# Non-Abelian gauge theory

In an earlier article, "What is a gauge theory?," we discussed Abelian gauge theory. Having provided the basics of the non-Abelian group in "$SU(2)$ Lie group and Lie algebra" and "Representations of the $SU(2)$ Lie algebra," in this article we discuss non-Abelian gauge theory, which is also called Yang-Mills theory.

Consider the following gauge transformation

$$\psi \to U\psi \tag{1}$$

where $U = e^{ig\theta^a T^a}$ (Einstein summation convention used)and $\psi$ are matrices, not single numbers (i.e. $1 \times 1$ matrix) as was the case in Abelian gauge theory.

What we want to have is the following:

$$D_\mu \psi \to U D_\mu \psi \tag{2}$$

Following the same logic as before, it goes without saying that we won't have the above transformation if $D_\mu \psi$ were simply $\partial_\mu \psi$. Therefore, we define the covariant derivative as follows:

$$D_\mu \psi = (\partial_\mu - igA_\mu)\psi \tag{3}$$

where $A_\mu$ is a Lie algebra-valued vector. i.e.

$$A_\mu = A_\mu^a T^a \tag{4}$$

where the Einstein summation convention for the repeated index is used.

Now, it can be easily shown that (2) is satisfied, if the following is satisfied:

$$A_\mu \to U A_\mu U^{-1} - \frac{i}{g}(\partial_\mu U)U^{-1} \tag{5}$$

Let's check this.

$$\begin{aligned} D_\mu \psi &\to \partial_\mu (U\psi) - ig(UA_\mu U^{-1} - \frac{i}{g}(\partial_\mu U)U^{-1})(U\psi) \\ &= (\partial_\mu U)\psi + U\partial_\mu \psi - igUA_\mu \psi - (\partial_\mu U)\psi \\ &= UD_\mu \psi \end{aligned} \tag{6}$$

Now observe the following:

$$D_\mu D_\nu \psi \to U D_\mu D_\nu \psi \tag{7}$$

Therefore, we can see that

$$[D_\mu, D_\nu]\psi \to U[D_\mu, D_\nu]\psi \tag{8}$$

which implies

$$[D_\mu, D_\nu] \to U[D_\mu, D_\nu]U^{-1} \tag{9}$$

since the above equation and (1) imply (8). Therefore, it seems that $[D_\mu, D_\nu]$ is a somewhat meaningful quantity, since its behavior under the gauge transformation is not as complicated as that in (5). We will see the meaning soon. So, let's calculate this:

$$\begin{aligned} D_\mu D_\nu \psi &= (\partial_\mu - igA_\mu)(\partial_\nu \psi - igA_\nu \psi) \\ &= \partial_\mu \partial_\nu \psi - igA_\mu \partial_\nu \psi - ig\partial_\mu(A_\nu \psi) - g^2 A_\mu A_\nu \psi \\ &= \partial_\mu \partial_\nu \psi - igA_\mu \partial_\nu \psi - igA_\nu \partial_\mu \psi - ig(\partial_\mu A_\nu - igA_\mu A_\nu)\psi \end{aligned} \tag{10}$$

Therefore we conclude:

$$\begin{aligned} [D_\mu, D_\nu] &= -ig(\partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]) \\ &= -ig(\partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu^a T^a, A_\nu^b T^b]) \\ &= -ig(\partial_\mu A_\nu - \partial_\nu A_\mu + gA_\mu^a A_\nu^b f^{abc} T^c) \\ &\equiv -igF_{\mu\nu} \end{aligned} \tag{11}$$

$F_{\mu\nu}$ is called the field strength. From (9), we see that:

$$F_{\mu\nu} \to UF_{\mu\nu}U^{-1} \tag{12}$$

Notice that in Abelian gauge theory, we have:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{13}$$

$$F_{\mu\nu} \to UF_{\mu\nu}U^{-1} = UU^{-1}F_{\mu\nu} = F_{\mu\nu} \tag{14}$$

which turns out to be the electromagnetic field. In fact, this is expected. Recall that $A_\mu$ was given by $(-\phi, A_x, A_y, A_z)$ in our earlier article "What is a gauge theory?" Therefore, (13) means:

$$\begin{aligned} \vec{E} &= -\nabla\phi - \frac{\partial \vec{A}}{\partial t} \\ \vec{B} &= \nabla \times \vec{A} \end{aligned} \tag{15}$$

if

$$\vec{E} = (F_{10} = -F_{01}, F_{20} = -F_{02}, F_{30} = -F_{03}) \tag{16}$$

$$\vec{B} = (F_{23} = -F_{32}, F_{31} = -F_{13}, F_{12} = -F_{21}) \tag{17}$$

Now, assume you are God, and you want to write a Lagrangian that involves $F_{\mu\nu}$. What choice do you have? (In the discussion following we assume that you are familiar with tensors.

If you aren't, please read the first six sections of my article "A short introduction to general relativity"). Surely, the Lagrangian must be a scalar quantity, so it should not have any free indices (i.e. indices that are not dummy.) This implies the following: Since $F_{\mu\nu}$ has two lower indices, we should contract it with a tensor with two upper indices. What tensors have two upper indices? What first comes into my mind is the metric tensor. However, it turns out that if you contract $F_{\mu\nu}$ with $g^{\mu\nu}$, you have some sort of variant of the Einstein-Hilbert action. Since the case at hand is Maxwell theory and not General Relativity, we should cross this out. Another choice you can have for a tensor with two upper indices is $F^{\mu\nu}$. You can obtain this tensor by raising indices of $F_{\mu\nu}$ with the metric tensor.

Therefore for the action of the Maxwell field, we have:

$$S = -\frac{1}{4}\int d^4x F_{\mu\nu}F^{\mu\nu} \tag{18}$$

where the factor $\frac{1}{4}$ is a convention.

The non-Abelian gauge theory version of this action turns out to be,

$$S = -\frac{1}{2}\int d^4x \mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}) \tag{19}$$

where the factor $\frac{1}{2}$ is a convention.

Let's check whether this action is a good one (i.e. invariant under gauge transformation). Equation (12) implies:

$$F^{\mu\nu} \to UF^{\mu\nu}U^{-1} \tag{20}$$

This in turn implies:

$$F_{\mu\nu}F^{\mu\nu} \to UF_{\mu\nu}F^{\mu\nu}U^{-1} \tag{21}$$

$$\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}) \to \mathrm{Tr}(UF_{\mu\nu}F^{\mu\nu}U^{-1}) = \mathrm{Tr}(U^{-1}UF_{\mu\nu}F^{\mu\nu}) = \mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}) \tag{22}$$

We indeed see that it is good.

So we have established a comparison between Abelian gauge theory and non-Abelian gauge theory (i.e. Yang-Mills theory).

Could we also find the Yang-Mills analog of $A_\mu \to A_\mu + \partial_\mu\theta$? This should be (5), but the comparison is not obvious. To make this more apparent, let's consider an infinitesimal $\theta$. Then, we have:

$$U = 1 + ig\theta^a T^a + O(\theta^2) \tag{23}$$

Plugging this to (5), we get:

$$\begin{aligned} A^a_\mu T^a \quad &\to (1 + ig\theta^b T^b)A^c_\mu T^c(1 - ig\theta^b T^b) - \frac{i}{g}(igT^a\partial_\mu\theta^a) + O(\theta^2) \\ &= A^a_\mu T^a + ig\theta^b A^c_\mu[T^b, T^c] + T^a\partial_\mu\theta^a + O(\theta^2) \end{aligned} \tag{24}$$

Therefore,

$$A^a_\mu \to A^a_\mu + \partial_\mu\theta^a - gf^{abc}\theta^b A^c_\mu \tag{25}$$

In this case, the field strength transforms as

$$\begin{aligned} F^a_{\mu\nu}T^a \quad &\to (1+ig\theta^b T^b)F^c_{\mu\nu}T^c(1-ig\theta^b T^b)+O(\theta^2) \\ &= F^a_{\mu\nu}T^a + ig\theta^b F^c_{\mu\nu}[T^b,T^c]+O(\theta^2) \end{aligned} \tag{26}$$

$$F^a_{\mu\nu} \to F^a_{\mu\nu} - gf^{abc}\theta^b F^c_{\mu\nu} \tag{27}$$

Notice that the Abelian analogs of (25) and (27) are $A_\mu \to A_\mu + \partial_\mu\theta$ and $F_{\mu\nu} \to F_{\mu\nu}$ as $f$ vanishes.

We can also obtain certain equations for the field strength as follows. By the Jacobi identity, we have:

$$\begin{aligned} &[D_\mu,[D_\nu,D_\lambda]]\psi + [D_\lambda,[D_\mu,D_\nu]]\psi + [D_\nu,[D_\lambda,D_\mu]]\psi = 0 \\ &D_\mu(F_{\nu\lambda}\psi) - F_{\nu\lambda}(D_\mu\psi) + \cdots = 0 \\ &(D_\mu F_{\nu\lambda})\psi + \cdots = 0 \\ &\quad D_\mu F_{\nu\lambda} + D_\lambda F_{\mu\nu} + D_\nu F_{\lambda\mu} = 0 \end{aligned} \tag{28}$$

Equation (28) is called the Bianchi identity. Its Abelian gauge theory (i.e. Maxwell theory) analog is the following:

$$\partial_\mu F_{\nu\lambda} + \partial_\lambda F_{\mu\nu} + \partial_\nu F_{\lambda\mu} = 0 \tag{29}$$

In our earlier article on general relativity, we saw the general relativity version of the Bianchi identity.

We can also obtain the equation of motion from the action (19). Varying it, we get:

$$0 \quad = \quad \delta S = -\int d^4x \mathrm{Tr}(\delta F_{\mu\nu}F^{\mu\nu}) = -\int d^4x \mathrm{Tr}\left((D_\mu\delta A_\nu - D_\nu\delta A_\mu)F^{\mu\nu}\right) \tag{30}$$

$$\quad = \quad -2\int d^4x \mathrm{Tr}\left(D_\mu\delta A_\nu F^{\mu\nu}\right) = 2\int d^4x \mathrm{Tr}\left(\delta A_\nu D_\mu F^{\mu\nu}\right) \tag{31}$$

$$0 \quad = \quad D_\mu F^{\mu\nu} \tag{32}$$

where we have used the anti-symmetricity of $F_{\mu\nu}$ and integration by part.

In later articles, we will show that Maxwell's equations, as well as all the formalisms in non-Abelian gauge theory, can be re-written succinctly and lucidly in the language of "differential forms." In particular, we will see that the fact that the field strength (e.g. the electromagnetic field in the Abelian case) is anti-symmetric (i.e. $F_{\mu\nu} = -F_{\nu\mu}$) makes this possible.

# Summary

- Under the following gauge transformation

$$\psi \to U\psi$$

the covariant derivative transforms covariantly as

$$D_\mu\psi \to UD_\mu\psi$$

. If $U$ is a non-Abelian group, we call it "non-Abelian gauge theory" or "Yang-Mills theory."

- The covariant derivative is defined by

$$D_\mu \psi = (\partial_\mu - igA_\mu)\psi$$

just as was the case in Abelian gauge theory, but now $A_\mu$ is a Lie-algebra valued, i.e.,

$$A_\mu = A_\mu^a T^a$$

- The commutator of two covariant derivatives gives the field strength, i.e.,

$$[D_\mu, D_\nu] \sim F_{\mu\nu}$$

- Under a gauge transformation, the field strength transforms as

$$F_{\mu\nu} \to U F_{\mu\nu} U^{-1}$$

- The action of non-Abelian gauge theory is given by

$$S = -\frac{1}{2}\int d^4x \mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}) \tag{33}$$

- From this we can obtain the equation of motion

$$D_\mu F^{\mu\nu} = 0$$

- From the Jacobi identity, we can obtain the Bianchi identity, i.e.,

$$D_\mu F_{\nu\lambda} + D_\lambda F_{\mu\nu} + D_\nu F_{\lambda\mu} = 0$$