Chapter 6

Gauge Theories in Particle Physics

In this chapter we will put to use the mathematical formalism of groups learnt in Chapter 2 to establish the $U(1)$, $SU(2)$, and $SU(3)$ symmetries of the Standard Model Lagrangian that give rise to the electromagnetic, weak, and strong interactions, respectively.

6.1 Gauge Invariance in Quantum Theory

We saw in Section 2.6 how a local phase transformation of the fermion wave function $\psi \to e^{iq\theta(x)}\psi$ alongwith a simultaneous redefinition of the eletromagnetic field $A^\mu \to A^\mu + \partial_\mu \theta(x)$ leaves the Lagrangian invariant. The result can, in fact, be interpreted to say that the local gauge invariance (*phase invariance* would be more appropriate, but we'll honor the historical legacy) requires the presence of a field $A^\mu = (V, \vec{A})$. In *Example 1* of Sec. 2.6, we solved for the scalar field as a static function of space. However, we could have written it just as well as a plane wave normalized to a single quantum of a defnite energy ω and momentum $\vec{k}$:

$$\phi = \frac{1}{\sqrt{2\omega}}\left(ae^{i(\vec{k}\cdot\vec{r}-\omega t)} + a^\dagger e^{-i(\vec{k}\cdot\vec{r}-\omega t)}\right) = \frac{1}{\sqrt{2\omega}}\left(ae^{ik^\mu x_\mu} + a^\dagger e^{-ik^\mu x_\mu}\right), \quad (6.1)$$

where $a^\dagger$ creates quanta associated with the field ϕ and a desroys them.

In a similar fashion, the solution for the vector field A^μ can be expanded in terms of particle creation and destruction operators. Consequently, there must be an associated particle, and since the field is described by a 4-vector, it must be associated with a vector, i.e. spin-1, particle. Since the same effect occurs with any charged particle (not just fermions), the interaction of the new particle, which we interpret as being the photon, is the same with any charged particle - it is a universal interaction. Thus, phase invariance of the theory for electrically charged particles requires that there must be a photon and an electromagnetic interaction of precisely the observed kind. Note, however, that the numerical value of the charge is undetermined.

In a sense, the existence and form of the electromagnetic interaction has been derived. If a particle carries a charge and the theory is invariant under certain phase transformations, then associated fields, called *gauge fields*, and associated spin-1 particles, called *gauge bosons* must exist. These allow us to write the associated interaction Lagrangians. There are three known gauge transformations under which the theory is invariant, and three associated sets of gauge bosons. Why these three and not others, or whether there are others, is not known.[1]

The Lagrangian (and wave equation) for a free charged particle can be modified to describe its interaction with a photon by replacing the ordinary 4-gradient with the *covariant derivative*:

$$D^\mu = \partial^\mu - iqA^\mu, \tag{6.2}$$

[2] This concept can be generalized to other (i.e. non-electromagnetic) charges as well. Suppose we want the theory to be invariant under a transformation where particle states change as

$$\psi' = U\psi \tag{6.3}$$

for some U. We want to define

$$D^\mu = \partial^\mu - igA^\mu, \tag{6.4}$$

where A^μ is the interacting field that has to be added to keep the theory invariant, but now we don't know how A^μ itself transforms. We also want

$$D^{\mu\prime}\psi' = U(D^\mu\psi), \tag{6.5}$$

i.e.

$$(\partial^\mu - igA^{\mu\prime})U\psi = U(\partial^\mu - igA^\mu)U\psi. \tag{6.6}$$

This can be solved for $A^{\mu\prime}$:

$$-igA^{\mu\prime}U\psi = -\partial^\mu(U\psi) + U\partial^\mu\psi - igUA^\mu\psi = -(\partial^\mu U)\psi - igUA^\mu\psi. \tag{6.7}$$

Since each term acts on an arbitrary state ψ, we can drop the ψ and multiply from the right by U^{-1}, so

$$A^{\mu\prime} = -\frac{i}{g}(\partial^\mu U)U^{-1} + UA^\mu U^{-1}. \tag{6.8}$$

This is how A^μ must transform for any U. (*Exercise: verify that this gives the expected answer for* $g = q$ *and* $U = e^{-iq\theta}$)

Equation 6.8 is very general and remains valid if U a matrix, rather than a scalar, in an internal space. The A^μ is also a matrix, so the order of multiplication is important.[3]

---

[1] Another way to phrase the interpretation of what we have observed is that we cannot distinguish between the effects of a local change in phase convention and the effects of a new vector field.

[2] For charged fermions, this can be seen by comparing Eqs. 3.31 and 3.35.

[3] Of course, if U and A^μ are not matrices, then $UA^\mu U^{-1} = A^\mu$, as is the case in electromagnetism.

6.2 Strong Isospin: an example of $SU(2)$

Let us take for example the strong isospin symmetry of nucleons, pions, and other hadrons, that plays an important role in the understanding of nuclei and of hadrons. This will serve as an important aid in "visualizing" the more fundamental weak isospin symmetry, which is what we're really after.

Consider the neutron (n) and the proton (p). Their masses,

$$m_n = 939.57 \text{ MeV}, \quad m_p = 938.27 \text{ MeV}, \tag{6.9}$$

differ by only $\sim$0.15%. No other particles have a similar mass. Both form nuclei and interact similarly. The only obvious difference is that the proton carries an electric charge and the neutron does not. However, their interactions in the nucleus are strong interactions. Strong interactions are not sensitive to electric charge and are $\sim$100 times stronger than electromagnetic ones. So, the electric charge is not of much consequence.

This kind of reasoning led to the idea of picturing n and p as two states of the same entity, a nucleon, N. One could associate an *internal quantum space* called the *strong isospin space*, where the nucleon points in some direction: "up" if it is a p, "down" if n. If strong interactions do not distinguish between a n and a p, it follows that the theory that describes strong interactions is invariant under rotations in the strong isospin space.[4]

Since there are two nucleon states, it is like spin-up and spin-down. So, we can try to put the p and n as states of a spin-like, or $SU(2)$, doublet:

$$N = \begin{pmatrix} p \\ n \end{pmatrix}. \tag{6.10}$$

Another example of a hadron classified as states in $SU(2)$ multiplets is the pion, which has states $\pi^\pm$ amd π^0, with masses $m_{\pi^\pm} = 139.57$ MeV, $m_{\pi^0} = 134.96$ MeV. It can be represented as an isospin-1 state, i.e. a triplet:

$$\pi = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{pmatrix}, \tag{6.11}$$

with charge states

$$\begin{aligned} \pi^\pm &= -\tfrac{1}{\sqrt{2}}(\pi_1 \mp i\pi_2), \\ \pi^0 &= \pi_3. \end{aligned} \tag{6.12}$$

As for nucleons, the pion states have the same strong interactions, and differences in mass and interactions of sizes typical of electromagnetic effects. Later we will see how W bosons can be put in a similar classification under weak isospin.

[4] This can only be an approximate symmetry of nature since it is broken, however slightly, by electromagnetic interactions.

For the strong isospin to be a valid symmetry, it must also hold for interactions. Let us try to write an interaction Lagrangian to describe the most general pion-nucleon interaction at the lowest order. Let $p^\dagger$ create a proton or destroy an antiproton, π^+ destroy a π^+ or create a π^-, n destroy a neutron or create an antineutron and so on. Then the most general 3-particle interaction Lagrangian that conserves neucleon number and electric charge is

$$\mathcal{L}_{\rm int} = g_{pn} p^\dagger n \pi^+ + g_{np} n^\dagger p \pi^- + g_{pp} p^\dagger p \pi^0 + g_{nn} n^\dagger n \pi^0. \tag{6.13}$$

For this Lagrangian to be invariant under rotations in the isospin space, certain relations must hold among the g's. For example, invariance under $p \to n$ rotation requires $g_{pp} = \pm g_{nn}$. Since π is a vector in the isospin space, we must make a vector from the neucleon, so the Lagrangian, made of the scalar product of the two, will be invariant. This is achieved by forming the vector $N^\dagger \vec{\sigma} N$, where σ_i are the Pauli spin matrices. Then

$$\mathcal{L}_{\rm int} = g\left(N^\dagger \vec{\sigma} N\right) \cdot \vec{\pi} = g N^\dagger \vec{\sigma} \cdot \vec{\pi} N \tag{6.14}$$

is invariant under rotations in the isospin space since it is the sacalar product of two vectors in that space. We can write the scalar product in an expanded form:

$$\begin{aligned}
\vec{\sigma}\cdot\vec{\pi} &= \sigma_1\pi_1 + \sigma_2\pi_2 + \sigma_3\pi_3 \\
&= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\pi_1 + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\pi_2 + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\pi_3 \\
&= \begin{pmatrix} \pi_3 & \pi_1 - i\pi_2 \\ \pi_1 + i\pi_2 & -\pi_3 \end{pmatrix} \\
&= \begin{pmatrix} \pi^0 & -\sqrt{2}\pi^+ \\ -\sqrt{2}\pi^- & -\pi^0 \end{pmatrix}.
\end{aligned} \tag{6.15}$$

Then

$$\begin{aligned}
N^\dagger\vec{\sigma}\cdot\vec{\pi}N &= \begin{pmatrix} p^\dagger & n^\dagger \end{pmatrix} \begin{pmatrix} \pi^0 & \sqrt{2}\pi^+ \\ -\sqrt{2}\pi^- & -\pi^0 \end{pmatrix} \begin{pmatrix} p \\ n \end{pmatrix} \\
&= \begin{pmatrix} p^\dagger & n^\dagger \end{pmatrix} \begin{pmatrix} \pi^0 p - \sqrt{2}\pi^+ n \\ -\sqrt{2}\pi^- p - \pi^0 n \end{pmatrix} \\
&= p^\dagger p \pi^0 - \sqrt{2} p^\dagger n \pi^+ - \sqrt{2} n^\dagger p \pi^- - n^\dagger n \pi^0.
\end{aligned} \tag{6.16}$$

Thus we see that for the interactions to be invariant under rotations in the strong isospin space, the couplings g_{pn}, g_{np}, g_{pp}, g_{nn} must be in ratios $1 : 1 : -\frac{1}{\sqrt{2}} : \frac{1}{\sqrt{2}}$.

This technique of writing interactions invariant under rotations in internal spaces to obtain the form of the interaction is used extensively. For weak isospin, it is the W bosons, rather than pions, that have isospin 1. That strong isospin is a nearly good symmetry turns out to be fortuitous, rather than a fundamental feature of nature.

6.3 Non-Abelian Gauge Theories

We can now put together the ideas of internal spaces and of phase invariance. For the moment, let us continue with the nucleon example. We can write a phase transformation where a rotation mixing the proton and neutron states is expressed as a unitary operator in the isospin space,

$$\begin{pmatrix} p' \\ n' \end{pmatrix} = e^{i\vec{\epsilon}\cdot\frac{\vec{\sigma}}{2}} \begin{pmatrix} p \\ n \end{pmatrix}. \tag{6.17}$$

The σ_i are the Pauli matrices and ϵ_i are three parameters that specify the rotation. We can expand the exponential as a power series. Since $\sigma_i^2 = 1$, all powers of any Pauli matrix are either itself or the unit matrix. Note, however, that the order of successive transformations matters, since the rotations do not commute. Formally, this is expressed by the commutator $[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k$. Whenever the order of transformations matters, they are called non-Abelian transformations.

We could equally well consider particles in a multiplet of any group, and demand invariance under the appropriate transformation. If particles a_1, a_2, and a_3 carry quantum numbers in an $SU(3)$ space, we could write

$$\begin{pmatrix} a_1' \\ a_2' \\ a_2' \end{pmatrix} = e^{i\vec{\alpha}\cdot\frac{\vec{\lambda}}{2}} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \tag{6.18}$$

where $\vec{\alpha} = (\alpha_1, \alpha_2, \cdots \alpha_8)$ are the eight rotation parameters, and λ_i, $(i = 1, 2, \cdots 8)$ are the Gell-Mann matrices. Quarks possess such a degree of freedom. It is called *color* because some of its properties are analogous to those of colors, although it has no connection to anything in classical physics nor with anything we experience in the everyday world.

As of now, there is no theoretical principle that tells us what internal spaces to examine. Each internal space where particles carry non-trivial quantum numbers leads to an interaction between particles, mediated by a new set of gauge bosons. In the Standard Model, the complete set of spaces comprises of the $SU(3)$ color space, and the $SU(2)$ and $U(1)$ electroweak spaces. these have been discovered empirically, and do an amazingly good job of describing the all fundamental phenomena and features of the world we live in. Next, we will examine the implications of demanding invariance under transformations in these spaces, but no one yet understands why it is these particular ones that apply and not others. Having gained some insight into the familiar case of the proton and the neutron, let us next examine the quarks and the leptons and their weak isospin.

6.4 Gauge Theories for Quarks and Leptons

Suppose that the quarks and leptons can be put in multiplets of a (*weak*) isospin space, and that the theory should be invariant under transformations of the form of Eq. 6.17. Proceeding as before, we demand invariance under local phase transformations. That is, technically, we make the parameters functions of space and time, $\epsilon_i(x^\mu)$ or $\alpha_i(x^\mu)$. That guarantees we can choose how we define the phase of the quark and lepton states at each point of space-time, rather than having a choice here fix how the phase must be defined somewhere else of sometime later. A theory with a local non-Abelian phase invariance is called a Yang-Mills gauge theory.

No free particle can have an invariance under a non-Abelian gauge transformation, since the derivatives in the wave equation will act on $\epsilon_i(x^\mu)$. We are led again to define a covariant derivative. All the logic of Sec. 6.1 carries over, but now instead of a function $q\theta(x)$ in the exponent of the wave function, we have a function $\epsilon_i\sigma_i$ that transforms non-trivially under a $SU(2)$ group or a function $\alpha_i\lambda_i$ that transforms non-trivially under a $SU(3)$ group. The states describe leptons or quarks.

To define the covariant derivative D^μ, it is necessary for the $SU(2)$ case to introduce a set of three fields, each of which behaves as a 4-vector under Lorentz transformations, in order that we can write a term that transforms as ∂^μ does. Before we needed one A^μ; now we need a W_i^μ for each σ_i. We can define

$$D^\mu = \partial^\mu - ig_2\frac{\vec{\sigma}}{2}\vec{W}^\mu. \tag{6.19}$$

This is the generalization of the Abelian case Eq. 6.4, to include the non-Abelian transformations. If both transformations were relevant, the appropriate terms would add in D^μ. The coupling g_2 is an arbitrary factor wich will determine the interaction strengths. The W_i^μ must be introduced if the theory is to be invariant under weak isospin transformations. Since they correspond to a particle transforming under space rotations as a vector, they should be realized as spin-1 particles, like the photon. Since σ_i is there, Eq. 6.19 is a 2×2 matrix equation.

We want to find how W_i^μ changes under a gauge transformation, since we have no previous answer analogous to $A^\mu \to A^\mu + \partial^\mu\theta(x)$ in the Abelian case. We start with the basic physics requirement that

$$D^{\mu\prime}\psi' = e^{i\vec{\epsilon}(x)\cdot\frac{\vec{\sigma}}{2}}D^\mu\psi, \tag{6.20}$$

since ψ itself transforms that way. Assume that W_i^μ transforms so that

$$W_i^{\prime\mu} = W_i^\mu + \delta W_i^\mu, \tag{6.21}$$

and we want to solve for δW_i^μ. The derivation is quite similar to the steps followed in Eq. 6.5 through Eq. 6.8, and leads to

$$\sigma_i \delta W_i^\mu = \frac{1}{g_2}(\partial^\mu \epsilon_i)\sigma_i + \frac{i}{2}\epsilon_i W_j^\mu[\sigma_i\sigma_j - \sigma_j\sigma_i]. \tag{6.22}$$

Recognizing the commutator as $2i\varepsilon_{ijk}\sigma_k$, this becomes

$$\sigma_i \left(\delta W_i^\mu - \frac{1}{g_2}\partial^\mu \epsilon_i - \varepsilon_{ijk}\epsilon_j W_k^\mu \right) = 0, \tag{6.23}$$

so we can conclude

$$\delta W_i^\mu = \frac{1}{g_2}\partial^\mu \epsilon_i - \varepsilon_{ijk}\epsilon_j W_k^\mu. \tag{6.24}$$

We will not use Eq. 6.24 in further derivations, though it would be useful in a more advanced treatment. For us, it demonstrates how a fully gauge invariant non-Abelian theory can be constructed.

In Eq. 6.19 the covariant derivative is written with the understanding that it will act on the doublet representation of $SU(2)$. That is appropriate for us as we will put left-handed fermions in such doublets. We have implicitly noted that by labeling the couplig g_2. Two generalizations are necessary.

First, though still in the internal $SU(2)$ weak isospin space, D^μ could act on a state in a different representation. If ψ is a state of weak isospin t with $2t+1$ components, let $\vec{T}$ be the $(2t+1) \times (2t+1)$ matrix operator representation in that basis. Then

$$D^\mu = \partial^\mu - ig_2 \vec{T} \cdot \vec{W^\mu}. \tag{6.25}$$

For spin-$\frac{1}{2}$, $\vec{T} = \frac{\vec{\sigma}}{2}$. We will interchangeably write $\vec{T} \cdot \vec{W^\mu}$ or $T_i W_i^\mu$, where summation over $i = 1, 2, 3$ is implied in the latter.

Second, we could consider a different internal space, where the interactions are invariant under another set of transformations. For a $SU(n)$ invariance, with group generators $\vec{F}$ in an (n^2-1)-dimensional space, and $[F_i, F_j] = ic_{ijk}F_k$, the appropriate D^μ to act on the n-dimensional matter state ψ is

$$D^\mu = \partial^\mu - ig_n \vec{F} \cdot \vec{G^\mu}, \tag{6.26}$$

where the G_μ are the (n^2-1) gauge bosons that must be introduced to have a gauge-invariant theory. We will interchangeably write $\vec{F} \cdot \vec{G^\mu}$ or $F_a G_a^\mu$, where summation over $a = 1, 2, \cdots 8$ is implied in the latter for $SU(3)$.

Apparently nature also knows about a $SU(3)$ internal space, which is called the "color" space as we have already mentioned, as well as the $SU(2)$ isospin space. The appropriate generators are the Gell-Mann matrices λ_i described in Section 2.3.

By adding several terms to ∂^μ we can guarantee that we obtain a covariant derivative D^μ that will allow us to write Lagrangians (and therefore equations) that are invariant under gauge transformations, simultaneously or separately, in all the internal spaces. the full covariant derivative that we are presently aware of can be written as

$$D^\mu = \partial^\mu - ig_1 \frac{Y}{2} B^\mu - ig_2 \frac{\sigma_i}{2} W_i^\mu - ig_3 \frac{\lambda_a}{2} G_a^\mu. \tag{6.27}$$

The couplings are arbitrary real numbers. For the Abelian $U(1)$ symmetry we have written the field that must be introduced as B^μ rather than as the electromagnetic field A^μ, since we do not know ahead of time that nature's $U(1)$ invariance corresponds precisely to electromagnetism. We will use physics arguments to make this association later. The $U(1)$ term has been written with a *generator* Y in a form analogous to the other terms. For $U(1)$, Y is just a number, though it can depend on the states on which D^μ operates. Y is called the $U(1)$ *hypercharge* generator.

It is worth emphasizing that for the non-Abelian transformations, once the g_i are fixed for any representation, they are known for all representations. For example, measuring g_2 with muon decay fixes it for quark interactions. Once the coupling of W or g to one fermion is measured, their coupling to all fermions and guage bosons is known.

The ∂^μ is a Lorentz 4-vector, as are all the terms in Eq. 6.27. The first two terms are singlets (i.e. they multiply the unit matrix) in the $SU(2)$ and $SU(3)$ spaces. The third term is a 2×2 matrix in $SU(2)$ and a singlet in the other spaces. The fourth term is a 3×3 matrix in $SU(3)$ and a singlet in the other spaces. There is no inconsistency in having different size matrices for different terms as they operate in different spaces.

Equation 6.27 is, in a sense, the main equation of the Standard Model. When used in a Lagrangian, it leads to the full theory of the SM. It is the culmination of several decades of creative thinking by a number of physicists, leading to the realization that the phase invariance of quantum theory must exist for transformations in new kinds of internal spaces, and that quarks and leptons apparently carry labels that distinguish among three internal spaces. The phase, or gauge, invariance is guaranteed by the form of D^μ, as we learned in Section 6.1. In each case, as in the discussion of gauge invariance for electromagnetism, additional spin-1 gauge boson fields B^μ, W_i^μ, and G_a^μ must exist (1, 3, and 8 respectively). All of these have been observed experimentally.