2)_\mathrm{L}$ symmetry breaking occurs. All of the essential features are in the covariant derivative $\mathcal{D}^\phi_\mu$. Written explicitly as a 2×2 matrix in $\mathrm{SU}(2)_\mathrm{L}$ space, we find that

$$\begin{aligned} \mathcal{D}^\phi_\mu &= \partial_\mu + i\frac{g'}{2} Y_\phi B_\mu + i\frac{g}{2}\boldsymbol{\tau}\cdot\boldsymbol{W}_\mu \\ &= \begin{pmatrix} \partial_\mu + i\frac{g'}{2}B_\mu + i\frac{g}{2}W^0_\mu & i\frac{g}{2}\left(W^1_\mu - iW^2_\mu\right) \\ i\frac{g}{2}\left(W^1_\mu + iW^2_\mu\right) & \partial_\mu + i\frac{g'}{2}B_\mu - i\frac{g}{2}W^0_\mu \end{pmatrix}, \end{aligned} \tag{10}$$

recalling that $Y_\phi = 1$.

We begin by defining the charged W boson fields, which are the non-zero elements of the $\mathrm{SU}(2)_\mathrm{L}$ raising and lowering operators:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp iW^2_\mu\right). \tag{11}$$

We then have B_μ and W^0_μ remaining along the leading diagonal of $\mathcal{D}^\phi_\mu$, which describe neutral-current electroweak interactions. These fields have the same quantum numbers and may mix. With a choice of $\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix}$, only the combination in the lower right corner of the matrix in Equation (10) acquires mass. With appropriate normalisation, we define this field to be Z_μ

$$Z_\mu = \frac{gW^0_\mu - g'B_\mu}{\sqrt{g^2 + g'^2}} = \cos\theta_\mathrm{W} W^0_\mu - \sin\theta_\mathrm{W} B_\mu, \tag{12}$$

where $\tan\theta_\mathrm{W} = g'/g$ defines the Weinberg angle θ_W. The orthogonal field is denoted A_μ:

$$A_\mu = \frac{g'W^0_\mu + gB_\mu}{\sqrt{g^2 + g'^2}} = \sin\theta_\mathrm{W} W^0_\mu + \cos\theta_\mathrm{W} B_\mu. \tag{13}$$

Exercise 3: Substitute the above definitions for $W_\mu^\pm$, Z_μ and A_μ into Equation (10). Use this to compute $\mathcal{D}_\mu^\phi$, and use the coefficient of A_μ to argue that the electric charge can be expressed as

$$e = g' \cos\theta_{\mathrm{W}} = g \sin\theta_{\mathrm{W}} \tag{14}$$

Answer: The substitutions for the off-diagonal elements of Equation (10) are straightforward, however the diagonal elements require more attention. We already know that the lower-right element contains only the field Z_μ:

$$\frac{g'}{2} B_\mu - \frac{g}{2} W_\mu^0 = -\frac{\sqrt{g^2 + g'^2}}{2} Z_\mu = -\frac{g}{2\cos\theta_{\mathrm{W}}} Z_\mu,$$

Where we have used the fact that $g/\sqrt{g^2+g'^2} = \cos\theta_{\mathrm{W}}$, derived from $\tan\theta_{\mathrm{W}} = g'/g$. The top-left element can be found most easily by inverting Equations (12) and (13):

$$\begin{aligned} B_\mu &= \cos\theta_{\mathrm{W}} A_\mu - \sin\theta_{\mathrm{W}} Z_\mu, \\ W_\mu^0 &= \sin\theta_{\mathrm{W}} A_\mu + \cos\theta_{\mathrm{W}} Z_\mu. \end{aligned}$$

Thus, the relevant combination of fields becomes

$$\begin{aligned} \frac{g'}{2} B_\mu + \frac{g}{2} W_\mu^0 &= \frac{g'}{2}\cos\theta_{\mathrm{W}} A_\mu + \frac{g}{2}\sin\theta_{\mathrm{W}} A_\mu - \frac{g'}{2}\sin\theta_{\mathrm{W}} Z_\mu + \frac{g}{2}\cos\theta_{\mathrm{W}} Z_\mu \\ &= \frac{g'}{2}\left(\cos\theta_{\mathrm{W}} + \frac{\sin\theta_{\mathrm{W}}}{\tan\theta_{\mathrm{W}}}\right) A_\mu + \frac{g}{2}\left(-\sin\theta_{\mathrm{W}}\tan\theta_{\mathrm{W}} + \cos\theta_{\mathrm{W}}\right) Z_\mu \\ &= g'\cos\theta_{\mathrm{W}} A_\mu + \frac{g}{2\cos\theta_{\mathrm{W}}}\left(\cos^2\theta_{\mathrm{W}} - \sin^2\theta_{\mathrm{W}}\right) Z_\mu \\ &= g'\cos\theta_{\mathrm{W}} A_\mu + \frac{g\cos 2\theta_{\mathrm{W}}}{2\cos\theta_{\mathrm{W}}} Z_\mu. \end{aligned}$$

Substituting these into Equation (10), we obtain the following:

$$\mathcal{D}_\mu^\phi = \begin{pmatrix} \partial_\mu + ig'\cos\theta_{\mathrm{W}} A_\mu + i\frac{g\cos 2\theta_{\mathrm{W}}}{2\cos\theta_{\mathrm{W}}} Z_\mu & i\frac{g}{\sqrt{2}} W_\mu^+ \\ i\frac{g}{\sqrt{2}} W_\mu^- & \partial_\mu - i\frac{g}{2\cos\theta_{\mathrm{W}}} Z_\mu \end{pmatrix}.$$

The imaginary part of the coefficient of A_μ in the top-left element of $\mathcal{D}_\mu^\phi$ is the electric charge of the upper element of the Higgs doublet (recall, for example, Equation (8) of the last tutorial). We already know that, by definition, the two elements of an $\mathrm{SU}(2)_{\mathrm{L}}$ doublet differ in their electrical charge by one unit. As the Higgs vev must be electrically neutral, this means that the upper element of the Higgs doublet has charge +1, and therefore $g'\cos\theta_{\mathrm{W}}$ is the unit of electric charge.

Exercise 4: Use the result of the previous exercise to compute $(\mathcal{D}^\phi_\mu\phi)^\dagger\mathcal{D}^{\phi\mu}\phi$ in the vacuum, i.e. with $H(x)=0$. Use this to show that the photon is massless and that

$$m_W = \frac{gv}{2}, \qquad\qquad m_Z = \frac{m_W}{\cos\theta_\mathrm{W}}. \tag{15}$$

Hint: Recall the Lagrangian density for the Proca equation from the first tutorial. What is $W_\mu^{+\dagger}$ equal to?

Answer: First, we compute $\mathcal{D}^\phi_\mu\phi$ in the vacuum

$$\begin{aligned}\mathcal{D}^\phi_\mu\phi &= \begin{pmatrix}\partial_\mu + ig'\cos\theta_\mathrm{W}A_\mu + i\frac{g\cos 2\theta_\mathrm{W}}{2\cos\theta_\mathrm{W}}Z_\mu & i\frac{g}{\sqrt{2}}W_\mu^+ \\ i\frac{g}{\sqrt{2}}W_\mu^- & \partial_\mu - i\frac{g}{2\cos\theta_\mathrm{W}}Z_\mu\end{pmatrix}\begin{pmatrix}0\\ \frac{v}{\sqrt{2}}\end{pmatrix} \\ &= \begin{pmatrix} i\frac{gv}{2}W_\mu^+ \\ -\frac{igv}{2\sqrt{2}\cos\theta_\mathrm{W}}Z_\mu\end{pmatrix}.\end{aligned}$$

To compute $(\mathcal{D}^\phi_\mu\phi)^\dagger\mathcal{D}^{\phi\mu}\phi$ we note that $Z_\mu^\dagger = Z_\mu$ (as for the photon), and that $W_\mu^{+\dagger} = W_\mu^-$. Therefore:

$$\begin{aligned}(\mathcal{D}^\phi_\mu\phi)^\dagger\mathcal{D}^{\phi\mu}\phi &= \begin{pmatrix}-i\frac{gv}{2}W_\mu^- & \frac{igv}{2\sqrt{2}\cos\theta_\mathrm{W}}Z_\mu\end{pmatrix}\begin{pmatrix} i\frac{gv}{2}W^{+\mu} \\ -\frac{igv}{2\sqrt{2}\cos\theta_\mathrm{W}}Z^\mu\end{pmatrix} \\ &= \left(\frac{gv}{2}\right)^2 W_\mu^- W^{+\mu} + \frac{1}{2}\left(\frac{gv}{2\cos\theta_\mathrm{W}}\right)^2 Z_\mu Z^\mu.\end{aligned}$$

These are mass terms for the W and Z bosons. In the first tutorial, the mass term for a vector field A^μ was written as $\frac{1}{2}m^2A^\mu A_\mu$, This applies to real fields, and so we can immediately read off the Z boson mass as $gv/(2\cos\theta_\mathrm{W})$. The W boson is electrically charged, i.e. it is a complex field, so just as for a scalar field the corresponding mass term should look like $m^2W^{+\mu}W_\mu^-$.[a] Thus, the W boson mass is $gv/2 = m_Z\cos\theta_\mathrm{W}$.

[a] Note that the charge ordering is irrelevant, as $W^{+\mu}W_\mu^- = \frac{1}{2}(W^{1\mu}W_\mu^1 + W^{2\mu}W_\mu^2) = W^{-\mu}W_\mu^+$.

Exercise 5: (Optional) Extend the previous exercise to the case of $H(x)\neq 0$. What is the coupling strength between the Higgs boson and a) two W bosons, b) two Z bosons, c) two photons?

Answer: The easiest way to answer this question is to see that we can simply use the last equation, replacing v^2 by $(v+H)^2$ throughout.

With this, we find:

$$
\begin{aligned}
(\mathcal{D}^{\phi}_{\mu}\phi)^{\dagger}\mathcal{D}^{\phi\mu}\phi &= \left(\frac{g}{2}\right)^2 (v+H)^2 W^-_{\mu} W^{+\mu} + \frac{1}{2}\left(\frac{g}{2\cos\theta_{\mathrm{W}}}\right)^2 (v+H)^2 Z_{\mu} Z^{\mu} \\
&= \left(\frac{gv}{2}\right)^2 W^-_{\mu} W^{+\mu} + \frac{1}{2}\left(\frac{gv}{2\cos\theta_{\mathrm{W}}}\right)^2 Z_{\mu} Z^{\mu} \\
&\quad + \frac{g^2 v}{2} H W^-_{\mu} W^{+\mu} + \left(\frac{g}{2\cos\theta_{\mathrm{W}}}\right)^2 v H Z_{\mu} Z^{\mu} \\
&\quad + \left(\frac{g}{2}\right)^2 H^2 W^-_{\mu} W^{+\mu} + \frac{1}{2}\left(\frac{g}{2\cos\theta_{\mathrm{W}}}\right)^2 H^2 Z_{\mu} Z^{\mu}.
\end{aligned}
$$

The first line has the gauge boson mass terms as before. The second line contains Yukawa (i.e. involving three bosons) couplings between the Higgs boson H and the gauge bosons, while the last line contains four-boson interactions. It is the Yukawa couplings that concern us here.

a), b) The H–W^+–W^- and H–Z–Z couplings are

$$
\begin{aligned}
g_{HWW} &= \frac{g^2 v}{2} = \frac{2m_W^2}{v}, \\
\text{and } g_{HZZ} &= \left(\frac{g}{2\cos\theta_{\mathrm{W}}}\right)^2 v = \frac{m_Z^2}{v}.
\end{aligned}
$$

As with the fermions (Equations (4c) to (4f)), the couplings of the gauge bosons to the Higgs boson are directly related to their masses.

c) There is no term proportional to $HA^{\mu}A_{\mu}$, and therefore at tree-level the coupling between the Higgs boson and the photon is zero. There are, however, important loop diagrams that produce an effective H–γ–γ vertex, which we will discuss later in the course.

3.3 Coupling of the Z and photon to fermions

Let's take another look at the diagonal terms of the covariant derivative in Equation (10). Generalising to other $\mathrm{SU}(2)_{\mathrm{L}}$ doublets, we can without loss of generality write the interaction part as[4]

$$
i\frac{g'}{2} Y_f B_{\mu} + igI_3 W^0_{\mu}, \tag{16}
$$

where I_3 is the third component of weak isospin for the field f (i.e. $\pm\frac{1}{2}$ for the two members of a weak doublet, or zero for a right-handed singlet).

Exercise 6: Express Equation (16) in terms of A_{μ} and Z_{μ}. Show that

[4] The quarks have strong interaction terms, but these are unaffected by what follows.

the coefficient of A_μ is equal to ieQ_f, where

$$Q_f = \frac{Y_f}{2} + I_3. \tag{17}$$

The values of Y_f in Table 1 are assigned based on this relationship. Show in turn that the coefficient of Z_μ is proportional to the "charge"

$$c_f = I_3 - Q_f \sin^2\theta_\mathrm{W}. \tag{18}$$

(The factor $g/\cos\theta_\mathrm{W}$ is conventionally absorbed into the Z boson coupling constant).

Answer: Inverting Equations (12) and (13), we can express Equation (16) in terms of the physical boson fields:

$$\begin{aligned}
i\frac{g'}{2}Y_f B_\mu + igI_3 W^0_\mu &= i\begin{pmatrix}\frac{g'}{2}Y_f & gI_3\end{pmatrix}\begin{pmatrix}B_\mu \\ W^0_\mu\end{pmatrix} \\
&= i\begin{pmatrix}\frac{g'}{2}Y_f & gI_3\end{pmatrix}\begin{pmatrix}\cos\theta_\mathrm{W} & -\sin\theta_\mathrm{W} \\ \sin\theta_\mathrm{W} & \cos\theta_\mathrm{W}\end{pmatrix}\begin{pmatrix}A_\mu \\ Z_\mu\end{pmatrix} \\
&= i\begin{pmatrix}\frac{g'}{2}Y_f\cos\theta_\mathrm{W} + gI_3\sin\theta_\mathrm{W} & gI_3\cos\theta_\mathrm{W} - \frac{g'}{2}Y_f\sin\theta_\mathrm{W}\end{pmatrix}\begin{pmatrix}A_\mu \\ Z_\mu\end{pmatrix}.
\end{aligned}$$

Recalling Equation (14), the factor multiplying A_μ simplifies to

$$\frac{g'}{2}Y_f\cos\theta_\mathrm{W} + gI_3\sin\theta_\mathrm{W} = e\left(\frac{Y_f}{2} + I_3\right),$$

confirming the statement from Equation (17). The coupling to the Z boson is more complicated. Expressing Y_f in terms of the weak isospin and electric charge, and expressing g' in terms of g and θ_W, the factor multiplying Z_μ becomes

$$\begin{aligned}
gI_3\cos\theta_\mathrm{W} - \frac{g'}{2}Y_f\sin\theta_\mathrm{W} &= gI_3\cos\theta_\mathrm{W} - g(Q_f - I_3)\tan\theta_\mathrm{W}\sin\theta_\mathrm{W} \\
&= \frac{g}{\cos\theta_\mathrm{W}}\left[I_3\cos^2\theta_\mathrm{W} - (Q_f - I_3)\sin^2\theta_\mathrm{W}\right] \\
&= \frac{g}{\cos\theta_\mathrm{W}}\left(I_3 - Q_f\sin^2\theta_\mathrm{W}\right).
\end{aligned}$$

As noted in the question, the factor $g/\cos\theta_\mathrm{W}$ is by convention absorbed into the coupling strength, and thus the weak "charge" for coupling to the Z boson is $(I_3 - Q_f\sin^2\theta_\mathrm{W})$.

4 Lepton masses: the CKM and PMNS matrices

We end this section of the course on gauge symmetry with a few notes about one of the most important effects of extending the SM to three generations. In this case, the notation of Section 2 needs to be extended, so that ℓ_{L}, e_{R}, q_{L}, u_{R} and d_{R} acquire an extra generation index $i \in \{1, 2, 3\}$. For example, $q_{2\mathrm{L}}$ contains the left-handed charm and strange quarks.

The first place in the Lagrangian density that the fermion fields enter is in line (4a), where the covariant derivatives act on them. When more than one generation exists, we can choose to align the generations such that the simple sum over fermions remains valid, producing terms like $i\overline{q_{1\mathrm{L}}}\gamma^\mu \mathcal{D}_\mu^{q_\mathrm{L}} q_{1\mathrm{L}}$ but not $i\overline{q_{2\mathrm{L}}}\gamma^\mu \mathcal{D}_\mu^{q_\mathrm{L}} q_{1\mathrm{L}}$, for example. These fields ($q_{2\mathrm{L}}$, $u_{1\mathrm{R}}$ etc.) define the *interaction basis* of the fermion fields, and each generation remains indepdendent.

However, by making this choice, we are then not able to prevent additional Higgs Yukawa terms such as $\overline{q_{2\mathrm{L}}}\phi y_{u21} d_{1\mathrm{R}}$. This mixes the first and second generations, but satisfies all SM gauge symmetries because all of the $\mathrm{SU}(2)_\mathrm{L}$ doublets $q_{i\mathrm{L}}$ transform in exactly the same way. After all flavour combinations are taken into account, and considering just quarks for the moment, this means that y_u and y_d now become 3×3 (complex) matrices, rather than just numbers.

These matrices can be diagonalised, as shown in the lecture notes, by rotating the quark fields in flavour space with a unitary transformation ($u'_{i\mathrm{L}} = U^\dagger_{uij} u_{j\mathrm{L}}$ etc.). This defines the *mass basis* of the flavours, so-called because it diagonalises the mass terms. It describes the states that would propagate freely in the absence of interactions.

Now we consider the gauge interactions between mass eigenstates. All of the terms in Equation (4a) involving right-handed fields are of the form $\overline{u_{i\mathrm{R}}} X u_{i\mathrm{R}}$ etc., where X is some operator, and are in fact unaffected by this transformation. Similarly, the electromagnetic and strong interaction terms for left-handed quarks are of the form $\overline{q_{i\mathrm{L}}} X q_{i\mathrm{L}}$, where X is proportional to the $\mathrm{SU}(2)_\mathrm{L}$ identity matrix. These terms are also unaffected by the transformation, so that these forces remain flavour-diagonal.

Exercise 7: Demonstrate that flavour-diagonal Lagrangian density terms remain flavour-diagonal under a rotation of the flavour basis. For example, you could consider the kinetic terms $i\overline{q_1}\gamma^\mu\partial_\mu q_1 + i\overline{q_2}\gamma^\mu\partial_\mu q_2$ under an arbitrary rotation between the (generically named) q_1 and q_2 fields.

Answer: The kinetic terms in the question can be conveniently written using matrix notation

$$i\overline{q_1}\gamma^\mu\partial_\mu q_1 + i\overline{q_2}\gamma^\mu\partial_\mu q_2 = i\overline{q}\gamma^\mu\partial_\mu q,$$

$$\text{where } q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}.$$

We then define a rotation of the two fields, which we can write using an orthogonal matrix R:

$$q' = Rq = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}.$$

The rotation matrix, as usual, satisfies the normalisation condition $R^{\mathrm{T}}R = 1$. Upon applying this rotation to the Lagrangian density terms, we find that it does not change:

$$\begin{aligned} i\overline{q'}\gamma^\mu\partial_\mu q' &= i\overline{q}R^{\mathrm{T}}\gamma^\mu\partial_\mu Rq \\ &= i\overline{q}\gamma^\mu\partial_\mu q. \end{aligned}$$

This works because R is a constant matrix, which therefore commutes with the differential operator.

It is a different matter when we come to weak interactions, in particular charged current interactions. The Lagrangian density term is again of the form $\overline{q_{i\mathrm{L}}}Xq_{i\mathrm{L}}$, but X has off-diagonal terms equal to those shown in Equation (10). When expanded, this results in terms like $\overline{u_{i\mathrm{L}}}\gamma^\mu W_\mu^- d_{i\mathrm{L}}$, which transform to $\overline{u'_{i\mathrm{L}}}U^*_{uij}\gamma^\mu W_\mu^- U_{djk}d'_{k\mathrm{L}}$.[5] The matrix $U_u^\dagger U_d$ is called the *CKM matrix*, after Cabbibo, Kobayashi and Maskawa, and it describes the changes between quark generations that can occur in charged-current weak interactions.

If we assume that a right-handed neutrino exists, then a similar matrix can be obtained for the leptons. This is the so-called *PMNS matrix* (after Pontecorvo, Maki, Nakagawa and Sakata), and could be responsible for neutrino oscillations. The form of the PMNS matrix is similar to the CKM matrix, although the numerical values of the elements are very different. However, the ν_{R} is essentially unobservable, as it is a singlet of all three Standard Model gauge groups, and so its existence remains unconfirmed. We will return to both of these topics later in the course.

[5] The first matrix is really $U_u^\dagger$, but the transposition has already been applied by swapping the indices: $(U_u^\dagger)_{ji} = U^*_{uij}$.