Lecture 3: More Special Relativity

In the previous lecture, we introduced special relativity from demanding that the speed of light is the same in every inertial reference frame. This resulted in the notions of space-time intervals, Lorentz transformations, and tensors that transform accordingly. In this lecture, we continue this discussion, but focused around electromagnetism and its Lagrangian formulation.

Maxwell's equations for electric and magnetic fields $\vec{E}$ and $\vec{B}$ in the presence of charge density ρ and current density $\vec{J}$ is:

$$\vec{\nabla}\cdot\vec{E} = \rho\,, \quad \vec{\nabla}\times\vec{B} - \frac{\partial\vec{E}}{\partial t} = \vec{J}\,, \quad \vec{\nabla}\cdot\vec{B} = 0\,, \quad \vec{\nabla}\times\vec{E} + \frac{\partial\vec{B}}{\partial t} = 0\,.$$

For compactness, I've just set $\varepsilon_0 = \mu_0 = c = 1$. This is also called "natural units", but in general, we'll keep factors of speed of light c around. This system of equations is Lorentz invariant but not "obviously" so. We often state that a symmetry is "manifest" if a system is, simply by eye, invariant to some transformation. The Lorentz group acts most naturally on four-vectors and tensors, so let's see if we can express Maxwell's equations in tensor notation.

First, note that the charge density ρ and current density $\vec{J}$ are naturally combined into a four-vector:

$$J = (\rho, \vec{J})$$

That this transforms sensibly under Lorentz transformations I won't explicitly check. However it should make sense. Under a rotation, charge density is just charge density, so ρ is unchanged, while the components of $\vec{J}$ will mix with one another. Under a boost, static charge is now moving, which turns ρ into $\vec{J}$.

The two Maxwell equations that depend on charge distributions, Gauss and Ampère, are actually a system of 4 equations; Because they are differential equations, this suggests that they can be represented as the action of the derivative four-vector on some object: The derivative four-vector ∂ is:

$$\partial = \left(\frac{\partial}{\partial t}, \vec{\nabla}\right) = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

Further, Gauss + Ampère only have derivatives that are in dot products or cross products with $\vec{E}$ or $\vec{B}$. That is, there are no "naked" derivative indices.

These observations motivate expressing two of Maxwell's equations as:

$$\partial_\mu F^{\mu\nu} = J^\nu$$

where $F^{\mu\nu}$ is some two-index tensor. We can construct $F^{\mu\nu}$ explicitly by matching with Maxwell. For example, let's set $\nu=0$ and see what we get. Recall that

$J^0 = \rho$ and so

$$\partial_\mu F^{\mu 0} = -\partial_0 F^{00} + \partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30} = J^0 = \rho$$

Therefore this corresponds to Gauss's Law if we set $F^{00}=0$ and $F^{10}=E_x$, $F^{20}=E_y$, $F^{30}=E_z$.

Then,

$$\partial_\mu F^{\mu 0} = \vec{\nabla}\cdot\vec{E} = \rho, \text{ exactly Gauss' Law.}$$

We can identify the other components of $F^{\mu\nu}$ similarly, but I'll just write the result:

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}^{\mu\nu} = -F^{\nu\mu}$$

$F^{\mu\nu}$ is a rank-two antisymmetric tensor. Its indices μ and ν represent its Lorentz transformation properties. We call $F^{\mu\nu}$ the field-strength tensor in electromagnetism, or the curvature tensor in geometry. With $F^{\mu\nu}$, Gauss' and Ampère's laws are:

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

The two other Maxwell equations follow from other considerations of $F^{\mu\nu}$. Maxwell's equations are actually redundant: there are a total of 8 equations (one for each vector component of $\vec{E}$ and $\vec{B}$ and two divergence equations), but there are only a total of 6 components of $\vec{E}$ and $\vec{B}$. This means that there must be a simpler object from which $\vec{E}$ and $\vec{B}$ can be defined. We can introduce a vector potential A_μ, where

$$A_\mu = (V, \vec{A})_\mu, \text{ for a potential } V \text{ and vector } \vec{A}.$$

With this object, we can trivially express some of Maxwell's equations. For example if we define

$\vec{B} = \vec{\nabla} \times \vec{A}$, then it is automatically divergence-free:

$\vec{\nabla} \cdot \vec{B} = \vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) = 0$, because the divergence of a curl is 0. (check in components!) The electric field can be written as: $\vec{E} = -\vec{\nabla} V - \frac{\partial \vec{A}}{\partial t}$

While I won't go through the details, in terms of A_μ, the field strength is very simple:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

This form makes the antisymmetry of $F_{\mu\nu}$ manifest. In geometry A_μ is called the connection and enjoys a transformation under which all observable quanties are invariant. We can perform a gauge transformation of A_μ which just adds an arbitrary gradient to A_μ:

$$A_\mu \to A_\mu + \partial_\mu \lambda, \text{ for any function } \lambda.$$

For example, in terms of $\vec{A}$, this transfomation is

$\vec{A} \to \vec{A} + \vec{\nabla}\lambda$ and the magnetic field transforms as:

$\vec{B} = \vec{\nabla} \times \vec{A} \to \vec{\nabla} \times (\vec{A} + \vec{\nabla}\lambda) = \vec{\nabla} \times \vec{A}$, because the curl of a gradient is 0. The field-strength tensor is gauge invariant:

$$F_{\mu\nu} \to \partial_\mu(A_\nu + \partial_\nu\lambda) - \partial_\nu(A_\mu + \partial_\mu\lambda) = F_{\mu\nu}$$

because partial derivatives commute: $\partial_\mu \partial_\nu = \partial_\nu \partial_\mu$.

A gauge transformation, because it doesn't affect physical quantities, should really be thought of as identifying different vector potentials. That is, if two potentials A_μ and A'_μ are related as:

$$A_\mu = A'_\mu + \partial_\mu \lambda$$

they produce the same physical consequences. So, the cost of unifying electric and magnetic fields is the introduction of classes of an infinite number of potentials that produce the same physics. This is actually a good trade-off. :)

In this notation, the field strength satisfies a relation called the Bianchi identity trivially:

$$\partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} = 0$$

This, when written in components are the final two of Maxwell's equations; $\vec{\nabla} \cdot \vec{B} = 0$ and $\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0$.

This is excellent: we have a covariant description of electromagnetism and can completely express the electric and magnetic fields in terms of a vector potential A_μ. For the rest of this lecture, we will introduce the Lagrangian formalism from which Maxwell's equations can be derived.

From classical mechanics, you are likely familiar with a Lagrangian, L. The Lagrangian is just the difference of kinetic and potential energies:

$$L = K - U.$$

For a particle with position $x(t)$ in a potential $U(x)$, the Lagrangian is: $L = \frac{1}{2}m\dot{x}^2 - U(x)$.

The <u>action</u> is the time integral of the Lagrangian

$$S[x] = \int dt\, L = \int dt \left[\tfrac{1}{2}m\dot{x}^2 - U(x)\right],$$

and is called a functional because it is a function of trajectory x which itself is a function of time. The equations of motion for the particle, Newton's 2^{nd} Law, follows from the <u>principle</u> of <u>least</u> <u>action</u>: physical trajectories minimize (actually, extremize) the action. To find the trajectories x that minimize S, we take a derivative. Now, you can go all Euler-Lagrange equations, but I never remember what goes where. So, let's take a derivative the old-fashioned way:

$$\frac{\delta S}{\delta x} = \lim_{\varepsilon \to 0} \frac{S[x+\varepsilon] - S[x]}{\varepsilon}, \quad \text{for some } \varepsilon \equiv \varepsilon(t).$$

Let's see what this is:

$$S[x+\varepsilon] = \int dt \left[\tfrac{1}{2}m(\dot{x}+\dot{\varepsilon})^2 - U(x+\varepsilon)\right]$$

$$= \int dt \left[\tfrac{1}{2}m\dot{x}^2 + m\dot{\varepsilon}\dot{x} - U(x) - \varepsilon\frac{\partial U}{\partial x}\right] + \mathcal{O}(\varepsilon^2)$$

product derivative rule

$$= S[x] + \int dt \left[\frac{d}{dt}(m\varepsilon\dot{x}) - m\varepsilon\ddot{x} - \varepsilon\frac{\partial U}{\partial x}\right] + \mathcal{O}(\varepsilon^2)$$

Therefore, the derivative of the action with respect to trajectory x is:

$$\frac{\delta S}{\delta x} = \lim_{\varepsilon \to 0} \frac{S[x+\varepsilon]-S[x]}{\varepsilon} = m\dot{x}\Big|_{t_0}^{t_f} - \int dt \left[m\ddot{x} + \frac{\partial u}{\partial x}\right] = 0$$

The first term is just the particle momentum evaluated at the initial and final time. This is just set by the boundary conditions, and not the dynamics of the trajectory, so we ignore it. (If momentum is conserved such a term is just 0.) For what remains to vanish for an arbitrary trajectory $x(t)$, we must demand:

$m\ddot{x} + \frac{\partial u}{\partial x} = 0$ or that $m\ddot{x} = -\frac{\partial u}{\partial x}$, which is just Newton's second law.

So, if we can formulate a Lagrangian and action for E+M, we have a procedure, via the principle of least action, for deriving the equations of motion, Maxwell's equations.

I won't derive the E+M Lagrangian here, but I'll present some plausibility arguments and understand what the consequence of gauge invariance of the action means.

The Lagrangian of E+M is still of the form kinetic minus potential energy. Kinetic energies involve two (time) derivatives. The field strength $F_{\mu\nu}$ is formed from single derivatives of A_μ, so we postulate that we need two $F_{\mu\nu}$'s. Further, the Lagrangian should be Lorentz invariant, so we need to contract all indices. Finally, we need to include the charge four-vector J^μ which couples directly to A_μ. These considerations motivate:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + A_\mu J^\mu .$$

The factor of $-\frac{1}{4}$ is there to correctly reproduce Maxwell's equations. This object is properly called a Lagrangian density because it consists of objects that fill space and time. That is, the vector potential is a function of the full space-time four-vector $x = (t, x, y, z)$:

$A_\mu = A_\mu(x)$. Such an object is called a field.

To construct the action we need to integrate over all of space and time:

$$S[A_\mu] = \int d^4x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + A_\mu J^\mu \right], \text{ where}$$

$d^4x = dt\, dx\, dy\, dz$.

Maxwell's equations then follow from the Bianchi identity and the principle of least action applied on $S[A_\mu]$

Finally, let's consider what gauge invariance means in this formalism. $F^{\mu\nu}$ is just gauge invariant, so the only thing that changes is the coupling to charge. Under a gauge transformation, we have

$$S[A_\mu] \supset \int d^4x\, A_\mu J^\mu \to \int d^4x\, (A_\mu + \partial_\mu \lambda) J^\mu$$

$$= \int d^4x\, A_\mu J^\mu + \int d^4x \left[\partial_\mu(\lambda J^\mu) - \lambda \partial_\mu J^\mu \right]$$

If the current J^μ vanishes at the boundary of spacetime, we can ignore the total derivative term. The action is gauge invariant if and only if: $\partial_\mu J^\mu = 0$,

which is the statement of charge conservation.