Relativity and Electromagnetism

Initial questions: How do different astronomical events look in different frames? How does this relate to cosmology and objects with large gravitational redshifts such as black holes and neutron stars?

We'll wrap up our lectures on fundamental electromagnetism by considering its relation to special relativity. Historically, Einstein was apparently motivated by the transformation properties of the Maxwell equations to put the finishing touches on special relativity. The approaches used in special relativity were then generalized (thus, general relativity) to deal with accelerated reference frames. We won't deal with general relativity per se, but we will use some of the geometrical concepts that give the relativity framework its power.

In a fundamental sense, relativity deals with how quantities transform in different frames. For example, how does the measured length of a bar change from frame to frame? Let's start with scalar quantities, which are the same in all frames and are obviously conserved.

An easy example of a scalar is a number. For example, if in a certain region one observer determines that there are five particles, all observers will agree that there are five particles.

Less clear is the case of vectors. To understand this better, let's consider another situation that is simpler. Suppose we have a plane, and in that plane we define x and y axes. In this plane we place a vector. The vector has a particular x component and a particular y component. If we now consider a different set of axes x' and y' rotated with respect to the first set, then the components of the vector will change. **Ask class:** what quantity about the vector will *not* change? Its length is invariant. If the x and y components of the vector are A^x and A^y, then the squared length is $A^xA_x + A^yA_y$; we're writing one index up and one down for reasons that only really become clear in general relativity. We can write this more compactly by $A^\alpha A_\alpha$, where $\alpha{=}x, y$ and it is implied that we sum over indices (this is called the Einstein summation convention). This quantity is a scalar; it doesn't change when new coordinates are picked. More generally, the dot product $A \cdot B = A^\alpha B_\alpha$ between two vectors is also a scalar. As described in this formula, any expression without any "free indices" (i.e., an expression in which any indices have been summed over, once up and once down) is a scalar.

We also know that if the x', y' system is rotated counterclockwise by an angle θ relative to the x, y system, then we can write the new components in terms of the old components: $A^{x'} = A^x \cos\theta + A^y \sin\theta$, $A^{y'} = -A^x \sin\theta + A^y \cos\theta$. We can then represent this as a matrix:

$$\begin{pmatrix} A^{x'} \\ A^{y'} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} A^x \\ A^y \end{pmatrix} \tag{1}$$

or, using summation notation, $A^{\beta'} = M^{\beta'}_{\ \alpha} A^\alpha$, where M is the matrix. Note that α and β' are both dummy indices, in the sense that any letters could have been used, but one convention in relativity is to use Greek letters to indicate summing over all four indices of spacetime.

What about in three dimensions? Now we define x, y, and z axes. Again, the vector components change. **Ask class:** what remains constant about the vector? The length, again, of course. We can write it as $A^x A_x + A^y A_y + A^z A_z = A^\alpha A_\alpha$. We can also write the new components in terms of the old for a given set of rotations, and could in fact have the same matrix multiplication as before.

In Newtonian physics, the spatial distance between any two quantities at an instant of time is a scalar, i.e., has the same value in any reference system. Similarly, the time between two events is measured to be the same by any two observers. Another way of saying this is that Newtonian mechanics is invariant under Galilean transformations.

However, Maxwell's equations are *not* invariant under Galilean transformations, but rather under the Lorentz transformations, which have the non-intuitive property that they mix time and space. For example, if the coordinates x, y, z, t are measured by one observer, an observer moving in the $+x$ direction at speed v measures

$$\begin{aligned} x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \\ t' &= \gamma(t - vx/c^2)\ . \end{aligned} \tag{2}$$

Here $\gamma = (1 - v^2/c^2)^{-1/2}$. Under these transformations the *interval* $x^2 + y^2 + z^2 - c^2t^2$ remains unchanged, so it is the equivalent of the squared length and is a scalar. However, it means that we can't consider space and time separately any more, but must think of time and space as forming a "four-vector". Let us specifically think of time as the "0-component" of the vector, and x, y, z as the 1,2,3 components. Then, in analogy to the length of vectors in a plane or in three dimensions, we can write the squared length as $A^\alpha A_\alpha$, where our rule is that $A^\alpha = A_\alpha$ unless $\alpha = 0$ (the time component), in which case $A_\alpha = -A^\alpha$. The transformation of a vector (i.e., the knowledge of what its components would be in one frame if you already know them in a different frame) is given by the Lorentz transformation matrix $\Lambda^{\beta'}_{\ \alpha}$, so that $A^{\beta'} = \Lambda^{\beta'}_{\ \alpha} A^\alpha$ and again there is an implied summation over indices that appear once up and once down.

An important four-vector is the four-velocity of a particle. $U^\alpha = (\gamma_u c, \gamma_u \mathbf{u})$, where $\mathbf{u}$ is the normal space velocity with three components and $\gamma_u = (1 - u^2/c^2)^{-1/2}$ where u is the magnitude of $\mathbf{u}$. If we compute the square of the four-velocity we find $U^\alpha U_\alpha = -\gamma_u^2 c^2 + \gamma_u^2 u^2 = -c^2$. This means that the four-velocity has a constant magnitude. In turn, this means that the four-acceleration has to be orthogonal to the four-velocity, $a^\mu U_\mu = 0$.

When we widen our scope to electromagnetism we find a number of examples of useful four-vectors. For example, energy divided by c is the 0 component of a four-vector in which the x, y, z components are the x, y, z components of momentum. For a photon this means that if a photon has frequency ω and wave vector $\mathbf{k}$, the four-vector could be written schematically as $(\omega, c\mathbf{k})$. Thus, we know how the frequency and wave vector transform between different frames. Another example is the electromagnetic potentials. If ϕ is the scalar potential and $\mathbf{A}$ is the vector potential, then the four-potential is $A^\alpha = (\phi, \mathbf{A})$. Yet another example is charge densities and current densities. If ρ is the charge density and $\mathbf{j}$ is the current density, then the four-current is $j^\alpha = (\rho c, \mathbf{j})$.

When we try to apply this to electromagnetic fields, however, we run into a puzzle. Consider the electric field. It has three components, so to turn it into a four-vector we would need one extra component. You might think it would be a simple matter of finding something that acts "time-like" to add to the three components of electric field, but a thought experiment shows you can't do that. Remember that one hint that time must be included with space in a four-vector is that when viewed in a moving frame, time and space are mixed. Now consider a single electron at rest. This clearly produces an electric field, but because it isn't moving there is no magnetic field. However, viewed in a moving frame the electron is moving so there is a current and hence a magnetic field. This means that motion mixes the electric and magnetic fields. But there are a total of six components (three electric, three magnetic), so clearly these can't all fit in a four-vector!

The way out is to use another entity, not a scalar or a vector but a grand generalization called a *tensor*. A tensor, like a scalar or a vector, is a geometric entity that has properties independent of any particular coordinate system, but in a particular coordinate system it can be thought of as a function that gives particular numbers at particular places for specific indices. For example, suppose we have the tensor $C^{\alpha\beta}$. At a particular location and time (i.e., point in spacetime), the xy component is C^{xy}, and so on. As an analogy (although this is *not* technically a tensor), think of an electric field. It has three components. For any of its three components you can take a derivative in any of three directions, so if you define a derivative $C^\alpha{}_\beta = \partial E^\alpha / \partial x^\beta$ you can get components. You could then take the second derivative in any direction to get $D^\alpha{}_{\beta\gamma} = \partial^2 E^\alpha / \partial x^\beta \partial x^\gamma$, and so on. To figure out the components of a tensor in one frame given the components in another frame, in special relativity we simply repeatedly apply the Lorentz transformation:

$$\begin{aligned} A'(=\text{scalar}) &= A \\ B^{\beta'} &= \Lambda^{\beta'}{}_\alpha B^\alpha \\ C^{\alpha'\beta'} &= \Lambda^{\alpha'}{}_\mu \Lambda^{\beta'}{}_\nu C^{\mu\nu} \end{aligned} \tag{3}$$

and so on. In general relativity, where spacetime is "curved" in a geometric sense, there are other terms, but we'll ignore those.

With all that background, we find eventually that the tensor associated with the electromagnetic field is the antisymmetric tensor

$$F_{\mu\nu} \equiv \partial A_\nu/\partial x^\mu - \partial A_\mu/\partial x^\nu \tag{4}$$

or in Cartesian coordinates

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} . \tag{5}$$

Okay. That's all very well, but one reason we got into this was that we wanted to know what quantities remain constant between frames. **Ask class:** how can we form a scalar out of the electromagnetic tensor, with the hint that if there are no free indices then we have a scalar? One way is to form the "square" of the tensor, $F_{\mu\nu}F^{\mu\nu}$, which is $2(B^2 - E^2)$, so $B^2 - E^2$ is invariant. Another one is the determinant, which is a scalar that can be obtained from any matrix. The determinant is $(\mathbf{E}\cdot\mathbf{B})^2$, so $\mathbf{E}\cdot\mathbf{B}$ is invariant. Yet another would be the trace (sum of diagonal elements), $F_\mu{}^\mu$, which is 0 and indeed conserved (but not very interesting!). Suppose that we again return to the case of a single electron that is stationary in a particular frame. **Ask class:** can we transform to a frame in which $\mathbf{E} = \mathbf{B} = 0$? No, because in such a frame $B^2 - E^2 = 0$, whereas in our original frame it was nonzero.

Why do we go to all this trouble? The original form of Maxwell's equations may have its complications, but at least it is relatively familiar and we have an idea of how to do manipulations with them. In contrast, we've now introduced a whole new formalism. The advantage is that in this formalism (which really isn't so bad once you get used to it), there are extremely well defined rules to figure out how observed quantities such as the electric and magnetic fields change under relative motion. In contrast, if you laboriously figure out $\mathbf{E}$ and $\mathbf{B}$ for one reference frame using Maxwell's equations, figuring out what they are in another frame drops you back to step 1 if you have to do it all from scratch. The transformation laws mean, in particular, that for any given situation you can pick the frame in which the original calculation is easiest, then transform away. This is also the key to computations in general relativity: pick an easy (local!) frame in which the answers are as clear as possible, then use the machinery of transformations as necessary.

Let's examine a particular example of how to use transformations. When we computed the Larmor formula for radiation, we restricted ourselves to slow-moving particles. Now we'd like to allow arbitrary relativistic motion. Consider a single particle in arbitrary motion. **Ask class:** at a given instant, what is the most convenient reference frame for this analysis? It's the one in which the particle is instantaneously at rest (but it may still have a nonzero acceleration). In that frame, we have our old formula

$$P = (2q^2/3c^3)|\mathbf{a}|^2 , \tag{6}$$

where **a** is the three-acceleration. Note, however, that in this frame the four-velocity is $U^\alpha = (c, \mathbf{0})$. Since $a^\mu U_\mu = 0$ this must mean that in this frame $a^0 = 0$. We can therefore write $P = (2q^2/3c^3)a^\alpha a_\alpha$. But we sum over all indices, so this is a scalar! The total radiated power is independent of the reference frame, and to compute it we just pick a frame and measure the four-acceleration. Does this make sense? Note that energy and time both transform the same way (they're the timelike components of a four-vector), meaning that since power is energy per time, the changes due to the transformation cancel. One can also use this approach to show that (as expected) the power is beamed in the direction of motion.

Our final task will be to consider how the phase volume and distribution function transform between frames. Consider a group of particles all close together in position and momentum. Their spatial volume is $d^3\mathbf{x}' = dx'dy'dz'$ and their momentum volume is $d^3\mathbf{p}' = dp'_x dp'_y dp'_z$. Their overall phase volume is then $d\mathcal{V} = d^3\mathbf{x}' d^3\mathbf{p}'$. Now consider how this phase volume is measured by an observer moving at some constant speed v (with Lorentz factor $\gamma = (1 - v^2/c^2)^{-1/2}$) with respect to our original frame. We have freedom to rotate our axes, so let's align them so that the motion is in the x direction. In that direction, the length is shortened by a factor γ (so that $dx = \gamma^{-1}dx'$), so $d^3\mathbf{x} = \gamma^{-1}d^3\mathbf{x}'$. However, the x-momentum is *increased* by the factor γ, so that $d^3\mathbf{p} = \gamma d^3\mathbf{p}'$. Therefore the phase volume $d\mathcal{V} = d\mathcal{V}'$; it's an invariant. It follows immediately that the distribution function $f = dN/d\mathcal{V}$ is also invariant, since the number of particles is obviously a scalar.

Using this we can (at long last!) explain why I_ν/ν^3 is constant. Consider the energy density per solid angle per frequency $u_\nu(\Omega)$ we defined earlier. We can also compute this using f, so we have

$$h\nu f p^2 dp\, d\Omega = u_\nu(\Omega) d\Omega\, d\nu \; . \tag{7}$$

Earlier we found $u_\nu(\Omega) = I_\nu/c$, and we know $p = h\nu/c$. Simplifying, we find that $I_\nu/\nu^3 \propto f$. But f is a Lorentz invariant, so we know I_ν/ν^3 is also a Lorentz invariant. The book contains other examples, but let's think of one in particular to sharpen our intuition.

Recommended Rybicki and Lightman problems: 4.1, 4.2, 4.4