7

Interaction Theory

Summary

In this chapter we will derive how different fields and particles interact with each other. This will enable us, for example, to describe how electrons, interact with photons[1].

[1] From a different point of view: how the electron field (= a massive spin $\frac{1}{2}$ field) interacts with the photon field (= a massless spin 1 field).

We will be guided to the correct form of the Lagrangians by internal symmetries, which are in this context often called gauge symmetries[2]. The starting point will be *local* $U(1)$ symmetry[3] and we end up with the Lagrangian

[2] This name will be explained in a moment.

[3] This means we multiply the field at each point in spacetime with a different factor: $e^{i\alpha(x)}$, instead of using the same transformation everywhere: $e^{i\alpha}$. In other words: the transformation parameter $\alpha = \alpha(x)$ is now a function of x and has therefore a different value for different points in spacetime.

$$\mathscr{L} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + A_\mu\bar{\Psi}\gamma^\mu\Psi + \frac{1}{2}\left(\partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu\right) .$$

This is the Lagrangian of **quantum electrodynamics**. This Lagrangian describes the interaction between charged, massive spin $\frac{1}{2}$ fields and a massless spin 1 field (the photon field). The Lagrangian is only then locally $U(1)$ invariant, if we avoid "mass terms" of the form $mA_\mu A^\mu$ in the Lagrangian. This coincides with the experimental fact that photons, described by A^μ, are massless. Using Noether's theorem, we can derive a new conserved quantity from $U(1)$ symmetry, which is commonly interpreted as **electric charge**.

Then we move on to local $SU(2)$ symmetry. For this purpose a two component object

$$\bar{\Psi} := \begin{pmatrix} \bar{\psi}_1 & \bar{\psi}_2 \end{pmatrix},$$

called **doublet**, is introduced. Such a doublet contains two spin $\frac{1}{2}$ fields, for example, the electron and the electron neutrino field that are "rotated" by $SU(2)$ transformations into each other.

Using this doublet notation we are able to write down a locally $SU(2)$ invariant Lagrangian[4]

[4] The object $W_j^{\mu\nu}$ is for the three fields W_j^μ what $F_{\mu\nu}$ is for the $U(1)$ gauge field A^μ.

J. Schwichtenberg, *Physics from Symmetry*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-319-66631-0_7

$$\mathscr{L} = i\bar{\Psi}\gamma_\mu \left(\partial^\mu - ig\mathcal{W}^\mu\right)\Psi - \frac{1}{4}Tr\big(\mathcal{W}_{\mu\nu}\mathcal{W}^{\mu\nu}\big),$$

which contains three spin 1 fields $\mathcal{W}^\mu \equiv (W^\mu)_i \frac{\sigma^i}{2}$, with $i = 1, 2, 3$. We need three fields to make the Lagrangian locally $SU(2)$ invariant, because we have three $SU(2)$ basis generators: $J_i = \frac{\sigma_i}{2}$. We will see that local $SU(2)$ symmetry can only be achieved without "mass terms" of the form $m\bar{\Psi}\Psi$, $m\mathcal{W}^\mu\mathcal{W}_\mu$, with some arbitrary mass **matrix** m, because Ψ is now a two-component object. So this time not only are the spin 1 fields $(W^\mu)_i$ required to be massless, but the spin $\frac{1}{2}$ fields, too. Also allowed are equal masses for the two spin $\frac{1}{2}$ fields, but from experiments we know that this is not the case: The electron mass is much bigger than the electron-neutrino mass. In addition, we know from experiments that the three spin 1 fields $(W^\mu)_i$ are not massless. This is commonly interpreted as the $SU(2)$ symmetry being **broken**.

This idea is the starting point for the **Higgs formalism**, which is introduced afterwards. This formalism enables us to get a locally $SU(2)$ invariant Lagrangian that includes mass terms. This works by adding the interaction with a spin 0 field, called Higgs field, into our considerations. The final locally $SU(2) \times U(1)$ invariant interaction Lagrangian describes the so called **electroweak interactions**. The notion electroweak interactions contains the electromagnetic interaction and a new type of interaction called the **weak interaction**. Through the Higgs formalism the $SU(2) \times U(1)$ is broken to a remnant $U(1)$ symmetry. The weak interaction is mediated by three massive spin 1 fields, called W^+,W^- and Z and the remnant $U(1)$ symmetry by one massless spin 1 field γ. Using Noether's theorem we will be able to derive from $SU(2)$ symmetry a new conserved quantity, called **isospin**, which is the charge of weak interactions analogous to electric charge for electromagnetic interactions.

Lastly, we will consider internal *local* $SU(3)$ symmetry, which will lead us to a Lagrangian describing another new interaction, called the **strong interaction**. For this purpose, we will introduce triplet objects

$$Q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix},$$

that are transformed by $SU(3)$ transformations and which contain three spin $\frac{1}{2}$ fields. These three spin $\frac{1}{2}$ fields are interpreted as **quarks** carrying different **color**, which is the strong interaction analogue to the electrical charge of the electromagnetic interaction or isospin of the weak interaction. Again, mass terms are forbidden, but this time this coincides with the experimental fact that the 8 corresponding bosons[5], called gluons, are massless. In addition, we know from experiments that the fields inside a $SU(3)$ triplet have the same mass, This is a good thing, because local $SU(3)$ symmetry forbids a term with arbitrary mass matrix m for terms like $m\bar{Q}Q$, but allows a term of the form $\bar{Q}\begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix}Q$, which means that the terms in the triplet have the same mass. Therefore local $SU(3)$ invariance provides no new obstacles regarding mass terms in the Lagrangian and we therefore say the $SU(3)$ symmetry is unbroken. From experiments we know that only quarks (spin $\frac{1}{2}$) and gluons (spin 1) carry color. The resulting Lagrangian

[5] 8 because $SU(3)$ has 8 basis generators.

$$\mathscr{L} = -\frac{1}{4}\mathcal{G}_{\alpha\beta}\mathcal{G}^{\alpha\beta} + \overline{Q}(i(\partial_\mu - ig\mathcal{G}_\mu)\gamma^\mu - m)Q$$

will only be cited, because the derivation is completely analogous to what we did before.

To summarize the summary:

$U(1)$ ⟶ 1 gauge field ⟶ **massless** photons ⟶ electric charge

$SU(2)$ ⟶ 3 gauge fields ⟶ **massive** W- and Z-bosons (Higgs needed) ⟶ isospin

$SU(3)$ ⟶ 8 gauge fields ⟶ **massless** gluons ⟶ color charge

7.1 $U(1)$ Interactions

To derive the correct interaction terms in the Lagrangian, we are going to use internal symmetrie that are usually called gauge symmetries. The notion gauge symmetry, is used for historic reasons and doesn't make much sense for the type of symmetry we are considering here. Weyl tried to derive electromagnetism[6]

> "as a consequence of spacetime symmetry, specifically symmetry under local changes of length scale."

[6] Frank Wilczek. Riemann-einstein structure from volume and gauge symmetry. *Phys. Rev. Lett.*, 80:4851–4854, Jun 1998. DOI: 10.1103/PhysRevLett.80.4851

Naming this kind of symmetry **gauge symmetry** makes sense, because this means, for example, that we can change the platinum bar that defines a standard meter (and which was used to gauge objects that measure length in experiments), arbitrarily without changing physics. This attempt was unsuccessful, but some time later, Weyl found the correct symmetry to derive electromagnetism and the name was kept.

7.1.1 Internal Symmetry of Free Spin $\frac{1}{2}$ Fields

Let's have a look again at the Lagrangian that we derived for a free spin $\frac{1}{2}$ theory (Eq. 6.16)

$$\mathscr{L}_{\text{Dirac}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi = \bar{\Psi}(i\gamma_\mu\partial^\mu - m)\Psi. \tag{7.1}$$

We derived it by demanding Lorentz symmetry, but if we take a sharp look we can discover another symmetry of this Lagrangian. The Lagrangian does not change if we transform the field Ψ as follows

$$\Psi \to \Psi' = e^{ia}\Psi$$

if we take into account that this implies that $\bar{\Psi}$ also gets transformed

$$\Rightarrow \bar{\Psi} \to \bar{\Psi}' = \Psi'^{\dagger}\gamma_0 = (e^{ia}\Psi)^{\dagger}\gamma_0 = \bar{\Psi}e^{-ia}. \tag{7.2}$$

The minus sign comes from the complex conjugation[7] and a is an arbitrary real number. To see that the Lagrangian is invariant under this transformation, we transform the Lagrangian explicitly:

[7] Remember: $\bar{\Psi} = \Psi^{\dagger}\gamma_0$.

$$\begin{aligned}\mathscr{L}'_{\text{Dirac}} &= -m\bar{\Psi}'\Psi' + i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' \\ &= -m(\bar{\Psi}e^{-ia})(e^{ia}\Psi) + i(\bar{\Psi}e^{-ia})\gamma_\mu\partial^\mu(e^{ia}\Psi) \\ &= -m\bar{\Psi}\Psi \underbrace{e^{-ia}e^{ia}}_{=1} + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi \underbrace{e^{-ia}e^{ia}}_{=1} \\ &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi = \mathscr{L}_{\text{Dirac}},\end{aligned} \tag{7.3}$$

where we used that e^{ia} is just a complex number, which we can move around freely[8]. Remembering that we learned in Chapter 3 that all unit complex numbers can be written as e^{ia} and form a group called $U(1)$, we can put what we just discovered into mathematical terms, by saying that the Lagrangian is $U(1)$ invariant. This symmetry is an internal symmetry, because it is clearly no spacetime transformation and therefore transforms the field internally. This internal symmetry does not look like a big thing. At a first glance it may seem like a nice, but rather useless, mathematical side note. However, quite surprisingly we will see in a moment that this observation is incredibly important!

[8] Speaking more technically: A complex number commutes with every matrix, like, for example, γ_μ.

Now let's have a deeper look at what we just discovered. We showed that we are free to multiply our field with an arbitrary unit complex number without changing anything. The symmetry transformation $\Psi \to \Psi' = e^{ia}\Psi$ is called a **global** transformation, because we multiply the field $\Psi = \Psi(x)$ at every point x with the same factor e^{ia}.

Now, why should this factor at one point in spacetime be correlated to the factor at another point in spacetime? The choice at one point in spacetime shouldn't fix this immediately in the whole universe. This would be strange, because special relativity tells us that no information can spread faster than light, as was shown in Section 2.4. For a global symmetry the choice would be fixed immediately for any point in the whole universe.

Let's check if our Lagrangian is invariant if we transform each point in spacetime with a different factor $a = a(x)$. This is called a **local** transformation.

If we transform

$$\Psi \rightarrow \Psi' = e^{ia(x)}\Psi$$
$$\Rightarrow \bar{\Psi} \rightarrow \bar{\Psi}' = e^{-ia(x)}\bar{\Psi}, \tag{7.4}$$

where the factor $a = a(x)$ now depends on the position, we get the transformed Lagrangian[9]

$$\begin{aligned}\mathscr{L}'_{\text{Dirac}} &= -m\bar{\Psi}'\Psi' + i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' \\ &= -m(\bar{\Psi}\underbrace{e^{-ia(x)})(e^{ia(x)}}_{=1}\Psi) + i(\bar{\Psi}e^{-ia(x)})\gamma_\mu\partial^\mu(e^{ia(x)}\Psi) \\ &\underbrace{=}_{\text{Product rule}} -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu(\partial^\mu\Psi)\underbrace{e^{-ia(x)}e^{ia(x)}}_{=1} + i(e^{-ia(x)}\bar{\Psi})\gamma_\mu\Psi(\partial^\mu e^{ia(x)}) \\ &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + i^2(\partial^\mu a(x))\bar{\Psi}\gamma_\mu\Psi \neq \mathscr{L}_{\text{Dirac}}\end{aligned} \tag{7.5}$$

Therefore, our Lagrangian is *not* invariant under local $U(1)$ symmetry, because the product rule produces an extra term. As discussed above, our Lagrangian should be locally invariant, but we just found out that it isn't. There is something we can do about it, but first we must investigate another symmetry.

[9] Maybe you wonder if the Lagrangian that includes both possible derivatives, which we neglected for brevity, is locally $U(1)$ invariant: $\mathscr{L} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + i(\partial^\mu\bar{\Psi})\gamma_\mu\Psi$. This Lagrangian is indeed locally $U(1)$ invariant, as you can check, but take note that the addition of the second and the third term yields zero:

$$i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + i(\partial^\mu\bar{\Psi})\gamma_\mu\Psi \underbrace{=}_{\text{integration by parts}} i(\partial^\mu\bar{\Psi})\gamma_\mu\Psi - i(\partial^\mu\bar{\Psi})\gamma_\mu\Psi = 0.$$

The correct Lagrangian that includes both possible derivatives has a minus sign between those terms: $\mathscr{L}_{\text{Dirac}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - i(\partial^\mu\bar{\Psi})\gamma_\mu\Psi$ and is therefore **not** locally $U(1)$ invariant.

7.1.2 Internal Symmetry of Free Spin 1 Fields

Next, let's take a look at the Lagrangian we derived for free spin 1 particles[10]

$$\mathscr{L}_{\text{Proca}} = \partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu + m^2 A^\mu A_\mu. \tag{7.6}$$

[10] See Eq. 6.26 and take note that we neglect, for brevity, a conventional factor $\frac{1}{2}$ here.

We can discover a **global** internal symmetry here, too. If we transform

$$A_\mu \rightarrow A'_\mu = A_\mu + a_\mu \tag{7.7}$$

with some arbitrary constants a_μ[11], the Lagrangian reads[12]

$$\begin{aligned}\mathscr{L}'_{\text{Proca}} &= (\partial^\mu A'^\nu\partial_\mu A'_\nu - \partial^\mu A'^\nu\partial_\nu A'_\mu) + m^2 A'^\mu A'_\mu \\ &= \partial^\mu(A^\nu + a^\nu)\partial_\mu(A_\nu + a_\nu) - \partial^\mu(A^\nu + a^\nu)\partial_\nu(A_\mu + a_\mu)) + m^2(A_\mu + a_\mu)(A^\mu + a^\mu) \\ &= \partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu + m^2(A_\mu + a_\mu)(A^\mu + a^\mu).\end{aligned} \tag{7.8}$$

[11] Remember, as always, $\mu = 0, 1, 2, 3$.

[12] The a_μ are constants and therefore $\partial^\mu a^\nu = 0$ etc.

We conclude this transformation is a **global** symmetry transformation of this Lagrangian, if we restrict ourselves to **massless** fields, i.e. $m = 0$.

What about **local** symmetry here? We transform

$$A_\mu \to A'_\mu = A_\mu + a_\mu(x) \tag{7.9}$$

and the transformed **massless** Lagrangian reads

$$\begin{aligned}
\mathscr{L}'_{\text{Maxwell}} &= (\partial^\mu A'^\nu \partial_\mu A'_\nu - \partial^\mu A'^\nu \partial_\nu A'_\mu) \\
&= \partial^\mu (A^\nu + a^\nu(x)) \partial_\mu (A_\nu + a_\nu(x)) - \partial^\mu (A^\nu + a^\nu(x)) \partial_\nu (A_\mu + a_\mu(x)) \\
&= \partial^\mu A^\nu \partial_\mu A_\nu + \partial^\mu a^\nu \partial_\mu A_\nu + \partial^\mu A^\nu \partial_\mu a_\nu(x) + \partial^\mu a^\nu(x) \partial_\mu a_\nu(x) \\
&\quad - \partial^\mu A^\nu \partial_\nu A_\mu - \partial^\mu A^\nu \partial_\nu a_\mu(x) - \partial^\mu a^\nu(x) \partial_\nu A_\mu - \partial^\mu a^\nu(x) \partial_\nu a_\mu(x) \, ,
\end{aligned} \tag{7.10}$$

which shows that $A_\mu \to A'_\mu = A_\mu + a_\mu(x)$ is *not* a local internal symmetry .

Nevertheless, we can find a local internal symmetry by considering the transformation $A_\mu \to A'_\mu = A_\mu + \partial_\mu a(x)$ instead. This means we add the derivative of some arbitrary function $\partial_\mu a(x)$ instead of an arbitrary function. This transformation has the following effect on the Lagrangian[13]

[13] The symmetry of partial derivatives $\partial_\nu \partial_\mu = \partial_\mu \partial_\nu$ is also known as "Schwarz's theorem".

$$\begin{aligned}
\mathscr{L}'_{\text{Maxwell}} &= \partial^\mu A'^\nu \partial_\mu A'_\nu - \partial^\mu A'^\nu \partial_\nu A'_\mu \\
&= \partial^\mu (A^\nu + \partial^\nu a(x)) \partial_\mu (A_\nu + \partial_\nu a(x)) - \partial^\mu (A^\nu + \partial^\nu a(x)) \partial_\nu (A_\mu + \partial_\mu a(x)) \\
&= \partial^\mu A^\nu \partial_\mu A_\nu + \partial^\mu (\partial^\nu a(x)) \partial_\mu A_\nu + \partial^\mu A^\nu \partial_\mu (\partial_\nu a(x)) + \partial^\mu (\partial^\nu a(x)) \partial_\mu (\partial_\nu a(x)) \\
&\quad - \partial^\mu A^\nu \partial_\nu A_\mu - \partial^\mu A^\nu \partial_\nu (\partial_\mu a(x)) - \partial^\mu (\partial^\nu a(x)) \partial_\nu A_\mu - \partial^\mu (\partial^\nu a(x)) \partial_\nu (\partial_\mu a(x)) \\
&\underbrace{=}_{\partial_\nu \partial_\mu = \partial_\mu \partial_\nu \text{ and renaming dummy indices}} \partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu = \mathscr{L}_{\text{Maxwell}} \, .
\end{aligned} \tag{7.11}$$

We see this is indeed an internal local symmetry transformation. Again this may look like a technical side note. Okay, we found some internal, local symmetry; so what?

7.1.3 Putting the Puzzle Pieces Together

Let's summarize what we found out so far:

- We discovered the Lagrangian for free spin $\frac{1}{2}$ fields has an internal global symmetry $\Psi \to \Psi' = e^{ia}\Psi$. Formulated differently: the Lagrangian for free spin $\frac{1}{2}$ fields is invariant under global $U(1)$ transformations.

- We saw that this symmetry is not local (although it should be), because for $a = a(x)$, we get an extra term in the Lagrangian of the form (Eq. 7.5).
$$-(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi. \tag{7.12}$$
In other words: The Lagrangian isn't locally $U(1)$ invariant.

- In the last section we found an internal local symmetry for massless spin 1 fields
$$A_\mu \to A'_\mu = A_\mu + \partial_\mu a(x), \tag{7.13}$$
which is only a local symmetry if we add the **derivative** of an arbitrary function $\partial_\mu a(x)$, instead of an arbitrary function $a_\mu(x)$.

This really looks like two pieces of a puzzle we should put together. When we transform Ψ, $\bar{\Psi}$ and A_μ simultaneously, an additional term $A_\mu\bar{\Psi}\gamma^\mu\Psi$ in the Lagrangian becomes

$$A_\mu\bar{\Psi}\gamma^\mu\Psi \to (A_\mu + \partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi = A_\mu\bar{\Psi}\gamma^\mu\Psi + \partial_\mu a(x)\bar{\Psi}\gamma^\mu\Psi\,. \tag{7.14}$$

Compare the second term here to Eq. 7.12. The new term in the Lagrangian, coupling Ψ, $\bar{\Psi}$ and A_μ together, therefore cancels exactly the term which stopped the Lagrangian for free spin $\frac{1}{2}$ from being locally $U(1)$ invariant. In other words: **By adding this new term we can make the Lagrangian locally $U(1)$ invariant**.

Let's study this in more detail. First take note that it's conventional to factor out a constant g in the exponent of the local $U(1)$ transformation: $e^{iga(x)}$. Then the extra term becomes

$$-(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi \to -g(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi. \tag{7.15}$$

This extra factor g accounts for an arbitrary coupling constant[14], as we will see now. We then add to the Lagrangian for free spin $\frac{1}{2}$ fields the new term

$$gA_\mu\bar{\Psi}\gamma^\mu\Psi,$$

where we included γ^μ to make the term Lorentz invariant[15] and inserted the **coupling constant**[16] g. This yields the Lagrangian

$$\mathscr{L}_{\text{Dirac+Extra-Term}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + gA_\mu\bar{\Psi}\gamma^\mu\Psi.$$

Transforming this Lagrangian according to the rules for local transformations of Ψ, $\bar{\Psi}$ and A_μ yields[17]

[14] A coupling constant always tells us how strong a given interaction is. Here we are talking about electromagnetic interactions and g determines its strength.

[15] Otherwise A_μ has an unmatched index μ and therefore wouldn't be Lorentz invariant.

[16] We can see here that g determines how strong Ψ, $\bar{\Psi}$ and A_μ couple together.

[17] The combined transformation of Ψ, $\bar{\Psi}$ and A_μ is called $U(1)$ **gauge transformation**.

$$
\begin{aligned}
\mathscr{L}'_{\text{Dirac+Extra-Term}} &= -m\bar{\Psi}'\Psi' + i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' + gA'_\mu\bar{\Psi}'\gamma^\mu\Psi' \\
&\underbrace{=}_{\text{See Eq. 7.5}} -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - g(\partial^\mu a(x))\bar{\Psi}\gamma_\mu\Psi + gA'_\mu\bar{\Psi}'\gamma^\mu\Psi' \\
&= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - g(\partial^\mu a(x))\bar{\Psi}\gamma_\mu\Psi \\
&\quad + g(A_\mu + \partial_\mu a(x))(e^{-iga(x)}\bar{\Psi})\gamma^\mu(e^{iga(x)}\Psi) \\
&= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - \cancel{g(\partial^\mu a(x))\bar{\Psi}\gamma_\mu\Psi} \\
&\quad + gA_\mu\bar{\Psi}\gamma^\mu\Psi + g\underbrace{\cancel{(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi}}_{=(\partial^\mu a(x))\bar{\Psi}\gamma_\mu\Psi} \\
&= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + gA_\mu\bar{\Psi}\gamma^\mu\Psi = \mathscr{L}_{\text{Dirac+Extra-Term}}.
\end{aligned}
\tag{7.16}
$$

Therefore, by adding an extra term we get a locally $U(1)$ invariant Lagrangian. To describe a system consisting of massive spin $\frac{1}{2}$ and massless spin 1 fields we must add the Lagrangian for free massless spin 1 fields to the Lagrangian as well. This gives us the complete Lagrangian[18]

[18] We use here the conventional "normalization" with an additional factor $\frac{1}{2}$ in front of the last terms.

$$
\begin{aligned}
\mathscr{L}_{\text{Dirac+Extra-Term+Maxwell}} &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + gA_\mu\bar{\Psi}\gamma^\mu\Psi \\
&\quad - \frac{1}{2}(\partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu).
\end{aligned}
\tag{7.17}
$$

To unclutter the notation it is conventional to introduce a new symbol

$$D^\mu \equiv i\partial^\mu - igA^\mu, \tag{7.18}$$

called **covariant derivative**. The Lagrangian then reads

$$
\begin{aligned}
\mathscr{L}_{\text{Dirac+Extra-Term+Maxwell}} &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\underbrace{(\partial^\mu - igA_\mu)}_{\equiv D^\mu}\Psi \\
&\quad - \frac{1}{2}(\partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu) \\
&= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu D^\mu\Psi \\
&\quad - \frac{1}{2}(\partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu).
\end{aligned}
\tag{7.19}
$$

This is the correct Lagrangian for the quantum field theory of electrodynamics, commonly called **quantum electrodynamics**. We are able to derive this Lagrangian simply by making use of internal symmetries of the Lagrangians describing free spin $\frac{1}{2}$ fields and free spin 1 fields.

The next question we have to answer is: What equations of motion follow from this Lagrangian?

7.1.4 Inhomogeneous Maxwell Equations and Minimal Coupling

To spoil the surprise: This Lagrangian gives us the inhomogeneous Maxwell equations in the presence of currents.

The process is again straightforward: we simply put the Lagrangian (Eq. 7.17)

$$\begin{aligned}\mathscr{L}_{\text{Dirac+Extra-Term+Maxwell}} = &- m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + gA_\mu\bar{\Psi}\gamma^\mu\Psi \\ &- \frac{1}{2}(\partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu).\end{aligned}$$

into the Euler-Lagrange equation for each field

$$\frac{\partial\mathscr{L}}{\partial\Psi} - \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\Psi)}\right) = 0$$

$$\frac{\partial\mathscr{L}}{\partial\bar{\Psi}} - \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\bar{\Psi})}\right) = 0$$

$$\frac{\partial\mathscr{L}}{\partial A_\rho} - \partial_\sigma\left(\frac{\partial\mathscr{L}}{\partial(\partial_\sigma A_\rho)}\right) = 0.$$

This yields

$$\bar{\Psi}(i\gamma_\mu\partial^\mu + m) + gA_\mu\bar{\Psi}\gamma^\mu = 0 \tag{7.20}$$

$$(i\gamma_\mu\partial^\mu - m)\Psi + gA_\mu\gamma^\mu\Psi = 0 \tag{7.21}$$

$$\partial_\nu(\partial^\nu A^\mu - \partial^\mu A^\nu) + g\bar{\Psi}\gamma^\mu\Psi = 0\,. \tag{7.22}$$

The first two equations describe the behavior of spin $\frac{1}{2}$ particles/fields in an external electromagnetic field. In many books the derivation of these equations uses the notion **minimal coupling**, by which is meant that in the presence of an external field the derivative ∂^μ has to be changed into the covariant derivative

$$\partial^\mu \to D^\mu = \partial^\mu - igA^\mu \tag{7.23}$$

to yield the correct equations. The word "minimal" is used, because only one gauge field A_μ, with four components $\mu = 0, 1, 2, 3$, is used.

Now that we have the equation that describes how Dirac spinors behave in the presence of an external electromagnetic field (Eq. 7.21), we can show something that we promised in Section 3.7.10. There we claimed that a transformation, which we called very suggestively charge conjugation, changes the electric charge of the object it describes. In other words, if Ψ describes something of charge $+e$, the charge conjugate spinor Ψ^C describes something of charge $-e$.

Electrical charge determines the coupling strength of spin $\frac{1}{2}$ particles/fields to an external spin 1 field and we therefore investigate now, which equation of motion holds for Ψ^C. Afterwards we will talk about the third equation, i.e. Eq. 7.22.

7.1.5 Charge Conjugation, Again

Before we can derive the corresponding equation, we need to find an explicit form of the charge conjugation operator for Dirac spinors. We derived in Section 3.7.10 the transformation (Eq. 3.244)

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \rightarrow \Psi^C = \begin{pmatrix} \xi_L \\ \chi_R \end{pmatrix}. \tag{7.24}$$

This transformation can now be described easily with the help of one of the γ_μ matrices. Using the definition of γ_2 in Eq. 6.13, we have

$$\Psi^C = i\gamma_2\Psi^\star = i\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix}\begin{pmatrix} \chi_L^\star \\ \xi_R^\star \end{pmatrix}, \tag{7.25}$$

because we can rewrite this, using that $i\sigma_2 = \epsilon$ is exactly the spinor metric

$$= \begin{pmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{pmatrix}\begin{pmatrix} \chi_L^\star \\ \xi_R^\star \end{pmatrix} = \begin{pmatrix} \epsilon\xi_R^\star \\ -\epsilon\chi_L^\star \end{pmatrix}. \tag{7.26}$$

This is equivalent to

$$= \begin{pmatrix} \xi_L \\ \chi_R \end{pmatrix}, \tag{7.27}$$

as was shown in Section 3.7.7, specifically Eq. 3.202. Therefore we start with Eq. 7.21:

$$(i\gamma_\mu\partial^\mu - m)\Psi + gA_\mu\gamma^\mu\Psi = \big(\gamma_\mu(i\partial^\mu + gA^\mu) - m\big)\Psi = 0, \tag{7.28}$$

and complex conjugate this equation, as a first step towards an equation for Ψ^C:

$$\rightarrow \big(\gamma_\mu^\star(-i\partial^\mu + gA^\mu) - m\big)\Psi^\star = 0. \tag{7.29}$$

Now, we multiply this equation from the left-hand side with γ_2 and add a $1 = \gamma_2^{-1}\gamma_2$ in front of $\Psi^\star$:

$$\rightarrow \gamma_2\big(\gamma_\mu^\star(-i\partial^\mu + gA^\mu) - m\big)\underbrace{\gamma_2^{-1}\gamma_2}_{=1}\Psi^\star = 0 \tag{7.30}$$

$$\rightarrow \big(\underbrace{\gamma_2\gamma_\mu^\star\gamma_2^{-1}}_{=-\gamma_\mu}(-i\partial^\mu + gA^\mu) - m\gamma_2\gamma_2^{-1}\big)\gamma_2\Psi^\star = 0 \tag{7.31}$$

$$\underbrace{\rightarrow}_{\text{Multiplying the equation with } i} \left(-\gamma_\mu(-i\partial^\mu + gA^\mu) - m \right) \underbrace{i\gamma_2\Psi^\star}_{=\Psi^C \text{ see Eq. 7.25}} = 0 \tag{7.32}$$

$$\rightarrow \left((\gamma_\mu(i\partial^\mu - gA^\mu) - m \right)\Psi^C = 0\,. \tag{7.33}$$

This is exactly the same equation of motion as for Ψ, but with opposite coupling strength $g \rightarrow -g$. This justifies the name charge conjugation[19].

[19] It is important to note that $\Psi^c \neq \bar{\Psi}$. $\Psi^c = i\gamma_2\Psi^\star$ and $\bar{\Psi} = \Psi^\dagger\gamma_0 = (\Psi^\star)^T\gamma_0$. Charge conjugation is the correct transformation that enables us to interpret things in terms of antiparticles, as we will discuss later in detail.

Next, we turn to the third equation of motion derived in the last section (Eq. 7.22) which is called **inhomogeneous Maxwell equation in the presence of an electric current**. To make the notion "electric current" precise we need again Noether's theorem.

7.1.6 Noether's Theorem for Internal $U(1)$ Symmetry

In Section 4.5.5 we learned that Noether's theorem connects each internal symmetry with a conserved quantity. What conserved quantity follows from the $U(1)$ symmetry we just discovered? Noether's theorem for internal symmetries tells us that a transformation of the form

$$\Psi \rightarrow \Psi' = \Psi + \delta\Psi$$

leads to a Noether current

$$J^\mu = \frac{\partial\mathscr{L}}{\partial(\partial_\mu\Psi)}\delta\Psi$$

which fulfils a continuity equation

$$\partial_\mu J^\mu = 0. \tag{7.34}$$

A global[20] $U(1)$ transformation is

$$\Psi \rightarrow \Psi' = \mathrm{e}^{iga}\Psi = (1 + iga + \ldots)\Psi.$$

[20] Recall that the Lagrangian for free spin $\frac{1}{2}$ fields was only *globally* $U(1)$ invariant. The final Lagrangian of the last section was *locally* $U(1)$ invariant. Global symmetry is a special case of local symmetry with a =const. Therefore, if we have a locally $U(1)$ invariant Lagrangian, it is automatically globally $U(1)$ invariant, too. Considering **global** $U(1)$ symmetry here will give us a quantity that is conserved for free **and** interacting fields.

We stop the series expansion of the exponential function, as usual, after the first term, because $U(1)$ is a Lie group and arbitrary transformations can be built of infinitesimal ones. An infinitesimal transformation reads

$$\Psi \rightarrow \Psi' = \Psi + iga\Psi.$$

Therefore we have $\delta\Psi = iga\Psi$ and as we derived in Section 4.5.5 the corresponding Noether current is

$$\begin{aligned} J^\mu &= \frac{\partial\mathscr{L}}{\partial(\partial_\mu\Psi)}\delta\Psi \\ &= \frac{\partial(-m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi)}{\partial(\partial_\mu\Psi)} iga\Psi \\ &= -\bar{\Psi}\gamma^\mu ga\Psi = -ga\bar{\Psi}\gamma^\mu\Psi. \end{aligned} \tag{7.35}$$

We can ignore[21] the arbitrary constant a, because the continuity equation holds for arbitrary a and therefore, we define

$$J^\mu \equiv -g\bar{\Psi}\gamma^\mu\Psi\,. \tag{7.36}$$

[21] We keep the conventional constant g, which is not arbitrary but has one fixed value that is determined in experiments.

This is usually called the **electric four-current**. The zeroth component is the **electric charge density**, which gives us if we integrate it, a quantity that is conserved in time[22]

$$Q = \int d^3x \underbrace{\rho}_{\text{Charge density}} = \int d^3x J^0 = -g\int d^3x \bar{\Psi}\gamma^0\Psi. \tag{7.37}$$

[22] This can be seen by following the same steps as in Eq. 4.39.

In the quantum framework the objects Ψ will be related to probability amplitudes and this interpretation requires that $\int d^3x\bar{\Psi}\gamma^0\Psi = 1$, because the overall probability must be $100\% = 1$. Therefore, the conserved quantity is in fact the coupling strength g, which is for electromagnetism the electric charge. **Therefore, global $U(1)$ symmetry leads to the conservation of electric charge**.

If we now take a look again at Eq. 7.22, we can write it, using the definition in Eq. 7.36, as

$$\partial_\nu(\partial^\nu A^\mu - \partial^\mu A^\nu) + \underbrace{g\bar{\Psi}\gamma^\mu\Psi}_{=-J^\mu} = 0$$

$$\rightarrow \partial_\nu(\partial^\nu A^\mu - \partial^\mu A^\nu) = J^\mu. \tag{7.38}$$

Using the electromagnetic tensor as defined in Eq. 6.23 this equation reads

$$\partial_\nu F^{\nu\mu} = J^\mu. \tag{7.39}$$

These are the inhomogeneous Maxwell equations in the presence of an external electromagnetic current. These equations[23], together with the homogeneous Maxwell equations, which follow immediately from the definition of $F^{\nu\mu}$, are the basis for the classical theory of electrodynamics[24].

[23] Plural, because we have an equation for each component μ.

[24] We will see this explicitly in Chapter 11.

Next we take a quick look at interactions of **massive** spin 1 and spin 0 fields.

7.1.7 Interaction of Massive Spin 0 Fields

Take note that the Lagrangian we derived for spin 0 fields

$$\mathscr{L} = \frac{1}{2}(\partial_\mu\Phi\partial^\mu\Phi - m^2\Phi^2)$$

is not $U(1)$ invariant, as we can see by transforming $\Phi \rightarrow \Phi' = e^{ia}\Phi$. Nevertheless, the complex scalar theory

$$\mathscr{L} = \frac{1}{2}(\partial_\mu\Phi^\star\partial^\mu\Phi - m^2\Phi^\star\Phi). \tag{7.40}$$

has $U(1)$ symmetry, because then we have $\Phi \to \Phi' = e^{ia}\Phi$ and $\Phi^\star \to (\Phi^\star)' = e^{-ia}\Phi^\star$. Therefore it is possible to derive, analogous to what we did in Section 7.1 for spin $\frac{1}{2}$ fields, an interaction theory for this Lagrangian. The derivation is completely analogous[25] and one gets

$$\mathscr{L} = \frac{1}{2}\left(\left((\partial_\mu + iqA_\mu)\Phi^\star\right)\left((\partial^\mu - iqA^\mu)\Phi\right) - m^2\Phi^\star\Phi\right). \qquad (7.41)$$

[25] The correct Lagrangian can be computed by substituting $\partial^\mu \to D^\mu = \partial^\mu - igA^\mu$ as introduced in Eq. 7.23.

Using the Euler-Lagrange equation

$$\frac{\partial \mathscr{L}}{\partial \Phi^\star} - \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi^\star)}\right) = 0,$$

we find the corresponding equation of motion

$$(\partial_\mu - iqA_\mu)(\partial^\mu - iqA^\mu)\Phi - m^2\Phi = 0, \qquad (7.42)$$

which describes a charged spin 0 field coupled to a massless spin 1 field.

7.1.8 Interaction of Massive Spin 1 Fields

The interaction of a **massive** spin 1 field with a **massless** spin 1 field is dictated by symmetry, too. The Lagrangian for massless spin 1 fields is given by (Eq. 6.25)

$$\mathscr{L}_{\text{Maxwell}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}.$$

To distinguish between a massless and a massive spin 1 field, we name the massive field B^μ and define

$$G^{\mu\nu} := \partial^\mu B^\nu - \partial^\nu B^\mu.$$

The Lagrangian for this massive spin 1 field reads (Eq. 6.19)

$$\mathscr{L}_{\text{Proca}} = -\frac{1}{4}G^{\mu\nu}G_{\mu\nu} + m^2 B^\mu B_\mu.$$

Lorentz symmetry dictates the interaction term in the Lagrangian to be of the form

$$\mathscr{L}_{\text{Proca-interaction}} = CG_{\mu\nu}F^{\mu\nu},$$

with a coupling constant C that we need to measure in experiments. If you're interested you can derive yourself the corresponding equations of motion, by using the Euler-Lagrange equations.

7.2 $SU(2)$ Interactions

Motivated by the success with $U(1)$ symmetry we want to answer the question: Is $U(1)$ the only internal symmetry of our Lagrangians?

It turns out that we can find an internal symmetry for **two massless** spin $\frac{1}{2}$ fields. We get the Lagrangian for two spin $\frac{1}{2}$ fields by adding two copies of the Lagrangian that we derived in Section 6.3. The final Lagrangian can be found in Eq. 6.16 and we recite it here for convenience:

$$\mathscr{L}_{\text{Dirac}} = \bar{\psi}(i\gamma_\mu \partial^\mu - m)\psi.$$

Here we neglect mass terms, which means $m = 0$, because otherwise the Lagrangian isn't invariant as we will see in a moment. We will see later how we can include mass terms, without spoiling the symmetry. The addition yields

$$\mathscr{L}_{\text{D1+D2}} = i\bar{\psi}_1 \gamma_\mu \partial^\mu \psi_1 + i\bar{\psi}_2 \gamma_\mu \partial^\mu \psi_2 \,. \tag{7.43}$$

This can be rewritten, if we define

$$\Psi := \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

$$\rightarrow \bar{\Psi} := \begin{pmatrix} \bar{\psi}_1 & \bar{\psi}_2 \end{pmatrix} ,$$

where the newly defined object Ψ is called a **doublet**:

$$\mathscr{L}_{\text{D1+D2}} = i\bar{\Psi}\gamma_\mu \partial^\mu \Psi. \tag{7.44}$$

This Lagrangian is invariant under **global** $SU(2)$ transformations[26]

$$\Psi \rightarrow \Psi' = \mathrm{e}^{ia_i \frac{\sigma_i}{2}} \Psi \tag{7.45}$$

$$\Rightarrow \bar{\Psi} \rightarrow \bar{\Psi}' = \bar{\Psi}\mathrm{e}^{-ia_i \frac{\sigma_i}{2}}, \tag{7.46}$$

where a sum over the index "i" is implicitly assumed, a_i denotes arbitrary real constants and $\frac{\sigma_i}{2}$ are the usual generators of $SU(2)$, with the Pauli matrices σ_i.

[26] To calculate the transformation behavior of $\bar{\Psi}$, we use $\sigma_i^\dagger = \sigma_i$, which we already noted in Eq. 3.213.

To see the invariance we take a look at the transformed Lagrangian[27]

$$\begin{aligned} \mathscr{L}'_{\text{D1+D2}} &= i\bar{\Psi}'\gamma_\mu \partial^\mu \Psi' \\ &= i\bar{\Psi}\mathrm{e}^{-ia_i \frac{\sigma_i}{2}} \gamma_\mu \partial^\mu \mathrm{e}^{ia_i \frac{\sigma_i}{2}} \Psi \\ &= i\bar{\Psi}\gamma_\mu \partial^\mu \Psi = \mathscr{L}_{\text{D1+D2}}, \qquad \checkmark \end{aligned} \tag{7.47}$$

where we got to the last line because our transformation $\mathrm{e}^{ia_i \frac{\sigma_i}{2}}$ acts on our newly defined two-component object Ψ, whereas γ_μ acts on the objects in our doublet, i.e. the Dirac spinors. We can express this using indices

$$\left[\left(e^{-ia_i \frac{\sigma_i}{2}}\right)_{ab} \delta_{\alpha\beta} \right] \left[\delta_{bc} \gamma_\mu^{\beta\delta} \right] \left[\left(e^{ia_i \frac{\sigma_i}{2}}\right)_{cd} \delta_{\delta\epsilon} \right] = \left[\delta_{ad} \gamma_\mu^{\alpha\epsilon} \right] \,.$$

[27] We neglect mass terms here, because they are, in general, not invariant under $SU(2)$ transformations. The mass terms would be $-m_1\bar{\Psi}_1\Psi_1$ and $-m_2\bar{\Psi}_2\Psi_2$ and we could write them, using the two component definition for Ψ and by defining

$$m := \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}$$

as

$$\mathscr{L}_{\text{D1+D2}} = -\bar{\Psi}m\Psi.$$

Such a term is not invariant under $SU(2)$ transformations, because, in general,

$$\mathscr{L}_{\text{D1+D2}} = -\bar{\Psi}'m\Psi' = \bar{\Psi}\underbrace{\mathrm{e}^{-ia_i \frac{\sigma_i}{2}} m \mathrm{e}^{ia_i \frac{\sigma_i}{2}}}_{\neq m}\Psi.$$

For equal masses $m_1 = m_2$ it would be invariant, but we are going to see how we can include arbitrary mass terms without violating this symmetry. We know from experiments that the two fields in the doublet do not create particles of equal mass, i.e. $m_1 \neq m_2$. This will be discussed later in detail.

This symmetry should be a *local* symmetry, too. The $SU(2)$ transformations mix the two components of the doublet. Later we will give these two fields names like electron and electron-neutrino field. Our symmetry here tells us that it does not matter what we call electron and what electron-neutrino field, because by using $SU(2)$ transformations we can mix them as we like. If this is only a global symmetry, as soon as we fix one choice[28], which means we decide what we call electron and what electron-neutrino field, this choice would be fixed immediately for the complete universe. Therefore we investigate if this is a local symmetry. Again we find that it isn't, but as for local $U(1)$ symmetry, we will do everything we can to make the Lagrangian locally $SU(2)$ invariant.

[28] This is known as **choosing a gauge**.

The problem here is again the derivative, which produces an extra term. To unclutter the notation, we define $U(x) \equiv e^{-ia_i(x)\frac{\sigma_i}{2}}$:

$$\Psi \to \Psi' = U(x)\Psi \tag{7.48}$$

$$\Rightarrow \bar{\Psi} \to \bar{\Psi}' = \bar{\Psi}U^\dagger(x). \tag{7.49}$$

Our transformed Lagrangian then reads

$$\begin{aligned} \mathscr{L}'_{\text{D1+D2}} &= i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' \\ &= i\bar{\Psi}U^\dagger(x)\gamma_\mu\partial^\mu(U(x)\Psi) \\ &\underbrace{=}_{\text{product rule}} i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + i\bar{\Psi}\gamma_\mu U^\dagger(x)(\partial_\mu U(x))\Psi \neq \mathscr{L}_{\text{D1+D2}}. \end{aligned} \tag{7.50}$$

We can see that the Lagrangian is not invariant because we get an additional term, $i\bar{\Psi}\gamma_\mu U^\dagger(x)(\partial_\mu U(x))\Psi$, after the transformation.

So how can we modify our Lagrangian $\mathscr{L}_{\text{D1+D2}}$ such that is invariant under local $SU(2)$ transformations? In principle the same method that we discovered in the last sections to ensure local $U(1)$ works again. However, some details are a bit more complicated. We already noted in Eq. 7.23 that the essence of how we need to change our Lagrangian can be summarized through the replacement of the ordinary derivative ∂^μ, with a "covariant derivative" $D^\mu = \partial^\mu - igA^\mu$. With this in mind, we replace the derivative ∂^μ in our Lagrangian $\mathscr{L}_{\text{D1+D2}}$ with a new object D^μ and then try to derive how D^μ looks like in order to ensure local $SU(2)$ invariance.

Concretely, we now write

$$\widetilde{\mathscr{L}}_{\text{D1+D2}} = i\bar{\Psi}\gamma_\mu D^\mu\Psi \tag{7.51}$$

and under a local $SU(2)$ transformation this Lagrangian becomes

$$\begin{aligned}\widetilde{\mathscr{L}}'_{\mathrm{D1+D2}} &= i\bar{\Psi}'\gamma_\mu(D^\mu\Psi)' \\ &= i\bar{\Psi}U^\dagger(x)\gamma_\mu(D^\mu\Psi)' \\ &\stackrel{!}{=} \widetilde{\mathscr{L}}_{\mathrm{D1+D2}}. \end{aligned} \tag{7.52}$$

We can see here that this new Lagrangian is invariant under the local $SU(2)$ transformation $U(x) = \mathrm{e}^{-ia_i(x)\frac{\sigma_i}{2}}$ if $(D^\mu\Psi)' = U(x)D^\mu\Psi$, because then[29]

$$\begin{aligned}\widetilde{\mathscr{L}}'_{\mathrm{D1+D2}} &= i\bar{\Psi}'\gamma_\mu(D^\mu\Psi)' \\ &= i\bar{\Psi}U^\dagger(x)\gamma_\mu(D^\mu\Psi)' \\ &= i\bar{\Psi}\gamma_\mu \underbrace{U^\dagger(x)U(x)}_{=1} D^\mu\Psi \\ &= \widetilde{\mathscr{L}}_{\mathrm{D1+D2}} \quad \checkmark. \end{aligned} \tag{7.53}$$

So our goal is to find an object D^μ that has exactly this transformation behavior $(D^\mu\Psi)' = U(x)D^\mu\Psi$. Take note that this property is the reason why we call D^μ the "covariant derivative"[30]. The covariant derivative of Ψ, which we denote by $D^\mu\Psi$, transforms exactly like Ψ, i.e. $\Psi' = U(x)\Psi$. Thus the form stays the same and we don't get additional terms, like we do if we use the ordinary derivative ∂^μ[31].

From our experience with local $U(1)$ symmetry, we already know that the crucial trick is to make use of spin 1 fields. However, there is one crucial difference. We saw in Eq. 7.50 that again the reason for the non-invariance is that the derivative ∂_μ produces an extra term $i\bar{\Psi}\gamma_\mu U^\dagger(x)(\partial_\mu U(x))\Psi$. Our local $SU(2)$ transformations $U(x) \equiv \mathrm{e}^{-ia_i(x)\frac{\sigma_i}{2}}$ are a bit more complicated than local $U(1)$ transformations $U(x) \equiv \mathrm{e}^{-i\alpha(x)}$, because of the generators $\sigma_i/2$ in the exponent. For local $SU(2)$ transformations, we have

$$\partial_\mu U(x) = \partial_\mu \mathrm{e}^{-ia_i(x)\frac{\sigma_i}{2}} = -i(\partial_\mu a_i(x))\frac{\sigma_i}{2}\mathrm{e}^{-ia_i(x)\frac{\sigma_i}{2}}. \tag{7.54}$$

We can see here that we actually get three extra terms[32], one for each Pauli matrix σ_i. Thus in contrast to the $U(1)$ case, to make the Lagrangian locally $SU(2)$ invariant we don't need one additional spin 1 field, but three! In addition, we can see that the troublesome terms, that are produced through the derivative in the Lagrangian ∂_μ, are proportional to the generators $\frac{\sigma_i}{2}$.

To summarize: we need three spin 1 fields to cancel the terms that make our Lagrangian non-invariant under local $SU(2)$ transformations and must introduce these new fields in such a way that they are able to cancel terms that involve the generators $\frac{\sigma_i}{2}$.

[29] The reason that we can change the positions of $U^\dagger(x)$ and γ_μ, as already discussed in the text below Eq. 7.47, is that γ_μ is a matrix that acts on the components of the Dirac spinors, whereas the $SU(2)$ transformation mixes the two Dirac spinors ψ_1 and ψ_2 inside the doublet Ψ.

[30] The notion "covariance" was discussed in Section 2.6 and means roughly that the form of an equation or an object is not changed under a given transformation.

[31] This was demonstrated in Eq. 7.50.

[32] We have here an implicit sum, in the sense of Einstein's summation convention, over the index i.

We therefore try

$$D^\mu = \partial^\mu - ig\frac{\sigma^a}{2}W_a^\mu \,, \tag{7.55}$$

where σ^a are the Pauli matrices and $a = 1,2,3$. With this ansatz, the requirement $(D^\mu\Psi)' \stackrel{!}{=} U(x)D^\mu\Psi$, can now be translated into a transformation law for the spin 1 fields W_a^μ:

$$\begin{aligned}
(D^\mu\Psi)' &\stackrel{!}{=} U(x)D^\mu\Psi \\
\underbrace{\rightarrow}_{\text{Eq. 7.55}} ((\partial^\mu - ig\frac{\sigma^a}{2}W_a^\mu)\Psi)' &\stackrel{!}{=} U(x)(\partial^\mu - ig\frac{\sigma^a}{2}W_a^\mu)\Psi \\
\rightarrow \partial^\mu\Psi' - ig\frac{\sigma^a}{2}W_a'^\mu\Psi' &\stackrel{!}{=} U(x)(\partial^\mu - ig\frac{\sigma^a}{2}W_a^\mu)\Psi \\
\underbrace{\rightarrow}_{\text{Eq. 7.48}} \partial^\mu(U(x)\Psi) - ig\frac{\sigma^a}{2}W_a'^\mu U(x)\Psi &\stackrel{!}{=} U(x)(\partial^\mu - ig\frac{\sigma^a}{2}W_a^\mu)\Psi \\
\underbrace{\rightarrow}_{\text{product rule}} (\partial^\mu U(x))\Psi + \cancel{U(x)(\partial^\mu\Psi)} - ig\frac{\sigma^a}{2}W_a'^\mu U(x)\Psi &\stackrel{!}{=} \cancel{U(x)\partial^\mu\Psi} - igU(x)\frac{\sigma^a}{2}W_a^\mu\Psi \\
\rightarrow (\partial^\mu U(x))\Psi - ig\frac{\sigma^a}{2}W_a'^\mu U(x)\Psi &\stackrel{!}{=} -igU(x)\frac{\sigma^a}{2}W_a^\mu\Psi \,.
\end{aligned} \tag{7.56}$$

The Lagrangian should be invariant for arbitrary Ψ and therefore, we write the last line in Eq. 7.56 without it

$$(\partial^\mu U(x)) - ig\frac{\sigma^a}{2}W_a'^\mu U(x) \stackrel{!}{=} -igU(x)\frac{\sigma^a}{2}W_a^\mu \,. \tag{7.57}$$

We can now calculate the correct transformation behavior of the spin 1 fields $W_a'^\mu$ by "solving" this equation for $W_a'^\mu$. To achieve this we multiply Eq. 7.57 from the right with $U^{-1}(x)$, which yields

$$\begin{aligned}
&(\partial^\mu U(x))U^{-1}(x) - ig\frac{\sigma^a}{2}W_a'^\mu \underbrace{U(x)U^{-1}(x)}_{=1} \stackrel{!}{=} -igU(x)\frac{\sigma^a}{2}W_a^\mu U(x)^{-1} \\
&\rightarrow \frac{\sigma^a}{2}W_a'^\mu \stackrel{!}{=} U(x)\frac{\sigma^a}{2}W_a^\mu U^{-1}(x) - \underbrace{\frac{1}{ig}}_{=\frac{i}{i^2 g}=-\frac{i}{g}}(\partial^\mu U(x))U^{-1}(x) \\
&\rightarrow \frac{\sigma^a}{2}W_a'^\mu \stackrel{!}{=} U(x)\frac{\sigma^a}{2}W_a^\mu U^{-1}(x) + \frac{i}{g}(\partial^\mu U(x))U^{-1}(x)
\end{aligned} \tag{7.58}$$

This is how our gauge fields $W_a'^\mu$, that we introduced in Eq. 7.55 as part of the covariant derivative $D^\mu = \partial^\mu - ig\frac{\sigma^a}{2}W_a^\mu$, need to transform such that $(D^\mu\Psi)' = U(x)D^\mu\Psi$. We saw in Eq. 7.53 that this transformation behavior is needed to get a locally $SU(2)$ invariant Lagrangian.

There is one last thing, we need to take care of. We introduced three new spin 1 fields W_a^μ and we saw that they need to have very specific transformation properties when we want a locally $SU(2)$ invariant Lagrangian. However, the "naive" Lagrangian for these spin 1 fields (Eq. 6.25)

$$\begin{aligned}\mathscr{L}_{3\text{xMaxwell}} &= \frac{1}{4}(W_{\mu\nu})_1(W^{\mu\nu})_1 + \frac{1}{4}(W_{\mu\nu})_2(W^{\mu\nu})_2 + \frac{1}{4}(W_{\mu\nu})_3(W^{\mu\nu})_3 \\ &= \frac{1}{4}(W_{\mu\nu})_i(W^{\mu\nu})_i \qquad (7.59)\end{aligned}$$

with

$$(W_{\mu\nu})_i = \partial_\mu(W_\nu)_i - \partial_\nu(W_\mu)_i$$

is not invariant under such transformations[33].

[33] This can be seen by using the explicit transformation behavior that we derived in Eq. 7.58.

We saw above that the demand for local $U(1)$ symmetry was powerful and yielded the correct Lagrangian that describes electromagnetic interactions. Thus instead of discarding the transformation behavior that successfully makes the spin $\frac{1}{2}$ part of of the Lagrangian invariant under local $SU(2)$ transformations, because it is not a symmetry of the free spin 1 field Lagrangian, we try to find a better Lagrangian for these spin 1 fields which has the desired symmetry. In other words, our final task is to derive a Lagrangian that describes how these new spin 1 fields behave when they are on their own that is invariant under the transformation in Eq. 7.58.

To find this Lagrangian, we need to note several things:

1. The new spin 1 fields $(W_\mu)_i$ always appeared in the previous Lagrangians[34] in combination with the generators $\frac{\sigma^i}{2}$. It is thus useful to introduce a new object $\mathcal{W}_\mu \equiv (W_\mu)_i\frac{\sigma^i}{2}$. Next, we might try to use this new object $\mathcal{W}_\mu$ instead of $(W_\mu)_i$ in the Lagrangian Eq. 7.59[35]. The important difference is that $\mathcal{W}_\mu$ is a matrix, because the generators are matrices. Therefore, when we define our field strength tensor in terms of the new object $\mathcal{W}_\mu$:

$$\mathcal{W}_{\mu\nu} = \partial_\mu \mathcal{W}_\nu - \partial_\nu \mathcal{W}_\mu \, , \qquad (7.60)$$

it becomes a matrix, too.

[34] See, e.g. Eq. 7.58.

[35] It will become clear in a moment, why this is useful, although at a first glance it seems to make things much more complicated.

2. We want a Lagrangian that is invariant under local $SU(2)$ transformations, and therefore need to combine our fields in such a way that their transformation behavior cancels exactly. Now that we have new objects $(\mathcal{W}_\mu, \mathcal{W}_{\mu\nu})$ that are matrices, we can try to make use of the following nice property of the trace of matrices[36]

$$\text{Tr}(ABCD) = \text{Tr}(DABC) = \text{Tr}(CDAB) = \text{Tr}(BCDA) \, . \qquad (7.61)$$

[36] This property is known as "cyclic property" of the trace. In words it means that the trace stays the same when we perform cyclic permutations among the matrices A, B, C, D that appear here. We always take the last element that appears in the product and put it at the beginning. However, take note that arbitrary permutations do not lead to the same trace. For example, $\text{Tr}(ACBD) \neq \text{Tr}(ABCD)$.

If we somehow manage that our field strength tensor transforms as $\mathcal{W}_{\mu\nu} \rightarrow U(x)\mathcal{W}_{\mu\nu}U^{-1}(x)$, the term $\text{Tr}(\mathcal{W}_{\mu\nu}\mathcal{W}_{\mu\nu})$ would be invariant, because[37]

[37] To get to the second line, we use the cyclic property of the trace.

$$\begin{aligned}\text{Tr}\big(\mathcal{W}_{\mu\nu}\mathcal{W}_{\mu\nu}\big) \rightarrow &\text{Tr}\big(U(x)\mathcal{W}_{\mu\nu}U^{-1}(x)U(x)\mathcal{W}_{\mu\nu}U^{-1}(x)\big)\\ &\underbrace{=}_{\text{Eq. 7.61}} \text{Tr}\big(\underbrace{U^{-1}(x)U(x)}_{=1}\mathcal{W}_{\mu\nu}\underbrace{U^{-1}(x)U(x)}_{=1}\mathcal{W}_{\mu\nu}\big)\\ &= \text{Tr}\big(\mathcal{W}_{\mu\nu}\mathcal{W}_{\mu\nu}\big) \quad \checkmark .\end{aligned} \tag{7.62}$$

3. Unfortunately, the naive field strength tensor that we wrote down in Eq. 7.60 does not transform so nicely. However, we can construct a different field strength tensor that has exactly the needed transformation behavior. Above, we derived the covariant derivative D^μ that transforms exactly in the way we need it to get a locally $SU(2)$ invariant Lagrangian for the spin $\frac{1}{2}$ fields: $(D^\mu\Psi)' = U(x)D^\mu\Psi$. The spin $\frac{1}{2}$ doublet Ψ transforms as[38] $\Psi \rightarrow U(x)\Psi$. Therefore, we can conclude that the covariant derivative transforms as $D^\mu \rightarrow U(x)D^\mu U^{-1}(x)$, because

[38] See Eq. 7.48.

$$\begin{aligned}(D^\mu\Psi)' &= D'^\mu\Psi'\\ &= U(x)D^\mu\underbrace{U^{-1}(x)U(x)}_{=1}\Psi\\ &= U(x)D^\mu\Psi \quad \checkmark\end{aligned} \tag{7.63}$$

Now the final crucial trick to get a locally $SU(2)$ invariant Lagrangian for the spin 1 fields is the observation that the object[39]

[39] At this point the way we introduce this correct field strength tensor may seem a bit like magic. However, of course, there is a deep reason why the correct field strength tensor must be the commutator of the covariant derivative. Unfortunately, a proper discussion lies beyond the scope of this text. It will become clear in the a moment, why we included the additional factor $\frac{i}{g}$ here.

$$\mathcal{W}_{\mu\nu} \equiv \frac{i}{g}[D^\mu, D^\nu] = \frac{i}{g}\big(D^\mu D^\nu - D^\nu D^\mu\big) \tag{7.64}$$

has exactly the correct transformation behavior[40] $\mathcal{W}_{\mu\nu} \rightarrow U(x)\mathcal{W}_{\mu\nu}U^{-1}(x)$:

$$\begin{aligned}\frac{g}{i}\mathcal{W}_{\mu\nu} &= D^\mu D^\nu - D^\nu D^\mu\\ &\rightarrow U(x)D^\mu\underbrace{U^{-1}(x)U(x)}_{=1}D^\nu U^{-1} - U(x)D^\nu\underbrace{U^{-1}(x)U(x)}_{=1}D^\mu U^{-1}\\ &= U(x)D^\mu D^\nu U^{-1} - U(x)D^\nu D^\mu U^{-1}\\ &= U(x)\big(D^\mu D^\nu - D^\nu D^\mu\big)U^{-1}\\ &= U(x)\mathcal{W}_{\mu\nu}U^{-1} \quad \checkmark\end{aligned} \tag{7.65}$$

[40] Maybe you wonder why we use $D^\mu D^\nu - D^\nu D^\mu$ and not just $D^\mu D^\nu$. In short: we need to include all terms that are allowed by our restrictions in the Lagrangian. For the spin 1 fields this means that we not only need $\partial_\mu \mathcal{W}_\nu$ but also $\partial_\nu \mathcal{W}_\mu$. To correctly account for both these possibilities, we need to use not only $D^\mu D^\nu$, but also $D^\nu D^\mu$. This will be made explicit below.

4. When we now calculate explicitly how our field strength tensor looks like in terms of the fields $\mathcal{W}_\mu$, we can see what we missed in our naive definition in Eq. 7.60[41]

[41] We include here an arbitrary test function $f(x)$, because our derivatives must act on something.

$$
\begin{aligned}
\mathcal{W}_{\mu\nu} f(x) &= \frac{i}{g}\big(D^\mu D^\nu - D^\nu D^\mu\big) f(x) \\
&= \left(\frac{i}{g}(\partial^\mu - ig\mathcal{W}^\mu)(\partial^\nu - ig\mathcal{W}^\nu) - \frac{i}{g}(\partial^\nu - ig\mathcal{W}^\nu)(\partial^\mu - ig\mathcal{W}^\mu)\right) f(x) \\
&= \Big(\cancel{\frac{i}{g}\partial^\mu\partial^\nu} + \partial^\mu \mathcal{W}^\nu + \mathcal{W}^\mu \partial^\nu - ig\mathcal{W}^\mu\mathcal{W}^\nu \\
&\quad - (\cancel{\frac{i}{g}\partial^\nu\partial^\mu} + \partial^\nu \mathcal{W}^\mu + \mathcal{W}^\nu \partial^\mu - ig\mathcal{W}^\nu\mathcal{W}^\mu)\Big) f(x) \\
&\underbrace{=}_{\text{product rule}} \Big(\partial^\mu \mathcal{W}^\nu + \cancel{\mathcal{W}^\nu\partial^\mu} + \cancel{\mathcal{W}^\mu\partial^\nu} - ig\mathcal{W}^\mu\mathcal{W}^\nu \\
&\quad - (\partial^\nu \mathcal{W}^\mu + \cancel{\mathcal{W}^\mu\partial^\nu} + \cancel{\mathcal{W}^\nu\partial^\mu} - ig\mathcal{W}^\nu\mathcal{W}^\mu)\Big) f(x) \\
&= \Big(\partial^\mu \mathcal{W}^\nu - \partial^\nu \mathcal{W}^\mu - ig\mathcal{W}^\mu\mathcal{W}^\nu + ig\mathcal{W}^\nu\mathcal{W}^\mu)\Big) f(x) \\
&= \big(\partial^\mu \mathcal{W}^\nu - \partial^\nu \mathcal{W}^\mu - ig[\mathcal{W}^\mu, \mathcal{W}^\nu]\big) f(x)\,. \qquad (7.66)
\end{aligned}
$$

The last term here is what we missed previously[42]. It is important for local $SU(2)$ symmetry, because the $\mathcal{W}^\mu$ are now matrices and therefore $\mathcal{W}^\mu\mathcal{W}^\nu \neq \mathcal{W}^\nu\mathcal{W}^\mu$. In contrast, in order to ensure local $U(1)$ symmetry there was no need for such a term. There was just one spin 1 field, hence no matrix and therefore this last term vanishes.

[42] In addition, we can see here, why we included the factor $\frac{i}{g}$ in the definition of $\mathcal{W}_{\mu\nu}$. this factor is exactly what we need to cancel the factors that come from the definition of D^μ.

This was quite a long journey, but now we have everything we need to write down a Lagrangian that is invariant under local $SU(2)$ transformations:

$$
\mathscr{L}_{\text{locally SU(2) invariant}} = i\bar{\Psi}\gamma_\mu D^\mu \Psi - \frac{1}{4} Tr\big(\mathcal{W}_{\mu\nu}\mathcal{W}^{\mu\nu}\big)\,, \qquad (7.67)
$$

where

$$
\begin{aligned}
D^\mu &\equiv \partial^\mu - ig\mathcal{W}^\mu \\
\mathcal{W}^{\mu\nu} &\equiv \partial^\mu \mathcal{W}^\nu - \partial^\nu \mathcal{W}^\mu - ig[\mathcal{W}^\mu, \mathcal{W}^\nu] \\
\mathcal{W}^\mu &\equiv (W^\mu)_i \frac{\sigma^i}{2}\,. \qquad (7.68)
\end{aligned}
$$

7.3 Mass Terms and "Unification" of $SU(2)$ and $U(1)$

In the last section we couldn't add mass terms like $m_1\bar{\Psi}\Psi$ and $m_2(W^\mu)_i(W_\mu)_i$ to the Lagrangian without destroying the $SU(2)$ symmetry. From experiments we know that the corresponding particles[43] have mass and this is conventionally interpreted as the $SU(2)$ symmetry being

[43] For example, the electron e^- and the electron-neutrino ν_e, described by Ψ and the three bosons, described by $(W^\mu)_i$.

broken. This means the symmetry exists at high energy and spontaneously breaks at lower energies.

So far we derived a locally $U(1)$ invariant Lagrangian and in this context it's conventional to name the corresponding spin 1 field B_μ:

$$\mathscr{L}_{\text{locally U(1) invariant}} = -m\bar{\psi}\psi + i\bar{\psi}\gamma_\mu(\partial^\mu - igB_\mu)\psi - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \quad (7.69)$$

with

$$B^{\mu\nu} := \partial^\mu B^\nu - \partial^\nu B^\mu$$

The spin 1 field B^μ is often called $U(1)$ **gauge field**, because it makes the Lagrangian $U(1)$ invariant.

The locally $SU(2)$ invariant Lagrangian is (Eq. 7.67)

$$\mathscr{L}_{\text{locally SU(2) invariant}} = i\bar{\Psi}\gamma_\mu D^\mu \Psi - \frac{1}{4}Tr\big(\mathcal{W}_{\mu\nu}\mathcal{W}^{\mu\nu}\big)\,. \quad (7.70)$$

As above, the three spin 1 fields $(W_\nu)_i$ are called the $SU(2)$ gauge fields, because they make the Lagrangian locally $SU(2)$ invariant.

We can combine them into one locally $U(1)$ **and** locally $SU(2)$ invariant Lagrangian[44]

$$\mathscr{L}_{\text{SU(2) and U(1)}} = i\bar{\Psi}\gamma_\mu(\partial^\mu - igB^\mu - ig'\mathcal{W}^\mu)\Psi - \frac{1}{4}Tr\big(\mathcal{W}_{\mu\nu}\mathcal{W}^{\mu\nu}\big) - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \quad (7.71)$$

[44] We use here again the notation $\mathcal{W}^\mu = \frac{\sigma_j}{2}W_j^\mu$.

Now, how can we add mass terms to this Lagrangian without spoiling the $SU(2)$ symmetry? The only ingredient we haven't used so far is a spin 0 field; so let's see. The globally $U(1)$ invariant Lagrangian for a complex spin 0 field is given by (Eq. 7.40)

$$\mathscr{L}_{\text{spin0}} = \frac{1}{2}\Big(\partial_\mu\phi^\dagger\partial^\mu\phi - m^2\phi^\dagger\phi\Big). \quad (7.72)$$

We can add to this Lagrangian the next higher power in ϕ without violating any symmetry constraints[45]. Thus we write, renaming the constants to their conventional names

[45] Recall that only higher order derivatives were really forbidden in order to get a sensible theory. Higher powers of ϕ describe the self-interaction of the field ϕ and were omitted in order to get a "free" theory.

$$\mathscr{L}_{\text{spin0+extraTerm}} = \partial_\mu\phi^\dagger\partial^\mu\phi + \rho^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2. \quad (7.73)$$

We already know from Eq. 7.41 how we can add a coupling term between this spin 0 field and a $U(1)$ gauge field B_μ, which makes the Lagrangian locally $U(1)$ invariant:

$$\mathscr{L}_{\text{spin0+extraTerm+spin1Coupling}} = \Big((\partial_\mu + igB_\mu)\phi^\dagger\Big)\Big((\partial^\mu - igB^\mu)\phi\Big) + \rho^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \quad (7.74)$$

with the symmetries[46]

[46] See Eq. 7.13 and recall that an overall constant in the Lagrangian has no influence.

$$B_\mu \to B'_\mu = B_\mu + \partial_\mu a(x) \tag{7.75}$$
$$\phi(x) \to \phi'(x) = e^{ia(x)}\phi(x). \tag{7.76}$$

In the same way how we derived in the last chapter the locally $SU(2)$ invariant Lagrangian for spin $\frac{1}{2}$ fields, we can write a locally $SU(2)$ invariant Lagrangian for doublets of spin 0 fields[47] as

[47] $\Phi = \begin{pmatrix}\phi_1 \\ \phi_2\end{pmatrix}$

$$\mathscr{L}_{\text{SU(2) and U(1)}} = \Big((\partial_\mu + ig'\mathcal{W}^\mu + igB_\mu)\Phi^\dagger\Big)\Big((\partial^\mu - ig'\mathcal{W}^\mu - igB^\mu)\Phi\Big) + \underbrace{\rho^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2}_{\equiv -V(\Phi)} \tag{7.77}$$

with the doublet $\Phi := \begin{pmatrix}\phi_1 \\ \phi_2\end{pmatrix}$ and the symmetries[48] (Eq. 7.58)

[48] We use again the abbreviation $U(x) = e^{ib_i(x)\frac{\sigma_i}{2}}$ for our local $SU(2)$ transformations.

$$\mathcal{W}^\mu \to \mathcal{W}^\mu = U(x)\mathcal{W}^\mu U^{-1}(x) + \frac{i}{g}(\partial^\mu U(x))U^{-1}(x) \tag{7.78}$$

and

$$\Phi \to \Phi' = U(x)\Phi\,, \quad \overline{\Phi} \to \overline{\Phi}' = \overline{\Phi}U^\dagger(x)\,. \tag{7.79}$$

We start with this locally $SU(2)$ invariant Lagrangian and investigate in the following how this Lagrangian gives us mass terms for the fields $\mathcal{W}^\mu = W_i^\mu \frac{\sigma_i}{2}$ and B^μ. Adding mass terms "by hand" to the Lagrangian does not work, because these terms spoil the symmetry and this leads to an insensible theory[49].

[49] The reason is quite complicated and will not be discussed in this book. In technical terms: We need a locally $SU(2)$ symmetric Lagrangian to get a renormalizable theory. You are encouraged to read about this in the books mentioned at the end of this chapter.

The term we defined above

$$\begin{aligned} V(\Phi) &= -\rho^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2 \\ &= -\rho^2\phi_1^\dagger\phi_1 + \lambda(\phi_1^\dagger\phi_1)^2 - \rho^2\phi_2^\dagger\phi_2 + \lambda(\phi_2^\dagger\phi_2)^2 \\ &= V_1(\phi_1) + V_2(\phi_2) \end{aligned} \tag{7.80}$$

is often called Higgs potential. A two-dimensional plot for different values of ρ, with $\lambda > 0$ can be seen in Fig. 7.1.

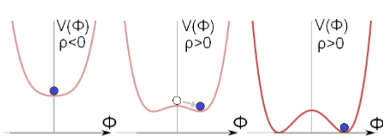

Fig. 7.1: Two-dimensional illustration of the Higgs potential for different values of ρ, which is believed to have changed as the universe cooled down as a result of the expansion of the universe. Figure adapted from "Spontaneous symmetry breaking" by FT2 (Wikimedia Commons) released under a CC BY-SA 3.0 licence: `http://creativecommons.org/licenses/by-sa/3.0/deed.en`. URL: `http://commons.wikimedia.org/wiki/File:Spontaneous_symmetry_breaking_(explanatory_diagram).png`, Accessed: 8.12.2014

The idea is that at very high temperatures, e.g. in the early universe, the potential looks like in the image to the left. The minimum, in this context called the **vacuum expectation value**, is without ambiguity at $\phi = 0$. With sinking temperature the parameters λ and ρ change, and with them the shape of the potential. After the temperature dropped below some critical value the potential no longer has its minimum at $\phi = 0$, as indicated in the pictures to the right. Now there is not only one location with the minimum value, but many.

In fact, the potential has an infinite number of possible minima. The minima of the potential can be computed in the usual way

$$V(\phi) = -\rho^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)(\phi^\dagger\phi) \tag{7.81}$$

$$\frac{\partial V(\phi)}{\partial \phi} = -2\rho^2\phi^\dagger + 2\lambda|\phi|^2\phi^\dagger \stackrel{!}{=} 0 \tag{7.82}$$

$$\to \phi^\dagger(-2\rho^2 + 4\lambda|\phi|^2) \stackrel{!}{=} 0 \tag{7.83}$$

$$\to |\phi|^2 \stackrel{!}{=} \frac{\rho^2}{2\lambda} \tag{7.84}$$

$$\to |\phi| \stackrel{!}{=} \sqrt{\frac{\rho^2}{2\lambda}} \tag{7.85}$$

$$\phi_{\text{min}} = \sqrt{\frac{\rho^2}{2\lambda}} e^{i\varphi}. \tag{7.86}$$

This is a minimum for every value of φ and we therefore have an infinite number of minima. All these minima lie on a circle with radius $\sqrt{\frac{\rho^2}{2\lambda}}$. This can be seen in the three dimensional plot of the Higgs potential in Fig. 7.2. Like a marble that rolls down from the top of a sombrero, spontaneously one new vacuum value is chosen out of the infinite possibilities.

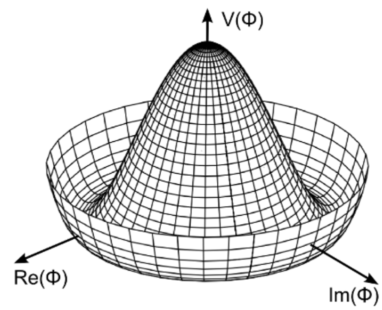

Fig. 7.2: 3-dimensional plot of the Higgs potential. Figure adapted from "Mexican hat potential polar" by Rupert Millard (Wikimedia Commons) released under a public domain licence. URL: `http://commons.wikimedia.org/wiki/File:Mexican_hat_potential_polar.svg` , Accessed: 7.5.2014

From Eq. 7.80 we see that for the doublet, both components have this choice to make. We therefore have for the doublet the minimum

$$\Phi_{\text{min}} = \begin{pmatrix} \phi_{1\text{min}} \\ \phi_{2\text{min}} \end{pmatrix}. \tag{7.87}$$

An economical **choice**[50] for the minimum is

$$\Phi_{\text{min}} = \begin{pmatrix} 0 \\ \sqrt{\frac{\rho^2}{2\lambda}} \end{pmatrix} \equiv \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}, \tag{7.88}$$

[50] Recall that symmetry breaking means that **one** minimum is chosen out of the infinite possibilities.

where the factor $\frac{1}{2}$ is just a convention to make computations easier and we define for brevity $v \equiv \sqrt{\frac{\rho^2}{\lambda}}$. We will learn later that in quantum field theory computations are always done as a series expansion around the minimum, because no exact solutions are available. In order to get sensible results, we must therefore shift the field Φ to the new minimum. We therefore consider the field

$$\Phi = \begin{pmatrix} \phi_{1r} + i\phi_{1c} \\ \frac{v}{\sqrt{2}} + \phi_{2r} + i\phi_{2c} \end{pmatrix}. \tag{7.89}$$

This can be rewritten as[51]

[51] This form is very useful as we will see in a moment.

$$\Phi = e^{i\theta_i \frac{\sigma_i}{2}} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}, \tag{7.90}$$

because if we consider the series expansion of the exponential function and the explicit form of the Pauli matrices σ_i (Eq. 3.80), we can see that in first order

$$\begin{aligned} e^{i\theta_i \frac{\sigma_i}{2}} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} &\approx (1 + i\frac{1}{2}\theta_i\sigma_i) \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \\ &= (1 + i\frac{1}{2}\theta_1\sigma_1 + i\frac{1}{2}\theta_2\sigma_2 + i\frac{1}{2}\theta_3\sigma_3) \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} 1 + i\frac{1}{2}\theta_3 & i\frac{1}{2}\theta_1 + \frac{1}{2}\theta_2 \\ i\frac{1}{2}\theta_1 - \frac{1}{2}\theta_2 & 1 - i\frac{1}{2}\theta_3 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} (i\frac{1}{2}\theta_1 + \frac{1}{2}\theta_2)\frac{v+h}{\sqrt{2}} \\ (1 - i\frac{1}{2}\theta_3)\frac{v+h}{\sqrt{2}} \end{pmatrix} \\ \text{redefinitions} \rightarrow &\equiv \begin{pmatrix} \phi_{1r} + i\phi_{1c} \\ \frac{v}{\sqrt{2}} + \phi_{2r} + i\phi_{2c} \end{pmatrix}. \end{aligned} \tag{7.91}$$

Writing the complex spin 0 doublet in this form is useful, because we can now use the local $SU(2)$ (gauge) symmetry to make computations simpler. In order to get physical results **one** gauge must be chosen and we prefer to work with a gauge that makes life the easiest[52].

A general local $SU(2)$ transformation is

$$\Phi \rightarrow \Phi' = e^{ib_i(x)\frac{\sigma_i}{2}}\Phi, \tag{7.92}$$

which enables us to eliminate the exponential factor in Eq. 7.90, by choosing appropriate $b_i(x)$. The complex scalar doublet is then, in this **unitary gauge**

$$\Phi_{un} = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}. \tag{7.93}$$

Another possible way to understand this is that of the original four components that appeared in our complex scalar doublet, three are equivalent to our $SU(2)$ gauge freedom[53]. Therefore, these three fields aren't physical[54] and can't be measured in experiments. What remains is **one** physical field h, which is called the **Higgs field**.

Next, we want to take a look at the implications of this symmetry breaking on the Lagrangian. We recite here the Lagrangian in question for convenience, which was derived in Eq. 7.74 and include an additional factor $\frac{1}{2}$ in front of the field B_μ to unclutter the notation in the calculations that follow[55]

[52] For different computations, different gauges can be useful. Here we will work with what is called the unitary gauge, that is particularly useful to understand the physical particle content of a theory.

[53] Take note that this is only possible, because we have a local $SU(2)$ theory, because our fields $\theta = \theta(x)$, of course depend on the location in spacetime. For a global symmetry, these components can't be gauged away, and are commonly interpreted as massless bosons, called Goldstone bosons.

[54] The local $SU(2)$ symmetry is nothing that can be measured in experiments. This is merely a symmetry of our equations and the gauge freedom disappears from everything that is measurable in experiments. Otherwise there would be no possible way to make predictions from our theory, because we would have an infinite number of equivalently possible predictions (that are connected by $SU(2)$ transformations). Nevertheless, this symmetry is far from being useless, because it guides us to the correct form of the Lagrangian.

[55] Such a factor changes nothing, because it corresponds simply to a redefinition of the coupling constant $\frac{1}{2}\tilde{g} \equiv g$. Thus, strictly speaking, we should use $\tilde{g}$ here instead of g. However, the name of the constant doesn't matter and we continue to call the coupling constant simply g.

$$\mathscr{L} = \Big((\partial_\mu + ig'\frac{\sigma_i}{2}(W_\mu)_i + i\frac{1}{2}gB_\mu)\Phi^\dagger\Big)\Big((\partial^\mu - ig'\frac{\sigma_i}{2}(W^\mu)_i - i\frac{1}{2}gB^\mu)\Phi\Big) - V(\Phi)\,. \tag{7.94}$$

We now substitute the field Φ with the shifted field in the unitary gauge, which was defined in Eq. 7.93. Of particular interest for us will be the newly appearing terms that include the constant vacuum value v. The other terms describe the self-interaction of the Higgs-field and the interaction of the Higgs field with the other fields, which we will not examine any further. If we put in the minimum value $\Phi \to \Phi_{\text{min}} = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}$, which means we ignore h, we get

$$\begin{aligned}
\Big((\partial_\mu + i\frac{\sigma_i}{2}g'(W_\mu)_i + i\frac{1}{2}gB_\mu)\Phi^\dagger_{\text{min}}\Big)&\Big((\partial^\mu - ig'\frac{\sigma_i}{2}(W^\mu)_i - i\frac{1}{2}gB^\mu)\Phi_{\text{min}}\Big) \\
&= \Big|\Big((\partial^\mu + ig'\frac{\sigma_i}{2}(W^\mu)_i + i\frac{1}{2}gB^\mu)\Phi_{\text{min}}\Big)\Big|^2 \\
&= \Big|\Big((\partial^\mu + ig'\frac{\sigma_i}{2}(W^\mu)_i + i\frac{1}{2}gB^\mu)\sqrt{\frac{1}{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}\Big)\Big|^2 \\
&= \frac{v^2}{8}\Big|\Big((g'\sigma_i(W^\mu)_i + gB^\mu)\begin{pmatrix} 0 \\ 1 \end{pmatrix}\Big)\Big|^2.
\end{aligned}$$

Now using that we have behind B^μ an implicit 2×2 identity matrix and the explicit form of the Pauli matrices[56] σ_i yields

[56] $\sigma_i W_i = \begin{pmatrix} W_3 & W_1 - iW_2 \\ W_1 + iW_2 & -W_3 \end{pmatrix}$

$$\begin{aligned}
&= \frac{v^2}{8}\Big|\begin{pmatrix} g'W_3^\mu + gB^\mu & g'W_1^\mu - ig'W_2^\mu \\ g'W_1^\mu + ig'W_2^\mu & -g'W_3^\mu + gB^\mu \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\Big|^2 \\
&= \frac{v^2}{8}\Big|\Big(\begin{pmatrix} g'W_1^\mu - ig'W_2^\mu \\ -g'W_3^\mu + gB^\mu \end{pmatrix}\Big)\Big|^2 \\
&= \frac{v^2}{8}\Big((g')^2\Big((W_1^\mu)^2 + (W_2^\mu)^2\Big) + (g'W_3^\mu - gB^\mu)^2\Big)
\end{aligned} \tag{7.95}$$

Next we define two new spin 1 fields from the old ones we have been using so far

$$W_+^\mu \equiv \frac{1}{\sqrt{2}}(W_1^\mu - iW_2^\mu) \tag{7.96}$$

$$W_-^\mu \equiv \frac{1}{\sqrt{2}}(W_1^\mu + iW_2^\mu), \tag{7.97}$$

where W_+^μ is the complex conjugate of W_-^μ. The first term in Eq. 7.95 is then

$$(W_1^\mu)^2 + (W_2^\mu)^2 = 2(W^+)_\mu(W^-)^\mu \tag{7.98}$$

and thus we have, including the constants,

$$\Big(\underbrace{\frac{g'v}{2}}_{\equiv m_W}\Big)^2 (W^+)_\mu (W^-)^\mu \tag{7.99}$$

which then looks like a typical "mass" term.

The second term in Eq. 7.95 can be written in matrix form

$$(g'W_3^\mu - gB^\mu)^2 = \begin{pmatrix} W_3^\mu, B_\mu \end{pmatrix} \underbrace{\begin{pmatrix} g'^2 & -gg' \\ -gg' & g^2 \end{pmatrix}}_{\equiv G} \begin{pmatrix} W_3^\mu \\ B_\mu \end{pmatrix} . \tag{7.100}$$

In order to be able to interpret this as mass-terms, we need to diagonalize[57] the matrix G. The standard linear-algebra way to do this needs the eigenvalues λ_1, λ_2 and normalized[58] eigenvectors $\vec{v}_1, \vec{v}_2$ of the matrix G, which are

[57] We will see in a moment that a diagonalized matrix gives us terms that look exactly like the other mass terms. This enables us to interpret the corresponding fields as physical fields that can be observed in experiments. We could work with fields W_3^μ and B_μ, but the physical interpretation would be much harder.

[58] Which means length 1, i.e. $\vec{v} \cdot \vec{v} = 1$.

$$\lambda_1 = 0 \rightarrow \vec{v}_1 = \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g \\ g' \end{pmatrix}$$

$$\lambda_2 = (g^2 + g'^2) \rightarrow \vec{v}_2 = \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g' \\ -g \end{pmatrix} .$$

The matrix G is then diagonalized by the matrix M build from the eigenvectors as its columns, i.e. $G_{diag} = M^{-1}GM$, with

$$M = \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g & g' \\ g' & -g \end{pmatrix} \tag{7.101}$$

and

$$G_{diag} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & (g^2 + g'^2) \end{pmatrix} . \tag{7.102}$$

The matrix M is orthogonal ($M^T = M^{-1}$), because we work with normalized eigenvectors:

$$M^T M = \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g & g' \\ g' & -g \end{pmatrix} \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g & g' \\ g' & -g \end{pmatrix}$$
$$= \frac{1}{(g^2 + g'^2)} \begin{pmatrix} g^2 + g'^2 & gg' - gg' \\ gg' - gg' & g^2 + g'^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{7.103}$$

We therefore add two unit matrices $1 = M^T M$, into Eq. 7.100:

$$\begin{pmatrix} W_3^\mu & , B_\mu \end{pmatrix} \underbrace{MM^T}_{=1} G \underbrace{MM^T}_{=1} \begin{pmatrix} W_3^\mu \\ B_\mu \end{pmatrix} = \begin{pmatrix} W_3^\mu & , B_\mu \end{pmatrix} M \underbrace{M^T G M}_{=G_{diag}} M^T \begin{pmatrix} W_3^\mu \\ B_\mu \end{pmatrix} . \tag{7.104}$$

The remaining task is then to evaluate $M^T \begin{pmatrix} W_3^\mu \\ B_\mu \end{pmatrix}$, in order to get the definition of two new fields, which have easily interpretable mass terms in the Lagrangian:

$$M^T \begin{pmatrix} W_3^\mu \\ B_\mu \end{pmatrix} = \frac{1}{\sqrt{g^2+g'^2}} \begin{pmatrix} g & g' \\ g' & -g \end{pmatrix} \begin{pmatrix} W_3^\mu \\ B_\mu \end{pmatrix}$$
$$= \frac{1}{\sqrt{g^2+g'^2}} \begin{pmatrix} (gW_3^\mu + g'B^\mu) \\ (g'W_3^\mu - gB^\mu) \end{pmatrix} \equiv \begin{pmatrix} A^\mu \\ Z_\mu \end{pmatrix}. \quad (7.105)$$

We can therefore write the second term as

$$\begin{pmatrix} A^\mu & Z_\mu \end{pmatrix} G_{diag} \begin{pmatrix} A^\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} A^\mu & Z_\mu \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & (g^2+g'^2) \end{pmatrix} \begin{pmatrix} A^\mu \\ Z_\mu \end{pmatrix}$$

$$= (g^2+g'^2)(Z^\mu)^2 + 0 \cdot (A^\mu)^2. \quad (7.106)$$

To summarize: We started with a Lagrangian, without mass terms for the spin 1 fields W_i^μ and B^μ (Eq. 7.71)

$$\mathscr{L}_{\text{SU(2) and U(1)}} = i\bar{\Psi}\gamma_\mu(\partial^\mu - igB^\mu - ig'\mathcal{W}^\mu)\Psi - \frac{1}{4}Tr(\mathcal{W}_{\mu\nu}\mathcal{W}^{\mu\nu}) - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \quad (7.107)$$

Then we included interactions with a doublet of spin 0 fields (Eq. 7.77) and after the process of spontaneous symmetry breaking, we have new terms in the Lagrangian that are interpreted as mass terms[59]

$$\underbrace{\frac{1}{4}v^2 g'^2}_{=M_W^2}(W^+)_\mu (W^-)^\mu + \underbrace{\frac{1}{8}v^2(g^2+g'^2)}_{=\frac{1}{2}M_Z^2} Z_\mu^2 + \underbrace{\frac{1}{8}}_{\text{photon mass }=0} v^2 0 \cdot A_\mu^2. \quad (7.108)$$

We can see that one of the spin 1 fields A_μ remains massless after spontaneous symmetry breaking. This is the photon field of electromagnetism and all experiments up to now verify that the photon is massless [60]. An important observation is that the field responsible for Z-bosons Z_μ and the field responsible for photons A_μ are orthogonal linear combinations of the fields B_μ and W_μ^3. Therefore we can see that both have a common origin!

The same formalism can be used to get mass terms for spin $\frac{1}{2}$ fields without spoiling the local $SU(2)$ symmetry, but before we discuss this, we need to talk about one very curious fact of nature: parity violation.

[59] For the first term here, we combine Eq. 7.95 with Eq. 7.98.

[60] Take note that I omitted some very important notions in this section: Hypercharge and the Weinberg angle. The Weinberg angle θ_W is simply defined as $\cos(\theta_W) = \frac{g}{\sqrt{g^2+g'^2}}$ or $\sin(\theta_W) = \frac{g'}{\sqrt{g^2+g'^2}}$. This can be used to simplify some of the definitions mentioned in this section. Hypercharge is a bit more complicated to explain and those who want to dig deeper are referred to the standard texts about quantum field theory, some of which are recommended at the end of Chapter 9.

7.4 Parity Violation

One of the biggest discoveries in the history of science was that nature is not invariant under parity transformations. In layman's terms this means that some experiments behave differently than their mirrored analogue. The experiment that discovered the violation of parity symmetry was the Wu experiment. A full description of this experiment, although fascinating, strays from our current subject, so let's just discuss the final result.

The Wu experiment discovered that the particles mediating the weak force (the W^+, W^-, Z bosons) only couple to left-chiral particles. In other words: **only left-chiral particles interact via the weak force**[61]. All particles produced in weak interactions are left-chiral. Neutrinos interact exclusively via the weak force and therefore it is possible that there are no right-chiral neutrinos[62]. All other particles can be produced via other interactions and therefore can be right-chiral, too.

[61] We will discuss at the end of this section why this means that parity is violated

[62] We will see in a moment that massive, left-chiral particles always get a right-chiral component during propagation. It is known from experiments that neutrinos have mass and therefore we would expect a right-chiral neutrino component. Nevertheless, this right-chiral component does not participate in any known interaction.

Up to this point, we used left-chiral and right-chiral as labels for objects transforming according to different representations of the Lorentz group. Although this seems like something very abstract, we can measure the chirality of particles, because there is a correlation to a more intuitive concept called helicity[63].

[63] We will not discuss this here, because the details make no difference for the purpose of the text. The message to take away is that it can be done. A very nice discussion of these matters can be found in Alessandro Bettini. *Introduction to Elementary Particle Physics*. Cambridge University Press, 2nd edition, 4 2014. ISBN 9781107050402

In fact, most of the time particles do not have a specified **chirality**, which means they aren't definitely left-chiral or right-chiral and the corresponding Dirac spinor Ψ has both components. Parity violation was **no** prediction of the theory and a total surprise for every physicist. Until the present day, no one knows why nature behaves so strangely. Nevertheless, it's easy to accommodate this discovery into our framework. We only need *something* that makes sure we always deal with left-chiral spinors when we describe weak interactions.

Recall that the symbols χ and ξ denote Weyl spinors (two component objects), ψ Dirac spinors (four component objects, consisting of two Weyl spinors)

$$\psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \tag{7.109}$$

and Ψ doublets of Dirac spinors

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} . \tag{7.110}$$

Then the "something" we need is the projection operator P_L:

$$P_L \psi = P_L \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = \begin{pmatrix} \chi_L \\ 0 \end{pmatrix} \equiv \psi_L . \tag{7.111}$$

Such a projection operator can be constructed using the matrix[64]

$$\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{7.112}$$

The matrix γ_5 is called the **chirality operator**, because states of pure chirality $\begin{pmatrix} \chi_L \\ 0 \end{pmatrix}$ or $\begin{pmatrix} 0 \\ \xi_R \end{pmatrix}$ are eigenstates of γ_5 with eigenvalue -1 and $+1$ respectively.

The projection operator P_L can then be written[65]

$$P_L = \frac{1-\gamma_5}{2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{7.113}$$

and we can define analogously

$$P_R = \frac{1+\gamma_5}{2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{7.114}$$

Now, in order to accommodate for the fact that only left-chiral particles interact via the weak force, we must simply include P_L into all terms of the Lagrangian that describe the interaction of $W^{\pm}_\mu$ and Z_μ with different fields. The corresponding terms were derived in Section 7.2, and the final result was Eq. 7.67, which we recite here for convenience:

$$\begin{aligned} \mathscr{L}_{\text{locally SU(2) invariant}} &= i\bar{\Psi}\gamma_\mu D^\mu \Psi - \frac{1}{4} Tr\big(\mathcal{W}_{\mu\nu}\mathcal{W}^{\mu\nu}\big) \\ &= i\bar{\Psi}\gamma_\mu(\partial^\mu - ig\mathcal{W}^\mu)\Psi - \frac{1}{4} Tr\big(\mathcal{W}_{\mu\nu}\mathcal{W}^{\mu\nu}\big). \end{aligned} \tag{7.115}$$

The relevant term that describes the interaction is $g\bar{\Psi}\gamma_\mu\mathcal{W}^\mu\Psi$ and we simply add P_L:

$$\rightarrow \mathscr{L} = i\bar{\Psi}\gamma_\mu(\partial^\mu - ig\mathcal{W}^\mu P_L)\Psi - \frac{1}{4} Tr\big(\mathcal{W}_{\mu\nu}\mathcal{W}^{\mu\nu}\big). \tag{7.116}$$

Here P_L acts on a doublet and is therefore

$$P_L\Psi = \begin{pmatrix} P_L & 0 \\ 0 & P_L \end{pmatrix}\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} (\psi_1)_L \\ (\psi_2)_L \end{pmatrix}. \tag{7.117}$$

One P_L is enough to project the left-chiral component out of both doublets $\bar{\Psi}$ and Ψ. To see this we need three identities:

- $(P_L)^2 = P_L$, which is obvious from the explicit matrix form and because every projection operator must have this property[66]. Projecting twice must be the same as projecting one time.

[64] Recall the definition of the γ_μ matrices in Eq. 6.13 and don't let yourself get confused about the missing γ_4 matrix. There is an alternative convention that uses γ_4 instead of γ_0 and to avoid interference between those conventions, the matrix here is commonly called γ_5.

[65] Maybe you wonder why we define P_L as so complicated and do not start with the explicit matrix form right away. We do this, because it's possible to work in a different basis where the matrices γ_μ look completely different. (For more information about this have a look at Section 8.10). Take note that in the Lagrangian the Dirac spinors appear always in combination with the matrices γ_μ. We can always add a $1 = U^{-1}U$, with some arbitrary invertible matrix U, between them. For example, $\partial_\mu\bar{\Psi}\gamma_\mu\Psi = \partial_\mu\bar{\Psi}\underbrace{U^{-1}U}_{=1}\gamma_\mu\underbrace{U^{-1}U}_{=1}\Psi = \partial_\mu\underbrace{\bar{\Psi}U^{-1}}_{=\bar{\Psi}'}\underbrace{U\gamma_\mu U^{-1}}_{\gamma'_\mu}\underbrace{U\Psi}_{\Psi'}$. Physics is of course completely independent of such transformations, but we can use this to simplify computations. The basis we prefer to work with in this text is called Weyl Basis. In other bases the two components of a Dirac spinor are mixtures of χ_L and ξ_R. Nevertheless, the projection operator defined as $P_L = \frac{1-\gamma_5}{2}$, always projects out the left-chiral component, because $P_L^{\text{Weyl}}\Psi^{\text{Weyl}} = \Psi_{\text{L}}^{\text{Weyl}} \Rightarrow P'_L\Psi' = \frac{1-\gamma'_5}{2}\underbrace{\Psi'}_{U\Psi^{\text{Weyl}}} = \frac{1-Ui\gamma_0U^{-1}U\gamma_1U^{-1}U\gamma_2U^{-1}U\gamma_3U^{-1}}{2}U\Psi^{\text{Weyl}} = \frac{U-Ui\gamma_0\gamma_1\gamma_2\gamma_3}{2}\Psi^{\text{Weyl}} = U\left(\frac{1-\gamma_5}{2}\right)\Psi^{\text{Weyl}} = U\Psi_L^{\text{Weyl}} = \Psi'_L$ ✓

[66] Another defining condition of any projection operator is $P_LP_R = P_RP_L = 0$, which is here fulfilled as you can check by using the explicit form of P_L,P_R and γ_5.

- $\{\gamma_5, \gamma_\mu\} = \gamma_5\gamma_\mu + \gamma_\mu\gamma_5 = 0$, which you can check by brute force computation[67].
- $(P_L)^\dagger = P_L$, because γ_5 is real, as can be seen from the explicit matrix form: $\gamma_5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

[67] Or using another identity $\{\gamma_\mu, \gamma_\nu\} = \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = \frac{1}{2}\eta_{\mu\nu}$, where $\eta_{\mu\nu}$ is the Minkowski metric and the definition $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$.

The second identity simply tells us that $\gamma_5\gamma_\mu = -\gamma_\mu\gamma_5$, i.e. that we can switch the position of γ_5 and any γ_μ matrix, as long as we include a minus sign. This tells us

$$\gamma_\mu P_L = \gamma_\mu \frac{1-\gamma_5}{2} = \frac{1+\gamma_5}{2}\gamma_\mu = P_R\gamma_\mu. \tag{7.118}$$

We can now rewrite the relevant term of Eq. 7.116:

$$\begin{aligned} g\bar{\Psi}\gamma_\mu \mathcal{W}^\mu P_L \Psi &= g\bar{\Psi}\gamma_\mu \mathcal{W}^\mu \underbrace{P_L^2}_{=P_L} \Psi \\ &= g \underbrace{\bar{\Psi}}_{\Psi^\dagger\gamma_0} \underbrace{\gamma_\mu P_L}_{P_R\gamma_\mu} \mathcal{W}^\mu \underbrace{P_L\Psi}_{\Psi_L} \\ &= g\Psi^\dagger \underbrace{\gamma_0 P_R}_{P_L\gamma_0} \gamma_\mu \mathcal{W}^\mu \Psi_L \\ &\underbrace{=}_{\text{Using } P_L^\dagger = P_L \text{ and } (AB)^\dagger = ((AB)^T)^\star = (B^T A^T)^\star = B^\dagger A^\dagger} (P_L\Psi)^\dagger \gamma_0\gamma_\mu \mathcal{W}^\mu \Psi_L \\ &= g(\underbrace{P_L\Psi}_{=\Psi_L})^\dagger \gamma_0\gamma_\mu \mathcal{W}^\mu \Psi_L \\ &= g\bar{\Psi}_L \gamma_\mu \mathcal{W}^\mu \Psi_L \qquad \checkmark \end{aligned} \tag{7.119}$$

We can see here that one projection operator is really enough to guarantee that both Ψ and $\bar{\Psi}$ are left-chiral.

Now we know how we can describe mathematically that only left-chiral fields interact via the weak force, but why does this mean that parity is violated? To understand this we need to parity transform this term, because if it isn't invariant, the physical system in question is different from its mirror image[68]. Here we need the parity operator for spinors[69] P_{spinor} and vectors[70] P_{vector}. The transformation yields

[68] This means an experiment, whose outcome depends on this term of the Lagrangian, will find a different outcome if everything in the experiment is arranged mirrored.

[69] The parity operator for spinors was derived in Section 3.7.9. Using the γ_μ matrices, we can write the parity operator derived there as $P = \gamma_0 = \begin{pmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

[70] The parity operator for vectors is simply $P_{\text{vector}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ as already mentioned in Eq. 3.140.

$$\begin{aligned} \underbrace{\bar{\Psi}}_{=\Psi^\dagger\gamma_0} \gamma_\mu \mathcal{W}^\mu P_L \Psi &\to (P_{\text{spinor}}\Psi)^\dagger \gamma_0\gamma_\mu (P_{\text{vector}}\mathcal{W}^\mu) P_L (P_{\text{spinor}}\Psi) \\ &= (\Psi)^\dagger \gamma_0\gamma_0\gamma_\mu (P_{\text{vector}}\mathcal{W}^\mu) P_L \gamma_0 \Psi \\ &\underbrace{=}_{\text{using } \{\gamma_5,\gamma_0\}=0 \text{ and } P_L = \frac{1-\gamma_5}{2}} (\Psi)^\dagger \gamma_0\gamma_0\gamma_\mu\gamma_0 (P_{\text{vector}}\mathcal{W}^\mu) P_R \Psi. \end{aligned} \tag{7.120}$$

Then we can use $\gamma_0\gamma_0\gamma_0 = \gamma_0$ and $\gamma_0\gamma_i\gamma_0 = -\gamma_i$, as you can check by looking at the explicit form of the matrices. Furthermore, we have $P_{\text{vector}}W^0 = W^0$ and $P_{\text{vector}}W^i = -W^i$, which follows from the explicit form of P_{vector}. We conclude these two minus signs cancel each other and the parity transformed term of the Lagrangian reads:

$$(P_{\text{spinor}}\Psi)^\dagger\gamma_0\gamma_\mu(P_{\text{vector}}\mathcal{W}^\mu)P_L(P_{\text{spinor}}\Psi) = \bar{\Psi}\gamma_\mu\mathcal{W}^\mu P_R\Psi \neq \bar{\Psi}\gamma_\mu\mathcal{W}^\mu P_L\Psi \tag{7.121}$$

Therefore this term is *not* invariant and **parity is violated**.

Parity violation has another important implication. Recall that we always write things below each other between two big brackets if they can transform into each other[71]. For example, we use four-vectors, because their components can transform into each other through rotations or boosts. In this section we learned that only left-chiral particles interact via the weak force and the correct term in the Lagrangian is $\bar{\Psi}\gamma_\mu\mathcal{W}^\mu P_L\Psi = \bar{\Psi}_L\gamma_\mu\mathcal{W}^\mu\Psi_L$. In physical terms this term means that the components of the left-chiral doublets, which means the two spin $\frac{1}{2}$ fields $(\psi_1)_L$ and $(\psi_2)_L$ can transform into each other through weak-interactions. Right-chiral fields do not interact via the weak force and therefore $(\psi_1)_R$, $(\psi_2)_R$ aren't transformed into each other. Therefore writing them below each other between two big brackets makes no sense. In mathematical terms this means that right-chiral fields form $SU(2)$ **singlets**, i.e. are objects transforming according to the 1 dimensional representation of $SU(2)$, which do not change at all[72]. So let's summarize:

[71] This is explained in Appendix A.

[72] This was explained in Section 3.6.2.

- Left-chiral fields are written as $SU(2)$ **doublets**: $\Psi_L = \begin{pmatrix} (\psi_1)_L \\ (\psi_2)_L \end{pmatrix}$, because they interact via the weak force and therefore can transform into each other. They transform under the two-dimensional representation of $SU(2)$:

$$\Psi_L \to \Psi'_L = e^{i\vec{a}\frac{\vec{\sigma}}{2}}\Psi_L\,. \tag{7.122}$$

- Right-chiral fields are described by $SU(2)$ **singlets**: $(\psi_1)_R$, $(\psi_2)_R$, because they do not interact via the weak force and therefore can't transform into each other. Therefore they transform under the one-dimensional representation of $SU(2)$:

$$\begin{aligned}(\psi_1)_R &\to (\psi_1)'_R = e^0(\psi_1)_R = (\psi_1)_R \\ (\psi_2)_R &\to (\psi_2)'_R = e^0(\psi_2)_R = (\psi_2)_R\,.\end{aligned} \tag{7.123}$$

Now we move on and try to understand how mass terms for spin $\frac{1}{2}$ particles can be added in the Lagrangian without spoiling any crucial symmetry.

7.5 Lepton Mass Terms

At the beginning of Section 7.2, we discovered that we can't include arbitrary mass terms $\bar{\Psi}m\Psi$ without spoiling the $SU(2)$ symmetry. Now we will see that parity violation makes this problem even bigger. After discussing the problem, we will see that again the Higgs mechanism is a solution.

In the last section we talked a bit about the chirality of the coupling term: $\bar{\Psi}\gamma_\mu\sigma_j W_j^\mu P_L \Psi$. What about the chirality of a mass term? Take a look again at the invariants without derivatives for spinors, which we derived in Eq. 6.7 and Eq. 6.8:

$$I_1 := (\chi_{\dot{a}})^T \xi^{\dot{a}} = (\chi_L)^\dagger \xi_R \qquad \text{and } I_2 := (\xi^a)^T \chi_a = (\xi_R)^\dagger \chi_L \tag{7.124}$$

We can write these invariants using Dirac spinors as[73]

$$\begin{aligned}
\bar{\psi}\psi &= \bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L \\
&= \begin{pmatrix} \chi_L^\dagger & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \xi_R \end{pmatrix} + \begin{pmatrix} 0 & \xi_R^\dagger \end{pmatrix} \begin{pmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{pmatrix} \begin{pmatrix} \chi_L \\ 0 \end{pmatrix} \\
&= \chi_L^\dagger \xi_R + \xi_R^\dagger \chi_L \quad \checkmark
\end{aligned} \tag{7.125}$$

[73] The Dirac spinors ψ_L and ψ_R are defined using the chiral-projection operators introduced in the last section: $\psi_L = P_L\psi$ and $\psi_R = P_R\psi$. And we have as always $\bar{\psi} = \psi^\dagger\gamma_0$.

We can see that Lorentz invariant mass terms always combine left-chiral with right-chiral fields. This is a problem, because left-chiral and right-chiral fields transform differently under $SU(2)$ transformations, as explained at the end of the last section. The left-chiral fields are doublets, whereas the right-chiral fields are singlets. The multiplication of a doublet and singlet is not $SU(2)$ invariant. For example

$$\underbrace{\bar{\Psi}_L}_{\text{doublet}} \underbrace{\psi_R}_{\text{singlet}} \to \bar{\Psi}'_L \psi'_R = \bar{\Psi}_L \mathrm{e}^{-ib_i \frac{\sigma_i}{2}} \psi_R \neq \bar{\Psi}_L \psi_R \tag{7.126}$$

From the experience with mass terms for spin 1 fields, we know what to do: Instead of considering terms as above, we add $SU(2)$ invariant coupling terms with 0 fields to the Lagrangian. Then, by choosing the vacuum value for the spin 0 field we break the symmetry and generate mass terms.

A $SU(2)$, $U(1)$ **and** Lorentz invariant term, coupling a spin 0 doublet and our spin $\frac{1}{2}$ fields together, is given by

$$\bar{\Psi}_L \Phi \psi_R. \tag{7.127}$$

To see the invariance we transform this term with a $SU(2)$ transformation[74]

[74] Remember $\Phi_L \to \Phi'_L = \mathrm{e}^{ib_i(x)\sigma_i}\Phi_L$ and $\sigma_i^\dagger = \sigma_i$

$$\bar{\Psi}_L \Phi \psi_R \to \bar{\Psi}'_L \Phi' \psi_R = \bar{\Psi}_L e^{-ib_i(x)\sigma_i} e^{ib_i(x)\sigma_i} \Phi \psi_R = \bar{\Psi}_L \Phi \psi_R \quad \checkmark$$

and equally for a $U(1)$ transformation:

$$\bar{\Psi}_L \Phi \psi_R \to \bar{\Psi}'_L \Phi \psi'_R = \bar{\Psi}_L e^{-ia(x)} \Phi e^{ia(x)} \psi_R = \bar{\Psi}_L \Phi \psi_R \quad \checkmark$$

The spin 0 field does not transform at all under Lorentz transformations[75] and therefore the term is Lorentz invariant, because we have the same Lorentz invariant terms as in Eq. 7.125.

[75] By definition a spin 0 field transforms according to the $(0,0)$ representation of the Lorentz group. In this representation all Lorentz transformations are trivially the identity transformation. This was derived in Section 3.7.4.

This kind of term is called **Yukawa term** and we add it, with the equally allowed Hermitian conjugate to the Lagrangian[76] $-\lambda_2$

$$\mathscr{L} = -\lambda_2 (\bar{\Psi}_L \Phi \psi_{2R} + \bar{\psi}_{2R} \bar{\Phi} \Psi_L). \tag{7.128}$$

[76] The strange name $-\lambda_2$ and why only add ψ_{2R} here will become clear in a moment, because terms including ψ_{1R} and $-\lambda_1$ will be discussed afterwards.

The coupling constant λ_2 is called a **Yukawa coupling**. This extra term does not only describe the interaction between the fermions and the Higgs field, but also leads to mass terms for the spin $\frac{1}{2}$ fields after the symmetry breaking. We put the expansion around the vacuum expectation value (Eq. 7.88)

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

into the Lagrangian, which yields

$$\begin{aligned}
\mathscr{L} &= -\frac{\lambda_2}{\sqrt{2}} \left(\left(\bar{\Psi}_{1L}, \bar{\Psi}_{2L} \right) \begin{pmatrix} 0 \\ v+h \end{pmatrix} \psi_{2R} + \bar{\psi}_{2R} \left(0, v+h \right) \begin{pmatrix} \Psi_{1L} \\ \Psi_{2L} \end{pmatrix} \right) \\
&= -\frac{\lambda_2 (v+h)}{\sqrt{2}} \left(\bar{\Psi}_{2L} \psi_{2R} + \bar{\psi}_{2R} \Psi_{2L} \right). \qquad (7.129)
\end{aligned}$$

Equation 7.125 tells us this is equivalent to

$$\begin{aligned}
&= -\frac{\lambda_2 (v+h)}{\sqrt{2}} \bar{\psi}_2 \psi_2 \qquad (7.130) \\
&= \underbrace{-\frac{\lambda_2 v}{\sqrt{2}} (\bar{\psi}_2 \psi_2)}_{\text{Fermion mass term}} - \underbrace{\frac{\lambda_f h}{\sqrt{2}} (\bar{\psi}_2 \psi_2)}_{\text{Fermion-Higgs interaction}}. \qquad (7.131)
\end{aligned}$$

We see that through the Higgs mechanism we get indeed the required mass terms. Again, we used symmetry constraints to add a term to the Lagrangian and this yields after spontaneous symmetry breaking mass terms for the spin $\frac{1}{2}$ fields. Take note that we only generated mass terms for the second field inside the doublet ψ_2. What about mass terms for the first field ψ_1?

To get mass terms for the first field ψ_1 we need to consider coupling terms to the charge-conjugated[77] Higgs field $\tilde{\Phi} = \epsilon \Phi^\star$, because

[77] Charge conjugation is explained in Section 3.7.10.

$$\Phi = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \rightarrow \tilde{\Phi} = \epsilon \Phi^\star = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{v+h}{\sqrt{2}} \\ 0 \end{pmatrix} . \tag{7.132}$$

Following the same steps as above with the charge conjugated Higgs field leads to mass terms for ψ_1:

$$\begin{aligned} \mathscr{L} &= -\lambda_f (\bar{\Psi}_L \tilde{\Phi} \psi_{1R} + \bar{\psi}_{1R} \bar{\tilde{\Phi}} \Psi_L) \\ &= -\frac{\lambda_1}{\sqrt{2}} \left(\left(\bar{\Psi}_{1L}, \bar{\Psi}_{2L} \right) \begin{pmatrix} \frac{v+h}{\sqrt{2}} \\ 0 \end{pmatrix} \Psi_{1R} + \bar{\Psi}_{1R} \begin{pmatrix} \frac{v+h}{\sqrt{2}} \\ 0 \end{pmatrix} \begin{pmatrix} \Psi_{1L} \\ \Psi_{2L} \end{pmatrix} \right) \\ &= -\frac{\lambda_1 (v+h)}{\sqrt{2}} \left(\bar{\Psi}_{1L} \Psi_{1R} + \bar{\Psi}_{1R} \Psi_{1L} \right) . \end{aligned} \tag{7.133}$$

To understand the rather abstract spin $\frac{1}{2}$ doublets better, we rewrite them more suggestively[78]

$$\Psi = \begin{pmatrix} \nu_e \\ e \end{pmatrix} \tag{7.134}$$

and equivalently for the other leptons μ, ν_μ and τ, ν_τ. This form of the doublets is suggested by experiments, because an electron e is always transformed by weak interactions into another electron e, with possibly different momentum, or an electron-neutrino v_e plus other particles. In weak interactions e and ν_e (equivalently μ and ν_μ or τ and ν_τ) always appear in pairs. This can be understood by looking at the coupling term $\bar{\Psi} \gamma_\mu \mathcal{W}^\mu P_L \Psi = \bar{\Psi} \gamma_\mu (W^\mu)_i \frac{\sigma^i}{2} P_L \Psi$. As discussed in the last section this can be rewritten using the explicit matrix form of the Pauli matrices[79] σ_i, which then gives us terms coupling the components of the doublets together:

$$\begin{aligned} \bar{\Psi} \gamma_\mu \sigma_j W_j^\mu P_L \Psi &= \begin{pmatrix} \bar{v}_e & \bar{e} \end{pmatrix} \gamma_\mu \begin{pmatrix} W_3^\mu & \sqrt{2} W_+ \\ \sqrt{2} W_- & -W_3^\mu \end{pmatrix} P_L \begin{pmatrix} \nu_e \\ e \end{pmatrix} \\ &\underbrace{=}_{\text{using Eq. 7.116}} \begin{pmatrix} (\bar{v}_e)_L & (\bar{e})_L \end{pmatrix} \gamma_\mu \begin{pmatrix} W_3^\mu & \sqrt{2} W_+ \\ \sqrt{2} W_- & -W_3^\mu \end{pmatrix} \begin{pmatrix} (\nu_e)_L \\ (e)_L \end{pmatrix} \\ &= (\bar{v}_e)_L \gamma_\mu W_3^\mu (\nu_e)_L + (\bar{v}_e)_L \gamma_\mu \sqrt{2} W_+ (e)_L \\ &\quad + (\bar{e})_L \gamma_\mu \sqrt{2} W_- (\nu_e)_L - (\bar{e})_L \gamma_\mu W_3^\mu (e)_L . \end{aligned} \tag{7.135}$$

If we want to consider all lepton generations at once, i.e. e, μ and τ, we need to write down three terms like this into the Lagrangian:

$$\bar{\Psi}_e \gamma_\mu \sigma_j W_j^\mu P_L \Psi_e + \bar{\Psi}_\mu \gamma_\mu \sigma_j W_j^\mu P_L \Psi_\mu + \bar{\Psi}_\tau \gamma_\mu \sigma_j W_j^\mu P_L \Psi_\tau , \tag{7.136}$$

[78] A neutrino is always denoted by a ν. In this step we simply give the two fields in the doublet ψ_1 and ψ_2 their conventional names: electron field e and electron-neutrino field ν_e.

[79] This gives us once more
$\sigma_i W_i^\mu = \begin{pmatrix} W_3^\mu & W_1^\mu - iW_2^\mu \\ W_1^\mu + iW_2^\mu & -W_3^\mu \end{pmatrix}$
which we can rewrite using
$W_\pm = \frac{1}{\sqrt{2}}(W_1 \mp W_2)$:
$\Rightarrow \sigma_i W_i^\mu = \begin{pmatrix} W_3^\mu & \sqrt{2} W_+ \\ \sqrt{2} W_- & -W_3^\mu \end{pmatrix}$

which can be written more compactly by introducing $\Psi_l = \begin{pmatrix} \nu_l \\ l \end{pmatrix}$, where $l = e, \mu, \tau$:

$$\bar{\Psi}_l \gamma_\mu \sigma_j W_j^\mu P_L \Psi_l.$$

Using the notation $l = \begin{pmatrix} l_L \\ l_R \end{pmatrix}$ the coupling term between the spin 0 and the electrically charged spin $\frac{1}{2}$ fields after spontaneous symmetry breaking reads

$$-\underbrace{\frac{\lambda_l v}{\sqrt{2}}}_{\text{Fermion mass } m_L}(\bar{l}l) \qquad - \qquad \underbrace{\frac{\lambda_l h}{\sqrt{2}}}_{\text{Fermion-Higgs interaction strength } c_L}(\bar{l}l)\,.$$

The terms for the corresponding neutrinos follow analogously[80].

[80] As already mentioned above, we can only write down such mass terms if there are right-chiral neutrinos. However, so far right-chiral neutrinos were never observed in experiments, although we know that neutrinos have mass. This is one of the open problems of the standard model.

This Lagrangian enables us to predict something about the Higgs field h that can be tested in experiments. For a given lepton l, the mass is given by

$$m_l = \frac{\lambda_l v}{\sqrt{2}} \rightarrow \lambda_l = \frac{m_l \sqrt{2}}{v} \tag{7.137}$$

and the coupling strength of this lepton to the Higgs is given by

$$c_l = \frac{\lambda_l h}{\sqrt{2}} \underbrace{=}_{\text{Eq. 7.137}} \frac{m_l h}{v}. \tag{7.138}$$

The last equation means that the coupling strength of the Higgs to a lepton is proportional to the mass of the lepton. The heavier the lepton the stronger the coupling. The same is true for all particles and the derivation is completely analogous.

There are other spin $\frac{1}{2}$ particles, called **quarks**, that interact via the weak force. In addition, quarks interact via a third force, called the strong force and this will be the topic of Section 7.8, but first we want to talk about mass terms for quarks. Luckily, these can be incorporated analogous to the lepton mass terms.

7.6 Quark Mass Terms

We learned in the last section that an $SU(2)$ doublet contains the particles that are transformed into each other via the weak force. For quarks[81] these are the up- and down quark:

[81] If you've never heard of quarks before, have a look at Section 1.3.

$$q = \begin{pmatrix} u \\ d \end{pmatrix} \tag{7.139}$$

and equally for the strange and charm or top and bottom quarks.

Again, we must incorporate the experimental fact that only left-chiral particles interact via the weak force. Therefore, we have left-chiral doublets and right-chiral singlets:

$$\underbrace{q_L}_{\text{doublet}} = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \quad \rightarrow \quad e^{ia_i \frac{\sigma_i}{2}} q_L \tag{7.140}$$

$$\underbrace{u_R}_{\text{singlet}} \quad \rightarrow \quad u_R$$

$$\underbrace{d_R}_{\text{singlet}} \quad \rightarrow \quad d_R. \tag{7.141}$$

Again, right-chiral particles do not interact via the weak force and therefore they aren't transformed into anything and form a $SU(2)$ singlet (=one component object).

The problem is the same as for leptons: To get something Lorentz invariant, we need to combine left-chiral with right-chiral spinors. Such a combination is not $SU(2)$ invariant and we use again the Higgs mechanism. This means, instead of terms like

$$\bar{q}_L u_R + \bar{q}_L d_R + \bar{u}_R q_L + \bar{d}_R q_L, \tag{7.142}$$

which aren't $SU(2)$ invariant, we consider the coupling of the quarks to a spin 0 field doublet Φ:

$$\lambda_u \bar{q}_L \tilde{\Phi} u_R + \lambda_d \bar{q}_L \Phi d_R + \lambda_u \bar{u}_R \tilde{\Phi} q_L + \lambda_d \bar{d}_R \Phi q_L, \tag{7.143}$$

with coupling constants λ_u, λ_d and the charge conjugated Higgs doublet[82] which is needed in order to get mass terms for the up quarks[83].

Then, everything is analogous to the lepton case: We put the expansion of the Higgs field around its minimum[84] into the Lagrangian, which gives us mass terms plus quark-Higgs coupling terms.

7.7 Isospin

Now it's time we talk about the conserved quantity that follows from $SU(2)$ symmetry. The free Lagrangians are only globally invariant and we need interaction terms to make them locally symmetric. Recall that global symmetry is a special case of local symmetry. Therefore we have global symmetry in every locally invariant Lagrangian and the corresponding conserved quantity is conserved for

[82] This is defined in Eq. 7.132: $\tilde{\Phi} = \epsilon \Phi^\star \underbrace{=}_{} \begin{pmatrix} \frac{v+h}{\sqrt{2}} \\ 0 \end{pmatrix}$.

[83] We defined the doublets as $\begin{pmatrix} u \\ d \end{pmatrix}$. Multiplication of this doublet with $\Phi = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}$ always results in terms proportional to d.

[84] $\Phi = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}$ and equivalently for the charge-conjugated Higgs field.

both cases. The result will be that global $SU(2)$ invariance gives us through Noether's theorem, a new conserved quantity called **isospin**. This is similar to electric charge, which is the conserved quantity that follows from global $U(1)$ invariance.

Noether's theorem for internal symmetry (Section 4.5.5, especially Eq. 4.57) tells us that

$$\partial_0 \underbrace{\int d^3x \, \frac{\partial \mathscr{L}}{\partial(\partial_0 \Psi)} \delta\Psi}_{=Q} = 0 \tag{7.144}$$

The Lagrangian is invariant under transformations of the form

$$\Psi \to e^{ia_i \frac{\sigma_i}{2}} \Psi = (1 + ia_i \frac{\sigma_i}{2} + \ldots)\Psi \tag{7.145}$$

Therefore our infinitesimal variation is $\delta\Psi = ia_i \frac{\sigma_i}{2}\Psi$, with arbitrary a_i. This tells us we get one conserved quantity for each generator, because the Lagrangian is invariant regardless of if two of the three a_i are zero and one isn't. For example, $a_2 = a_3 = 0$ and $a_1 \neq 0$ or $a_1 = a_2 = 0$ and $a_3 \neq 0$. Of course we get another conserved quantity for $a_1 \neq 0$, $a_2 \neq 0$ and $a_3 \neq 0$, which is just the sum of the conserved quantities we get from the individual generators. In other words: we get three independently conserved quantities, one for each generator of $SU(2)$.

The globally invariant, free Lagrangian (Eq. 7.44) is

$$\mathscr{L}_{\text{D1+D2}} = i\bar{\Psi}\gamma_\mu \partial^\mu \Psi.$$

The corresponding conserved quantities Q_i, for example for the electron-neutrino doublet, are[85]

[85] See Eq. 7.144 and as always defined without the arbitrary constants a_i.

$$\begin{aligned} Q_i &= i\bar{\Psi}\gamma_0 \frac{\sigma_i}{2}\Psi \\ &= \begin{pmatrix} v_e \\ e \end{pmatrix}^\dagger \underbrace{\gamma_0\gamma_0}_{=1} \frac{\sigma_i}{2} \begin{pmatrix} v_e \\ e \end{pmatrix}. \end{aligned} \tag{7.146}$$

Recall that only σ_3 is diagonal. This means we are only able to assign a definite value to the two components of the doublet (v_e, e) for the conserved quantity $i = 3$. For the other generators, σ_1 and σ_2, our two components v_e and e aren't eigenstates. We are of course free to choose a different basis, where for example σ_2 is diagonal. Then we can simply redefine what we call v_e and e and get the same result. The thing to take away is that although we have three conserved quantities, one for each generator, we can only use one at a time to label our particles/states.

For $i = 3$ we have

$$Q_3 = \begin{pmatrix} v_e \\ e \end{pmatrix}^\dagger \frac{\sigma_3}{2} \begin{pmatrix} v_e \\ e \end{pmatrix}$$
$$= \frac{1}{2} \begin{pmatrix} v_e \\ e \end{pmatrix}^\dagger \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} v_e \\ e \end{pmatrix}$$
$$= \frac{1}{2} v_e^\dagger v_e - \frac{1}{2} e^\dagger e \tag{7.147}$$

This means we can assign $Q_3(v_e) = \frac{1}{2}$ and $Q_3(e) = -\frac{1}{2}$ as new particle labels. In contrast for $i = 1$, we have

$$Q_1 = \begin{pmatrix} v_e \\ e \end{pmatrix}^\dagger \frac{\sigma_1}{2} \begin{pmatrix} v_e \\ e \end{pmatrix}$$
$$= \frac{1}{2} \begin{pmatrix} v_e \\ e \end{pmatrix}^\dagger \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v_e \\ e \end{pmatrix}$$
$$= \frac{1}{2} v_e^\dagger e + \frac{1}{2} e^\dagger v_e \tag{7.148}$$

and we can't assign any particle labels here, because the matrix σ_1 isn't diagonal.

7.7.1 Labelling States

Recall that in Section 3.5, we introduced the notion of Cartan generators, which is the set of diagonal generators of a given group. In the last section we learned that these generators become especially useful if we want to give new labels to our particles inside a doublet[86] object. A typical $SU(2)$ doublet is of the form

$$\begin{pmatrix} v_e \\ e \end{pmatrix}. \tag{7.149}$$

We can't diagonalize two or more elements of the Lie algebra $\mathfrak{su}(2)$ at the same time and thus the "Cartan subalgebra" consists of only only one element[87]. It is conventional to choose $J_3 = \frac{1}{2}\sigma_3$ as diagonal. The corresponding eigenvalues are $+\frac{1}{2}$ and $-\frac{1}{2}$. A left-chiral neutrino $\begin{pmatrix} v_e \\ 0 \end{pmatrix}$ is an eigenstate of this generator, with eigenvalue $+\frac{1}{2}$ and a left-chiral electron $\begin{pmatrix} 0 \\ e \end{pmatrix}$ an eigenstate of this generator, with eigenvalue $-\frac{1}{2}$. These are new particle labels, which are called the isospin of the neutrino and the electron.

[86] In a later section we will learn that the same can be done for triplets and the conserved quantities following from $SU(3)$ invariance.

[87] This can be seen by noting that there are no commuting generators in $\mathfrak{su}(2)$. Only commuting generators can be diagonalized at the same time.

Following the same line of thoughts we can assign an isospin value to the right-chiral singlets. These transform according to the one-dimensional representation of $SU(2)$, and the generators are in this representation simply zero[88]: $J_i = 0$. Therefore, in this one-dimensional representation, the singlets are eigenstates of the Cartan generator J_3 with eigenvalue zero. The right-chiral singlets, like e_R carry isospin zero. This coincides with the remarks above that right-chiral fields do not interact via the weak force. Just as electrically uncharged objects do not interact via electromagnetic interactions, fields without isospin do not take part in weak interactions.

[88] The right-chiral singlets do not transform at all as explained in Section 3.7.4.

Finally, we can assign isospin values to the three gauge fields $W_+^\mu, W_-^\mu, W_3^\mu$. The three gauge fields form a $SU(2)$ triplet

$$W^\mu = \begin{pmatrix} W_+^\mu \\ W_-^\mu \\ W_3^\mu \end{pmatrix}, \tag{7.150}$$

which transforms according to the three dimensional representation of $SU(2)$. In this representation the Cartan generator J_3 has eigenvalues[89] $+1, -1, 0$ and therefore we assign $Q_3(W_+^\mu) = 1$, $Q_3(W_-^\mu) = -1$, $Q_3(W_3^\mu) = 0$. This is the isospin of the W_+ and the W_- bosons.

[89] This can be seen directly from the explicit matrix form of J_3 in Eq. 3.129: $J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Take note that the triplet $\begin{pmatrix} W_1^\mu & W_2^\mu & W_3^\mu \end{pmatrix}$ simply belongs to a different basis, where J_3 isn't diagonal. This can be seen as another reason for our introduction of $W_\pm^\mu$.

If this is unclear, have a look at how we introduced the three gauge fields W_i^μ. They were included into the Lagrangian in combination with the generators $\sigma_i W_i^\mu$. This can be seen as a basis expansion of some objects W^μ in terms of the basis σ_i:

$$\mathcal{W}^\mu = \sigma_i W_i^\mu = \sigma_1 W_1^\mu + \sigma_2 W_2^\mu + \sigma_3 W_3^\mu,$$

analogous to how we can write a vector in terms of basis vectors:

$$\vec{v} = v_1 \vec{e}_1 + v_2 \vec{e}_2 + v_3 \vec{e}_3.$$

The generators σ_i live in the Lie algebra[90] of $SU(2)$ and consequently our object $\mathcal{W}^\mu$ lives there, too. Therefore, if we want to know how W^μ transforms, we need to know the representation of $SU(2)$ on this vector space, i.e. on its own Lie algebra. In other words: We need to know how the group elements of $SU(2)$ act on their own Lie algebra elements, i.e. its generators. This may seem like a strange idea at first, but actually is a quite natural idea. Recall how we defined a representation: a representation is a map[91] from the group to the space of linear operators over a vector space. So far we only looked

[90] The Lie algebra is a vector space!

[91] To be precise: A homomorphism, which is a map that satisfies some special conditions.

at "external" vector spaces like Minkowski space. The only intrinsic vector space that comes with a group is its Lie algebra[92] ! Therefore it isn't that strange to ask what a group representation on this vector space looks like. This very important representation is called the **adjoint representation**.

[92] A group itself is in general no vector space. Although we can take a look at how the group acts on itself, this is not a representation, but a realization of the group.

Gauge fields (like W_+, W_-, W_3) are said to live in the adjoint representation of the corresponding group. For $SU(2)$ the Lie algebra is three dimensional, because we have three basis generators and therefore the adjoint representation is three dimensional. Exactly how we are able to write the components of a vector between two brackets[93], we can write the component of W^μ between two brackets[94], which is what we call a triplet . The generators in the adjoint representation are connected to the three dimensional generator we derived earlier through a basis transformation.

[93] $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$

[94] $W^\mu = \begin{pmatrix} W_1^\mu \\ W_2^\mu \\ W_3^\mu \end{pmatrix}$

In the following section we move on to the "next higher" internal symmetry group $SU(3)$. Demanding local $SU(3)$ invariance of the Lagrangian gives us the correct Lagrangian which describes strong interactions.

7.8 $SU(3)$ Interactions

For **three** fermion fields we can find a locally $SU(3)$ invariant Lagrangian in exactly the same way we did in the last chapter for two fields and $SU(2)$. This symmetry is not broken and the corresponding spin 1 fields, called gluon fields, are massless. $SU(3)$ is the group of all unitary 3×3 matrices with unit determinant, i.e.

$$U^\dagger U = UU^\dagger = 1 \qquad \det U = 1. \tag{7.151}$$

As usual for Lie groups we can write this as an exponential function[95]

[95] As already noted in Section 2.4, Capital Roman letters $A, B, \ldots$ are always summed from 1 to 8.

$$U = \mathrm{e}^{iT_A\theta_A}. \tag{7.152}$$

The defining equations (Eq. 7.151) of the group require, as for[96] $SU(2)$, the generators to be Hermitian and traceless

[96] See Eq. 3.79 and the following text plus equations, where the basis was given by the 2×2 Pauli matrices.

$$T_A^\dagger = T_A \tag{7.153}$$

$$\mathrm{tr}(T_A) = 0\,. \tag{7.154}$$

A basis for those traceless, Hermitian generators is, at least in *one* representation, given by eight[97] 3×3 matrices, called **Gell-Mann matrices**:

[97] It can be shown that in general for $SU(N)$ the Lie-Algebra is $N^2 - 1$ dimensional.

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{7.155}$$

The generators of the group are connected to these Gell-Mann matrices, just as the Pauli matrices were connected to the generators[98] of the $SU(2)$ group via $T_A = \frac{1}{2}\lambda_A$. The Lie algebra for this group is given by

$$[T_A, T_B] = if^{ABC}T^C, \tag{7.156}$$

where we adopted the standard convention that capital letters like A, B, C can take on every value from 1 to 8. f^{ABC} are called the structure constants of $SU(3)$, which for $SU(2)$ were given by the Levi-Civita symbol ϵ_{ijk}. They can be computed by brute-force computation, which yields[99]

$$f^{123} = 1 \tag{7.157}$$

$$f^{147} = -f^{156} = f^{246} = f^{257} = f^{345} = -f^{367} = \frac{1}{2} \tag{7.158}$$

$$f^{458} = f^{678} = \frac{\sqrt{3}}{2}, \tag{7.159}$$

where all others can be computed from the fact that the structure constants f^{ABC} are antisymmetric under permutation of any two indices. For example

$$f^{ABC} = -f^{BAC} = -f^{CBA}. \tag{7.160}$$

All other possibilities, which cannot be computed by permutation, vanish.

Analogous to what we did for $SU(2)$ in Section 7.2, we introduce a triplet of spin $\frac{1}{2}$ fields

$$Q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \tag{7.161}$$

and exactly as for $SU(2)$ we get new labels for the objects inside this triplet. We will discuss these new labels in the next section.

[98] $J_i = \frac{\sigma_i}{2}$, see Eq. 3.81 and the explanations there.

[99] This is not very enlightening, but we list it here for completeness.

To make the Lagrangian

$$\mathscr{L} = i\overline{Q}\partial_\mu \gamma^\mu Q - \overline{Q}mQ \tag{7.162}$$

locally $SU(3)$ invariant, one again adds coupling terms between the spin $\frac{1}{2}$ fields and new spin 1 fields. The derivation is analogous as for $SU(2)$, but the computations are quite cumbersome, so we just quote the final Lagrangian[100]

[100] Remember: the sum over capital letters $(A, B, C, ...)$ runs from 1 to 8

$$\mathscr{L} = -\frac{1}{4}\mathcal{G}_{\alpha\beta}\mathcal{G}^{\alpha\beta} + \overline{Q}(iD_\mu\gamma^\mu - m)Q, \tag{7.163}$$

and the field strength tensor $\mathcal{G}_{\alpha\beta}$ for the spin 1 gluon fields $\mathcal{G}_\alpha \equiv T^C G^C_\alpha$ is defined as

$$\mathcal{G}_{\alpha\beta} = \partial_\alpha \mathcal{G}_\beta - \partial_\beta \mathcal{G}_\alpha - g[\mathcal{G}_\alpha, \mathcal{G}_\beta]\,. \tag{7.164}$$

Here T^C denotes the generators of $SU(3)$ that were defined at the beginning of this section. Furthermore, D_α is defined as

$$D_\alpha = \partial_\alpha - igT^C G^C_\alpha = \partial_\alpha - ig\mathcal{G}_\alpha\,. \tag{7.165}$$

As you can check every term here is completely analogous to the $SU(2)$ case, except we now have different generators with different commutation properties.

7.8.1 Color

From global $SU(3)$ symmetry we get through Noether's theorem new conserved quantities. This is analogous to what we discussed for $SU(2)$ in Section 7.7. Following the same lines of thought as for $SU(2)$ tells us that we have 8 conserved quantities, one for each generator. Again, we can only use the conserved quantities that belong to the diagonal generators as particle labels. $SU(3)$ has two Cartan[101] generators $\frac{1}{2}\lambda_3$ and $\frac{1}{2}\lambda_8$. Therefore, every particle that interacts via the strong force carries two additional labels, corresponding to the eigenvalues of the Cartan generators.

[101] Recall, Cartan generators = diagonal generators.

The eigenvalues[102] of $\frac{1}{2}\lambda_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ are $+\frac{1}{2}, -\frac{1}{2}, 0$. For $\lambda_8 = \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$ the eigenvalues[103] are $\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{-1}{\sqrt{3}}$. Therefore if we arrange the strong interacting fermions into triplets (in the basis spanned by the eigenvectors of the Cartan generators), we can assign them the following labels, with some arbitrary spinor ψ:

[102] The eigenvectors are of course $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

[103] Corresponding to the same eigenvectors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

$$\left(+\frac{1}{2}, \frac{1}{2\sqrt{3}}\right) \text{ for } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \psi,$$

where one usually defines **red**:= $\left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$. This means something of the form $\begin{pmatrix} \Psi \\ 0 \\ 0 \end{pmatrix}$ is called red[104]. Analogously

$$\left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right) \text{ for } \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \psi$$

with **blue**:= $\left(\frac{-1}{2}, \frac{1}{2\sqrt{3}}\right)$ and for the third possibility we define **green**:= $\left(0, \frac{-1}{\sqrt{3}}\right)$.

[104] Although we use here the familiar sounding labels "red", blue" and "green", there is absolutely no connection with the usual meaning of these words. In the context of $SU(3)$ interactions, we use these words simply as convenient labels. For a nice discussion on why we use color labels, see Chapter 1 in Francis Halzen and Alan D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1st edition, 1 1984. ISBN 9780471887416.

Completely analogous to what we did for $SU(2)$ we assign the color-charge zero to all $SU(3)$ singlets, which are then particles that do not interact via the strong force. Formulated differently: they are colorless. In addition, one can use the (8-dimensional[105]) adjoint representation of $SU(3)$ to assign color to the gauge fields G_A^μ, i.e. the gluons, completely analogous to how we assigned isospin to the W-Bosons in Section 7.7.

[105] The adjoint representation of $SU(3)$ is 8 dimensional, because we have 8 generators.

7.8.2 Quark Description

Recall that spin $\frac{1}{2}$ particles[106], which interact via the strong force are called quarks. If we want to talk about quarks we have to consider quite a lot of things:

[106] Spin $\frac{1}{2}$ are created by spin $\frac{1}{2}$ fields, as we will learn in Chapter 6.

- Quarks are $SU(3)$ triplets, denoted by Q. Inside a triplet we have the same quark, say an up-quark, in different colors: $U = \begin{pmatrix} u_r \\ u_b \\ u_g \end{pmatrix}$.
  The triplets always appear in pairs $\bar{Q}Q$ in order to get something $SU(3)$ invariant, exactly as we always need doublet pairs in order to get something $SU(2)$ invariant. Instead of writing $\bar{Q}Q$, we can use an index notation: $\bar{Q}Q = \bar{q}_c q_c$, where the index c stands for color $c = r, g, b$.
- In addition, left-chiral quarks are $SU(2)$ doublets, because they interact via the weak force, too. Each object in this doublet[107] (=each quark) is a triplet: $q = \begin{pmatrix} u_c \\ d_c \end{pmatrix}$. This can become very confusing, very fast and therefore the color index c is suppressed unless strong interactions are considered.
- As if this weren't enough, we need to remember that each quark is described by a Dirac spinor, which consits of two Weyl spinors

[107] Each quark doublet consists of two different quarks, for example an up- and a down-quark or a top- and a bottom-quark.

$u_c = \begin{pmatrix} (\chi_u^L)_c \\ (\xi_u^R)_c \end{pmatrix}$. The upper Weyl-Spinor describes a left-chiral and the lower component the same quark with right-chirality.

Having talked about this, let's return to $SU(3)$ interactions. Happily, there is no experimental need for mass terms for the gauge bosons in the Lagrangian, because all experiments indicate that the gauge bosons of $SU(3)$, called gluons, are massless. Therefore $SU(3)$ is not broken.

Furthermore, the $SU(3)$ symmetry poses no new problems regarding mass terms for the fermions in the triplet, because a term of the form

$$\bar{Q}mQ \tag{7.166}$$

is $SU(3)$ invariant, as long as all particles in the triplet have equal mass. This means m is proportional to the unit matrix[108]. The objects inside a triplet describe the same quark in different colors, which indeed have equal mass. For example, for an up-quark the triplet is

[108] $m = m\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, instead of $m = m\begin{pmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{pmatrix}$

$$U = \begin{pmatrix} u_r \\ u_b \\ u_g \end{pmatrix}, \tag{7.167}$$

where u_r denotes a red, u_b a blue and u_g green up-quark, which all have the same mass.

The other spin $\frac{1}{2}$ particles, like electrons or neutrinos, do not carry color and therefore do not couple to gluons. The interactions following from local $SU(3)$ invariance are called **strong interactions**, because the coupling constant is much bigger than for electromagnetic ($U(1)$) or weak ($SU(2)$) interactions.

7.9 The Interplay Between Fermions and Bosons

This section summarizes what we discovered in this chapter and puts it in a more physical context. We will learn later that spin $\frac{1}{2}$ fields create and destroy spin $\frac{1}{2}$ particles. Analogously, spin 1 fields create and destroy spin 1 particles. In this chapter we derived the Lagrangians that describe how different fields and therefore particles interact with each other.

As already mentioned in Section 1.3 we call spin $\frac{1}{2}$ particles fermions and spin 1 particles bosons. The standard interpretation is that fermions make up matter and bosons mediate the forces between matter. We can now understand how this comes about.

We started the chapter with Lagrangians describing free fields, which we derived in Chapter 6. Then we discovered internal symmetries for the Lagrangian describing one, two or three free spin $\frac{1}{2}$ fields. These internal symmetries are only global symmetries, which is quite unconvincing because of special relativity. More natural would be local symmetries.

We then discovered that we could make the Lagrangians locally invariant by introducing additional coupling terms. These coupling terms describe the interaction of our spin $\frac{1}{2}$ fields with new spin 1 fields. For historic reasons the internal symmetries here are called gauge symmetries and we therefore call these new spin 1 fields, gauge fields. Through Noether's theorem we get for each internal symmetry new conserved quantities. These are interpreted as charges, analogous to electric charge that follows for $U(1)$ symmetry.

- To get a locally $U(1)$ invariant Lagrangian, we need one gauge field A^μ. The final Lagrangian describes correctly electromagnetic interactions. $U(1)$ symmetry tells us that electric charge is conserved.
- To get a locally $SU(2)$ invariant Lagrangian, we need three gauge fields $W_1^\mu, W_2^\mu, W_3^\mu$. The final Lagrangian describes correctly weak interactions. $SU(2)$ symmetry tells us that isospin is conserved.
- To get a locally $SU(3)$ invariant Lagrangian, we need eight such fields $G_1^\mu, G_2^\mu, \ldots$. The final Lagrangian describes correctly strong interactions. $SU(3)$ symmetry tells us that color is conserved.

Different gauge bosons (spin 1 particles) are responsible for a different kind of force. The electromagnetic force is mediated by photons, which is created by the $U(1)$ gauge field A_μ. The weak force is mediated by W^+, W^- and Z bosons and the strong force by 8 different gluons, which are created by the corresponding $SU(2)$ and $SU(3)$ gauge fields.

In addition, we discovered that $SU(2)$ symmetry forbids mass terms in the Lagrangian. From experiments we know this is incorrect. The solution that enables us to include mass terms without spoiling any symmetry is the Higgs mechanism. It works by including additional terms, describing coupling of our spin 1 and spin $\frac{1}{2}$ fields to a new spin 0 field, called the Higgs field. By breaking $SU(2)$ symmetry spontaneously and expanding the Higgs field around a new, no longer symmetric minimum, we get the required mass terms in the Lagrangian.