June 6

Yang-Mills Theory

In this course, we have surveyed rather thoroughly the principles and applications of the theory of electromagnetism. But electromagnetism is not the only force of Nature. In fact it is only one of four ~~that~~ are now well characterized — and there may be still more forces that we have not yet uncovered. Gravity has its own subtleties and deserves a whole course in its own right. But the other two forces, the strong and weak interactions of elementary particle physics, turn out to be based on equations that are generalizations of Maxwell's equations. In this final lecture of the course, I would like to explain how to construct generalized Maxwell equations that are applicable to these other forces of Nature.

Earlier in 122, we derived Maxwell's equations in a strange but powerful way. We took as fundamental the postulate that a basic complex valued field — for example, the Schrödinger wavefunction of an electron — had a phase that could be freely redefined at each point in space. In other words, we took

$$\psi(x) \to e^{i\alpha(x)}\psi(x) \qquad \text{"local gauge invariance"}$$

as a fundamental symmetry of Nature. We then tried to construct a Lagrangian for $\psi(x)$ invariant under this symmetry.

$$\int d^4x \; \psi^* \psi(x)$$

is invariant, but $\int d^4x \; \vec{\nabla}\psi^* \cdot \vec{\nabla}\psi$ is not,

because $\vec{\nabla}\psi(x) \;\rightarrow\; \vec{\nabla} e^{i\alpha(x)}\psi = e^{i\alpha(x)}(\vec{\nabla}\psi + i\vec{\nabla}\alpha\,\psi)$

To write derivative terms in the Lagrangian, we must define a <u>covariant</u> <u>derivative</u> $\vec{D}$ s.t.

$$\vec{D}\psi \;\rightarrow\; e^{i\alpha(x)}\,\vec{D}\psi(x)$$

We can do this by defining (now I will switch to 4-vector notation)

$$D_\mu = (\partial_\mu + i a_\mu)$$

where a_μ has the transformation law

$$a_\mu \rightarrow a'_\mu = a_\mu - \partial_\mu \alpha$$

then

$$D_\mu\psi \rightarrow (\partial_\mu + i a_\mu - i\partial_\mu\alpha)\, e^{i\alpha(x)}\psi$$

$$= e^{i\alpha(x)}(\partial_\mu + i a_\mu)\psi + i\partial_\mu\alpha\,\psi - i\partial_\mu\alpha\,\psi$$

$$= e^{i\alpha(x)}\, D_\mu\psi \quad \text{as required}$$

then terms like $\int d^4x \; D_\mu\psi^\dagger D^\mu\psi$ can be used to construct invariant Lagrangians. We can also build a Lagrangian for a_μ. Since this object must be gauge-invariant, it must be built from

$$F_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$$

a suitable Lorentz-invariant term is $\int d^4x \, (F_{\mu\nu})^2$. To give the correct coupling of electromagnetism to the Schrödinger equation as discussed at the end of 120, we identify

$$a_\mu = \frac{e}{\hbar} A_\mu$$

where A_μ is the electromagnetic vector potential. Then the covariant derivative is

$$D_\mu = \left(\partial_\mu + i \frac{e}{\hbar} A_\mu \right)$$

and the Lagrangian of the A field – correctly normalized – is

$$\int d^4x \, \left(-\frac{1}{4\mu_0 c} \left(F_{\mu\nu}\right)^2 \right)$$

where $F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu)$

In 1954, Yang and Mills showed that this argument generalizes in a profound way. If $\psi(x)$ is viewed as a set of two real numbers, the operation $\psi \to e^{i\alpha}\psi$

is a rotation in 2-dimensional space. Yang and Mills showed how to generalize the construction from this case to that of a general compact continuous symmetry group G. I will now describe a little of the structure of such a group, and then we will review their construction.

To understand the structure of a general symmetry group, let's first look at the group of rotations in 3 dimensions. Three particular sets of elements of this group are

$$\text{rotations about } \hat{1} \quad R_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$$

$$\text{about } \hat{2} \quad R_2(\theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}$$

$$\text{about } \hat{3} \quad R_3(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

It is useful to look at the infinitesimal form of the rotation about $\hat{1}$

$$R_1(\theta) \cong \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \theta \\ 0 & -\theta & 0 \end{pmatrix} + \cdots$$

$$= \underline{1} + i\,\theta\, T_1 + \cdots$$

$$\text{where} \quad T_1 = \begin{pmatrix} 0 & & \\ & 0 & -i \\ & i & 0 \end{pmatrix}$$

similarly

$$T_2 = \begin{pmatrix} & & i \\ & 0 & \\ -i & & \end{pmatrix}$$

$$T_3 = \begin{pmatrix} 0 & -i & \\ i & 0 & \\ & & 0 \end{pmatrix}$$

describe infinitesimal rotations about $\hat{2}$ and $\hat{3}$

a general infinitesimal rotation about an arbitrary axis is

$$R \cong 1 + i\alpha^a T_a$$

where $|\vec{\alpha}|$ is the small angle and $\hat{\alpha}$ is the axis of rotation.

Notice that the T_a have a simple form:

$$(T_a)_{kl} = i\,\epsilon^{kal}$$

It is also interesting to compute the commutator of two T's:

$$[T_1, T_2] = (T_1 T_2 - T_2 T_1)$$

$$= \begin{pmatrix} & & \\ & & -i \\ & i & \end{pmatrix}\begin{pmatrix} & & i \\ & & \\ -i & & \end{pmatrix} - \begin{pmatrix} & & i \\ & & \\ -i & & \end{pmatrix}\begin{pmatrix} & & \\ & & -i \\ & i & \end{pmatrix}$$

$$= \begin{pmatrix} & & \\ -1 & & \\ & & \end{pmatrix} - \begin{pmatrix} & -1 & \\ & & \\ & & \end{pmatrix}$$

$$= \begin{pmatrix} & 1 & \\ -1 & & \\ & & \end{pmatrix} = i\,T_3$$

and similarly

$$[T_a, T_b] = i\,\epsilon^{abc}\,T_c$$

It turns out that this structure generalizes to an arbitrary continuous symmetry group. An infinitesimal element can be written as

$$V = 1 + i\alpha^a T_a + \cdots$$

where, for a compact group, the α^a are real parameters and the T_a are Hermitian matrices. The finite transformation is

$$V = e^{i\alpha^a T_a} \quad \left(\text{the solution of } \frac{d}{d\lambda} V = i\alpha^a T_a V \text{ at } \lambda = 1\right)$$

note that

$$V^\dagger = e^{-i\alpha^a T_a} = V^{-1}$$

so V is unitary. The multiplication law of V's is entirely determined by the commutation relation of the T_a's, which I will write as

$$[T_a, T_b] = i f^{abc} T_c$$

It is possible to choose the T's so that

$$\operatorname{tr} T_a T_b = (\text{const})\, \delta_{ab} \qquad (\text{const}) > 0$$

and so that f^{abc} is totally antisymmetric. The commutator above is also called the Lie algebra. A simple Lie algebra is one in which there is no subset of the T's that commutes with another subset. In particular, we can exclude $T = c.\underline{1}$, which commutes with everything, by insisting that $\operatorname{tr} T_a = 0$

Not every antisymmetric structure f^{abc} defines the Lie algebra of a group. The possible f^{abc} ("structure constants") were classified in the late 19^{th} century by Cartan and Killing. It turns out that there are three infinite families of possibilities for simple Lie algebras:

$SU(n)$ $\quad V \in \{ n\times n$ unitary matrices, not including $V = e^{i\alpha}\cdot \underline{1} \}$

these preserve $\xi^\dagger \xi$

$$(\xi^\dagger\xi \to (V\xi)^\dagger V\xi = \xi^\dagger V^\dagger V \xi = \xi^\dagger \xi)$$

$SO(n)$ $\quad n\times n$ unitary matrices which also preserve $\vec{\xi}\cdot\vec{\xi}$ $\quad$ (= rotations in n dimensions)

$Sp(n)$ $\quad n\times n$ unitary matrices which preserve an antisymmetric inner product $\xi^T E \xi$ $\quad$ (n = even only)

plus five more "exceptional" cases called

G_2, F_4, E_6, E_7, E_8

The largest exceptional group E_8 has nontrivial applications to coding theory, the theory of finite groups, and string theory.

For any consistent set of structure constants, the matrices

$$(T^a)_{mn} = i f^{man}$$

satisfy the commutation relations $[T^a, T^b] = i f^{abc} T^c$

This is called the "adjoint representation" of the commutation relations, and it gives the transformation law of the T_a themselves.

Now consider a multicomponent field ϕ e.g.

$\phi = \begin{pmatrix} \text{neutrino} \\ \text{electron} \end{pmatrix}$ acted on by the group G. The infinitesimal action is

$$\phi \to \phi' = (1 + i\alpha^a T_a)\phi$$

Note ~~that~~ $\phi^\dagger\phi$ is a group-invariant, since

$$\phi^\dagger\phi \to \phi^\dagger(1 - i\alpha^a T_a + \cdots)(1 + i\alpha^a T_a + \cdots)\phi$$

$$= \phi^\dagger(1 - i\alpha^a T_a + i\alpha^a T_a + \cdots)\phi$$

$$= \phi^\dagger\phi + O(\alpha^2)$$

so that $\frac{\partial}{\partial\alpha^a}\phi^\dagger\phi = 0$. Because we can build up finite group elements from infinitesimal ones, a quantity invariant to order α is totally invariant.

Now $\partial_\mu\phi$ is not invariant to local group transformations:

$$\partial_\mu\phi \to \partial_\mu(1 + i\alpha_a(x)T_a)\phi$$

$$= (1 + i\alpha_a T_a)\partial_\mu\phi + i(\partial_\mu\alpha_a)T_a\phi$$

To write a Lagrangian for ϕ which is locally invariant to G, we need to define a covariant derivative D_μ so that

$$D_\mu\phi \to (1 + i\alpha^a(x)T_a)D_\mu\phi$$

To do this, write

$$D_\mu = \left(\partial_\mu - i\frac{g}{\hbar}A_\mu^a T^a\right)$$

one new vector field for each T_a!

and give A_μ^a the transformation law

$$A_\mu^a \to A_\mu^{a\prime} = A_\mu^a + \delta A_\mu^a$$

where δA_μ^a is to be determined. Then

$$D_\mu\phi \to \left(\partial_\mu - i\frac{g}{\hbar}A_\mu^a T^a - i\frac{g}{\hbar}\delta A_\mu^a T_a\right)(1 + i\alpha^b(x)T^b)\phi$$

$$= (1 + i\alpha^b T^b)\left(\partial_\mu - i\frac{g}{\hbar}A_\mu^a T^a\right)\phi$$

$$+ i\,\partial_\mu\alpha^b(x)\,T_b\,\phi$$

$$+ \frac{g}{\hbar}\left(A_\mu^a T^a\,\alpha^b T^b - \alpha^b T^b\,A_\mu^a T^a\right)\phi$$

$$- i\frac{g}{\hbar}\,\delta A_\mu^a T^a\phi$$

$$= (1 + i\alpha^b T_b)\,D_\mu\phi \quad \text{if}$$

$$\delta A_\mu^a T^a = \frac{\hbar}{g}\,\partial_\mu\alpha^a T_a - i\left[A_\mu^a T^a, \alpha^b T^b\right]$$

Using the Lie algebra of G, this becomes

$$\delta A_\mu^a T^a = \frac{\hbar}{g}\left[\partial_\mu \alpha^c + \frac{g}{\hbar} f^{abc} A_\mu^a \alpha^b\right] T_c$$

so

$$\delta A_\mu^c = \frac{\hbar}{g}\left[\partial_\mu \alpha^c + \frac{g}{\hbar} f^{abc} A_\mu^a \alpha^b\right]$$

note that this expression is actually

$$= \frac{\hbar}{g} D_\mu \alpha^c = \frac{\hbar}{g}\left(\partial_\mu - ig A_\mu^a T_a\right)\alpha$$

with $(T^a)_{bc} = i f^{bac}$ in the adjoint representation.

Now

$$\int d^4x \; D_\mu \phi^\dagger D^\mu \phi$$

is an invariant Lagrangian. Can we also write an invariant Lagrangian for A_μ^a alone. Notice that our previous trick of writing

$$\partial_\mu A_\nu^a - \partial_\nu A_\mu^a$$

is too simple for the new transformation law of A_μ^a. But we can generalize this trick as follows: Note that

$$D_\mu \rightarrow (1 + i\alpha^a T_a) D_\mu (1 - i\alpha^b T_b)$$

then $$D_\mu D_\nu \rightarrow (1 + i\alpha^a T_a) D_\mu D_\nu (1 - i\alpha^b T_b)$$

and

$$[D_\mu, D_\nu] \longrightarrow (1 + i\alpha^a T_a)\,[D_\mu D_\nu]\,(1 - i\alpha^b T_b)$$

but $[D_\mu D_\nu]$ actually has no derivatives!

$$D_\mu D_\nu = \left(\partial_\mu - i\frac{g}{\hbar} A^a_\mu T_a\right)\left(\partial_\nu - i\frac{g}{\hbar} A^b_\nu T_b\right) - (\mu \leftrightarrow \nu)$$

$$= -i\frac{g}{\hbar}(\partial_\mu A^b_\nu) T_b + \left(-i\frac{g}{\hbar}\right)^2 A^a_\mu T^a A^b_\nu T^b$$

$$- (\mu \leftrightarrow \nu)$$

$$= -i\frac{g}{\hbar}\left[(\partial_\mu A^b_\nu - \partial_\nu A^b_\mu) T^b - i\frac{g}{\hbar}\left[A^a_\mu T^a, A^b_\nu T^b\right]\right]$$

$$= -i\frac{g}{\hbar}\left[\partial_\mu A^c_\nu - \partial_\nu A^c_\mu + \frac{g}{\hbar} f^{abc} A^a_\mu A^b_\nu\right] T^c$$

$$= -i\frac{g}{\hbar} F^c_{\mu\nu} T^c$$

where

$$F^c_{\mu\nu} = \partial_\mu A^c_\nu - \partial_\nu A^c_\mu + \frac{g}{\hbar} f^{abc} A^a_\mu A^b_\nu$$

generalizes the electromagnetic field strength. This object is nonlinear in A, and the nonlinearity is directly related to the geometrical structure of the symmetry group G!

From the relation at the top of the page, $F^a_{\mu\nu}$

transforms as

$$F_{\mu\nu}^{a}T^{a} \rightarrow (1+i\alpha^{a}T_{a})\, F_{\mu\nu}^{c}T^{c}\,(1-i\alpha^{b}T_{b})$$

$$= F_{\mu\nu}^{a}T^{a} + i\,[\alpha^{a}T^{a}, F_{\mu\nu}^{b}T_{b}]$$

$$= F_{\mu\nu}^{a}T^{a} - f^{abc}\alpha^{a}F_{\mu\nu}^{b}T^{c}$$

or

$$F_{\mu\nu}^{a} \rightarrow F_{\mu\nu}^{a} - f^{abc}\alpha^{b}F_{\mu\nu}^{c}$$

Since f^{abc} is totally antisymmetric

$$(F_{\mu\nu}^{a})^{2} \rightarrow (F_{\mu\nu}^{a})^{2} + (-2f^{abc})\,\alpha^{b}F_{\mu\nu}^{a}F_{\mu\nu}^{c}$$

$$= (F_{\mu\nu}^{a})^{2} + \mathcal{O}(\alpha^{2})$$

so $(F_{\mu\nu}^{a})^{2}$ is invariant and we can write the Lagrangian for A_{μ}^{a}

$$\int \mathcal{L} = \int d^{4}x \left(-\frac{1}{4\mu_{0}c}(F_{\mu\nu}^{a})^{2}\right)$$

that generalizes electromagnetism. You can show that the equation of motion that follows from this Lagrangian is:

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

a nonlinear generalization of Maxwell's equations.

So, how do we use Yang-Mills theory in physics? Let's begin with the strong interactions. The basic particles of the strong interactions are quarks, the constituents of protons, neutrons, pions, etc. Each type of quark comes in 3 distinct states, called "colors"
$$q = \begin{pmatrix} q_r \\ q_b \\ q_g \end{pmatrix}$$

An SU(3) group acts on the color labels. If we make color rotations a local gauge symmetry

$$q \to (1 + i\alpha^a \underset{\substack{\uparrow \\ 3\times 3 \text{ Hermitian}}}{T_a}) \begin{pmatrix} q_r \\ q_b \\ q_g \end{pmatrix}$$

we obtain an SU(3) Yang-Mills theory. The vector particles of this theory, called "gluons", are the carriers of the strong interaction, the analogues of photons. Gluons are not so easy to observe at GeV energies, because of the strong nonlinearity of the theory. However, at very high energies (10 GeV and above) gluons are radiated just like photons, with a Weizsäcker-Williams distribution. In e^+e^- annihilation experiments at 90 GeV, one clearly observes the process

$$e^+e^- \to q\bar{q} \to q\bar{q}\,g$$

quark-antiquark + gluon

the gluons are emitted approximately collinearly to the quark and antiquark directions, with the distribution

$$\int dx\ \frac{(\frac{4}{3}\alpha_s)}{2\pi}\ \frac{1+(1-x)^2}{x}\ \log\left(\frac{90\text{ GeV}}{1\text{ GeV}}\right)^2$$

where $\frac{4}{3}\alpha_s$ replaces α and characterizes the strength of the ~~strong~~ interaction at high energy.

For the weak interactions there is a slightly more complex ~~story~~. The weak interactions are short-range forces, mediated by particles that obey the Klein-Gordon equation and produce Yukawa potentials. To construct such particles, we need to generate a term

$$\int d^4x\ A^a_\mu\, m^2\, A^a_\mu$$

in the Lagrangian. But this seems to be forbidden by local gauge symmetry.

The solution to this problem is to assume that, for some reason, some field in Nature does not respect the basic symmetry of the problem. I'll demonstrate how this can work in a model called the Georgi–Glashow model, which is not a correct theory of weak interactions but is a simpler model that demonstrates the basic principles. This model has an SO(3) gauge symmetry (3-d rotations) and an extra field $\vec{\Phi}$ which is a 3-vector under these rotations. The kinetic term of $\vec{\Phi}$ is

$$\int d^4x \ \tfrac{1}{2} D_\mu \vec{\Phi} \cdot D^\mu \vec{\Phi}$$

$$= \int d^4x \ \tfrac{1}{2} \left(\vec{\Phi} \, (-D_\mu D^\mu) \vec{\Phi} \right)$$

We now assume that $\vec{\Phi}$ sits down, in its lowest-energy state, in a particular direction in the 3-d. space:

$$\vec{\Phi} = v \, \hat{3}$$

Then

$$\tfrac{1}{2} \vec{\Phi} \, (-D_\mu D^\mu) \vec{\Phi} = -\tfrac{1}{2} v^2 \ \hat{3} \left(\partial_\mu - i \tfrac{g}{\hbar} A_\mu^a T^a \right) \left(\partial_\mu - i \tfrac{g}{\hbar} A_\mu^a T^a \right) \hat{3}$$

$$= \tfrac{1}{2} v^2 \left(\tfrac{g}{\hbar} \right)^2 \left[\hat{3} \, T^a T^b \, \hat{3} \right] A_\mu^a A^{\mu b}$$

Now, this is $SO(3)$ so

$$(T^a)_{ij} = i\,\epsilon^{iaj}$$

Then

$$\hat{3}\, T^a T^b\, \hat{3} = -\,\epsilon^{3aj}\epsilon^{jb3}$$

$$= \delta^{ab}\delta^{33} - \delta^{a3}\delta^{b3}$$

So

$$\tfrac{1}{2}\,\vec{\Phi}\cdot D^\mu D_\mu \vec{\Phi} = \tfrac{1}{2}\left(\frac{g v}{\hbar}\right)^2 \left(A^1_\mu A^{\mu 1} + A^2_\mu A^{\mu 2}\right)$$

This is a remarkable result! A^1_μ and A^2_μ get an m^2 term and lead to Yukawa potentials which can be associated with the weak interactions. A^3_μ remains massless and obeys Maxwell's equations. The covariant derivative involving A^3_μ is

$$\left(\partial_\mu - i\frac{g}{\hbar} A^3_\mu T^3\right)$$

where $T^3 = \begin{pmatrix} & -i & \\ +i & & \\ & & 0 \end{pmatrix}$

this is a Hermitian matrix with eigenvectors

$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\\ 0\end{pmatrix} \to \lambda = +1 \qquad \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\\ 0\end{pmatrix}\ \lambda = -1 \qquad \begin{pmatrix}0\\ 0\\ 1\end{pmatrix}\ \lambda = 0$

In the basis of eigenvectors

$$T^3 = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}$$

If we now set $g = e$, A_μ^3 is a photon field that couples to <u>quantized</u> <u>electric</u> charges. Thus, we obtain a unified theory of weak and electromagnetic interactions, in which electric charge is quantized by virtue of the group structure! The true story in Nature is slightly more complicated, but it shares these basic principles.

The theory of electromagnetism thus provides the foundation for understanding all of the microscopic laws of Nature. But, to go on and understand microphysics more deeply, there is a big ingredient that is still missing — <u>quantum</u> <u>mechanics</u>. That is the next major topic you must study. Enjoy!