MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DEPARTMENT OF PHYSICS
8.033 FALL 2024

LECTURE 11
A COVARIANT FORMULATION OF ELECTROMAGNETICS (PART I)

11.1 Electric and magnetic fields and forces: Background

Our pivot from Galileo's relativity to Einstein's relativity began by considering electrodynamics. Let's write out again the critical equations which govern electrodynamics — the Maxwell equations which connect the the fields to their sources, and the Lorentz force law which shows how these fields act on charges:

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0 \;, \qquad \nabla \cdot \mathbf{B} = 0 \;, \tag{11.1}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \;, \qquad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \;; \tag{11.2}$$

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right) \;. \tag{11.3}$$

It should be emphasized very strongly that these equations are *fully* compatible with special relativity. Indeed, all of the modifications to various physical concepts that Einstein's relativity requires were introduced because it became clear that important aspects of Newtonian mechanics were not compatible with electrodynamics. Electrodynamics is one of the most successfully and accurate theories of nature we have developed. Once it has been updated to account for the fact that our universe is quantum mechanical in nature (a topic for a different course!), we end up with a version of electrodynamics that is perhaps humanity's most precisely-tested description of nature.

That said, Eqs. (11.1), (11.2), and (11.3) are not written in a way that makes it clear they are compatible with Lorentz covariance. The fields and the force are written using 3-vectors, which depend upon us choosing a particular observer's "space" coordinates; the field equations are expressed using a particular observer's time and space derivatives. These equations are formulated for one particular reference frame, and it is not obvious how they will transform to another reference frame. The goal of the next two lectures is to think how to organize the structures expressed in Eqs. (11.1), (11.2), and (11.3) in a way that clearly shows electrodynamics is a Lorentz covariant theory.

11.2 How to organize the fields

11.2.1 General considerations

So far, when we've translated a physical quantity into Lorentz covariant language, we have found a way of taking quantities which are 3-vectors and mapping them into 4-vectors. Examples so far are displacement (add ct as the "zeroth" component), the 4-velocity (change d/dt to $d/d\tau$ so that we use a clock whose meaning is invariant to describe time derivatives; add $c\,dt/d\tau = \gamma c$ as the zeroth component), and the 4-momentum (add energy as the zeroth component, dividing by c to make sure the dimensions are sensible). Can we do this with the electric and magnetic fields?

We have several problems here. First, we know that $\mathbf{E}$ and $\mathbf{B}$ fields must transform into one another when we change frames: what is pure magnetic field in one frame is a mixture of magnetic and electric fields in another; and vice versa. The classic example of this is a charge moving in a magnetic field. Consider a charge moving parallel to a current-carrying wire, as illustrated in Fig. 1:

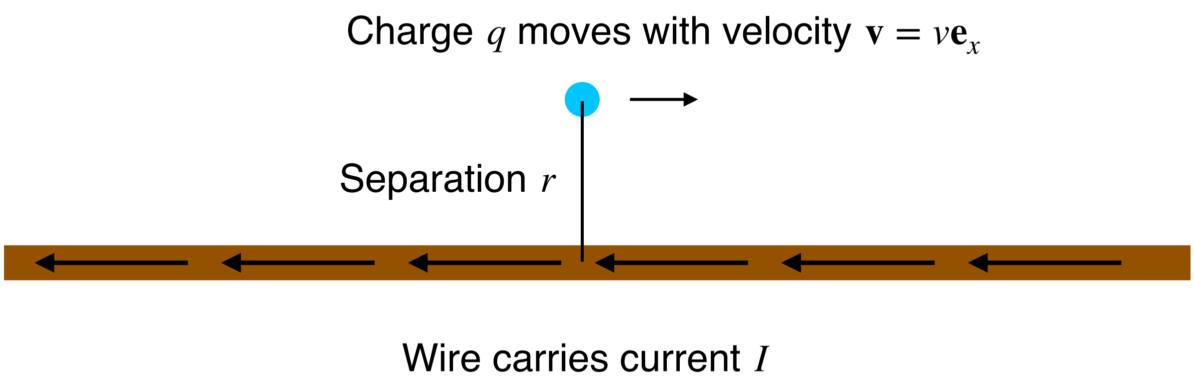

Figure 1: A charge q moving parallel to a wire carrying a current I.

For concreteness, let's define $\mathbf{e}_x$ as pointing to the right, $\mathbf{e}_y$ as pointing into the page, and $\mathbf{e}_z$ as pointing up. Then, in what we will call the "lab" frame L, we have a charge q that moves to the right. The charge is a distance r from a wire that carries a current flowing to the left. As we learned in 8.02/8.022, this wire generates a magnetic field that circulates around the wire. At the location of the charge, this field takes the value

$$\mathbf{B} = \frac{\mu_0 I}{2\pi r}\mathbf{e}_y \; . \tag{11.4}$$

The wire is neutral, so the charge q does not feel any electric force — it only feels a magnetic force, whose value is

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B} = \frac{\mu_0 q I v}{2\pi r}\mathbf{e}_z \; . \tag{11.5}$$

This force points "up" in the figure — the charge is repelled from the wire.

Let's now change frames, and think about what must happen. First, we require q to be the same in all reference frames. If changing frames changed the value of charge, the elementary charge would vary for moving charges. Imagine the effect this would cause for a system in which there are members whose charges are equal and opposite, but are moving at different relative speeds. A system which is neutral when its members "sit still" might have net charge when they are in motion! In addition to feeling absurd, the fact is that we have no experimental evidence for anything like this whatsoever. Observations and measurements indicate that a body's charge is unchanged no matter how fast we observe it to move.

So, let's jump into a reference frame that moves with $\mathbf{v} = v\mathbf{e}_x$ — i.e., the frame C in which the charge is at rest. In this frame *there can be no magnetic force*. The magnetic force is proportional to the charge's speed. If the speed is zero, the magnetic force is zero. However, a repulsive force in one frame of reference is *not consistent* with no force in another.

The details of how the force behaves in this frame might differ[1] (perhaps its magnitude will be different), but there still must be an overall repulsive force. If there is no magnetic force, then there must instead be an *electric* force.

This means there must be an electric field in the charge's rest frame, even though there was no such field in the lab frame. Something that we measured to be pure magnetic field transforms to a mixture of electric and magnetic field. Whatever "entity" we will use to describe electric and magnetic fields in special relativity must be able to transform magnetic fields into electric fields (and vice versa).

11.2.2 A covariant representation of the force and fields

Our root issue is essentially one of simple counting. We have had success fitting important physical quantities into 4-vectors so far, but it just isn't going to work for the electric and magnetic field. They have 6 components, and we cannot fit these 6 pieces of information into the 4 components of a 4-vector. We need something bigger.

A simple example of a bigger object is a 2nd-rank tensor, which has 16 components. That's too many; but, we can reduce the number of free components by imposing symmetry. If we use a symmetric tensor, then it has 10 free components — still too many. But an antisymmetric 2nd-rank tensor has 6 free components — exactly what we need.

So let's think how we can fit the 6 components $(E^x,\ E^y,\ E^z)$, $(B^x,\ B^y,\ B^z)$ into an antisymmetric 2nd-rank tensor which we will call $F^{\alpha\beta}$. To guide us, let's deduce how the Lorentz force law, $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, can be written in a fully covariant manner.

First, we "upgrade" the force. We start with $\mathbf{F} = d\mathbf{p}/dt$. Clearly, we will want to take the 3-momentum $\mathbf{p}$ over to the 4-momentum, whose components are p^α. We also need to upgrade the time derivative with one that uses a notion of time that all frames are happy to use as a point of reference. Just as we did in defining the 4-velocity, let's replace d/dt with $d/d\tau$, where τ is the proper time measured by the body which is experiencing the force.

What about the right-hand side, $q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$? This a quantity that is linear in q, linear in the fields, and — if we think about this carefully — linear in the components of the velocity. "Wait," I imagine you protesting, "the $\mathbf{B}$ term is linear in components of velocity, but what about the $\mathbf{E}$ term?" Note that $\mathbf{E}$ and $\mathbf{B}$ have different dimensions: the dimensions of $\mathbf{E}$ are force over charge, but the dimensions of $\mathbf{B}$ are force over speed times charge. When we assemble these quantities into a single tensor, we'll need to account for the difference in dimensions. We often do this by throwing in factors of the speed of light. This suggests that we think about the Lorentz force law as

$$\mathbf{F} = q\left[c\left(\frac{\mathbf{E}}{c}\right) + \mathbf{v} \times \mathbf{B}\right] \ . \tag{11.6}$$

Bearing in mind that the components of u^α are given approximately by (c, v^x, v^y, v^z) for a body that is not moving very fast relative to us, this suggests that in the Lorentz force law, the electric field is being multiplied by the *timelike* component of the 4-momentum.

Putting all this together, we want the covariant formulation of the Lorentz force to be

$$\frac{dp^\alpha}{d\tau} = qF^{\alpha\beta}u_\beta \ . \tag{11.7}$$

[1] In a few lectures we will look carefully at forces and accelerations in special relativity; we briefly introduce a handful of important issues a little later in this lecture.

Let's now figure out how to "fill up" the tensor $F^{\alpha\beta}$ so that this is consistent with the Lorentz force law that we learned about in 8.02/8.022 by going through the spatial components, $\alpha = 1, 2, 3$, one by one. (We'll come back to the $\alpha = 0$ component later.) First look at $\alpha = 1$, or $\alpha = x$:

$$\frac{dp^x}{d\tau} = q\left(F^{10}u_0 + F^{12}u_2 + F^{13}u_3\right) \ . \tag{11.8}$$

There is no F^{11} term because of this tensor's antisymmetry — all diagonal elements are zero. Let's further use the fact that $u_0 = -u^0 = -\gamma c$, $u_2 = \gamma(dy/dt)$, and $u_3 = \gamma(dz/dt)$:

$$\frac{dp^x}{d\tau} = \gamma q\left(-cF^{10} + F^{12}\frac{dy}{dt} + F^{13}\frac{dz}{dt}\right) \ . \tag{11.9}$$

Next, divide by sides by γ and use the fact that an interval of time dt measured by clocks in this frame is $\gamma d\tau$:

$$\frac{dp^x}{dt} = q\left(-cF^{10} + F^{12}\frac{dy}{dt} + F^{13}\frac{dz}{dt}\right) \ . \tag{11.10}$$

Compare this to the x component of the Lorentz force law:

$$\frac{dp^x}{dt} = q\left(E^x + B^z\frac{dy}{dt} - B^y\frac{dz}{dt}\right) \ . \tag{11.11}$$

This allows us to read off

$$F^{10} = -E^x/c \ , \qquad F^{12} = B^z \ , \qquad F^{13} = -B^y \ . \tag{11.12}$$

Repeating this exercise for the y and z force components and noting that the tensor is antisymmetric allows us to fill it in entirely:

$$F^{\alpha\beta} \doteq \begin{pmatrix} 0 & E^x/c & E^y/c & E^z/c \\ -E^x/c & 0 & B^z & -B^y \\ -E^y/c & -B^z & 0 & B^x \\ -E^z/c & B^y & -B^x & 0 \end{pmatrix} \ . \tag{11.13}$$

This tensor is often called the *Faraday* tensor. It replaces the 3-vectors which describe electric and magnetic fields according to some particular observer's reference frame with a geometric object whose components can be readily translated to any reference frame; and, it connects to 4-vectors whose components can likewise be readily translated to any reference frame.

11.3 A brief aside on forces and accelerations

In this lecture, we've been talking about a specific force without yet having discussed forces in special relativity in broader terms. We will discuss forces, accelerations, and the properties of accelerated observers in more detail in an upcoming lecture. Certain aspects of this discussion are needed now, so we pause in our discussion of electric and magnetic fields for a brief digression to talk about forces and accelerations.

As we have discussed, a body of mass m moving with 4-velocity $\vec{u}$ has a 4-momentum $\vec{p} = m\vec{u}$. As you have seen in our discussion above, this momentum changes if the body is acted on by a force or, more properly, a 4-force:

$$\vec{F} = \frac{d\vec{p}}{d\tau} \ . \tag{11.14}$$

If the body's mass cannot change, then this leads to the body having a 4-acceleration:

$$\vec{a} = \frac{1}{m}\vec{F} = \frac{d\vec{u}}{d\tau} \; . \tag{11.15}$$

When we discuss 3-velocities $\mathbf{u}$ and 3-accelerations $\mathbf{a}$, these quantities can have largely any value that we want them to have: the value of $\mathbf{u}$ is essentially an initial condition to our analysis, and the value of $\mathbf{a}$ is only constrained by the mechanism providing the force $\mathbf{F}$.

Not so for the 4-velocity and the 4-acceleration: there is a very interesting and important *constraint* which these two quantities must always satisfy. To see where this comes from, begin with the invariant that we can construct from $\vec{u}$:

$$\vec{u} \cdot \vec{u} = -c^2 \; . \tag{11.16}$$

Take $d/d\tau$ of both sides of this equation:

$$\vec{a} \cdot \vec{u} + \vec{u} \cdot \vec{a} = 0 \; , \tag{11.17}$$

or

$$\vec{a} \cdot \vec{u} = 0 \; . \tag{11.18}$$

The 4-velocity and the 4-acceleration are always "orthogonal" in spacetime. This important constraint has important implications for the nature of any 4-force that you may compute — if at the end of your analysis, you find that $\vec{F} \cdot \vec{u} \neq 0$, you have made a mistake or have overlooked something important.

11.4 Some details of the electromagnetic 4-force

With the above discussion in mind, let's examine the electromagnetic 4-force that we have worked out. Is it the case that $\vec{F} \cdot \vec{u} = 0$? The answer is yes, and we can show this using a little bit of "index gymnastics":

$$\begin{aligned}
\vec{F} \cdot \vec{u} &= qF^{\alpha\beta} u_\beta u_\alpha & (11.19)\\
&= -qF^{\beta\alpha} u_\beta u_\alpha & (11.20)\\
&= -qF^{\beta\alpha} u_\alpha u_\beta & (11.21)\\
&= -qF^{\alpha\beta} u_\beta u_\alpha \; . & (11.22)
\end{aligned}$$

Let's step through these lines of analysis carefully. On the first line, we have have contracted the definition of the electromagnetic 4-force, Eq. (11.7), with the 4-velocity in order to make the inner product. On the second line, we have used the fact that the Faraday tensor is antisymmetric to swap the order of the indices on the tensor, introducing a minus sign. On the third line, we have used the fact that $u_\alpha u_\beta$ is symmetric to swap the order of their indices. On the final line, we have used the fact that α and β are "dummy" indices — they are being summed over, so it doesn't matter how we label them. We can in fact change α for β and β for α, as long as we do this *consistently* throughout the expression.

Now compare the first line with the fourth line. Their right-hand sides are *identical* ... except for a minus sign. This is thus an expression of the form $x = -x$, whose only solution is $x = 0$. We conclude that

$$\vec{F} \cdot \vec{u} = 0 \; , \tag{11.23}$$

So our 4-force indeed is spacetime orthogonal with the 4-velocity — as it should be.

Two remarks on this calculation:

- This is our first encounter with a trick that gets used a lot: whenever you contract all free indices of a totally antisymmetric mathematical object, like $F^{\alpha\beta}$, against a totally *symmetric* mathematical object, like $u_\alpha u_\beta$, the result is zero.

  If this makes you nervous and you want to be totally confident in the result, you can always go through an exercise like the one that I did above. The key point is that by combining symmetric with antisymmetric, we add up terms that are equal and opposite. If you expand out the Einstein summation that I did above, you find that you can combine terms in pairs: $F^{10}u_1u_0 + F^{01}u_0u_1$, $F^{23}u_2u_3 + F^{32}u_3u_2$, etc. The members of each pair will always be equal in magnitude and opposite in sign.

- When a force law is set up properly, it generally works out "automatically" that we find $\vec{F}\cdot\vec{u} = 0$, in much the way that it did for this electromagnetic 4-force. Finding $\vec{F}\cdot\vec{u} = 0$ does not guarantee that your force law is correct, but *not* finding this guarantees that your force law is wrong.

Before moving on to other aspects of the covariant formulation of electric and magnetic fields, let's clean up one last detail. We saw in our calculation above that the $\alpha = 1$, 2, and 3 components of the 4-force correspond perfectly to the x, y, and z components of the Lorentz force. What is the $\alpha = 0$ component? Let's write this out:

$$\begin{aligned}\frac{dp^0}{d\tau} &= qF^{0\beta}u_\beta = q\left(F^{01}u_1 + F^{02}u_2 + F^{03}u_3\right)\\ &= \frac{\gamma q}{c}\left(E^x(\mathbf{u})^x + E^y(\mathbf{u})^y + E^z(\mathbf{u})^z\right)\\ &= \frac{\gamma q}{c}\mathbf{E}\cdot\mathbf{u}\,. \end{aligned} \tag{11.24}$$

Using the fact that $p^0 = E/c$, where E with no indices and no boldface means the energy[2] of the charged body, and using $dt = \gamma d\tau$, this becomes

$$\frac{dE}{dt} = q\mathbf{E}\cdot\mathbf{u}\,. \tag{11.25}$$

This expression tells us about the rate at which *work* is done on the charge by the electric field. If you need a reminder of where this comes from, remember that the differential of work done in moving through a 3-displacement $d\mathbf{r}$ in an $\mathbf{E}$ field is

$$dW = \mathbf{F}\cdot d\mathbf{r} = q\mathbf{E}\cdot d\mathbf{r}\,. \tag{11.26}$$

If the charge does this in a time dt, then

$$\frac{dW}{dt} = q\mathbf{E}\cdot\frac{d\mathbf{r}}{dt}\,, \tag{11.27}$$

in agreement with Eq. (11.25).

[2]The letter "E" is doing double duty here, standing for both energy and electric field. Sometimes people use U for energy in circumstances like this, in order to reduce the likelihood of any confusion.

11.5 Transforming electric and magnetic fields

By fitting the electric and magnetic fields into a rank-2 tensor, it becomes simple to deduce how these fields transform when we change frames. Let observer $\mathcal{O}$ measure fields described by the tensor $F^{\alpha\beta}$; let $\mathcal{O}'$ in a different inertial frame measure fields described by the tensor $F^{\mu'\nu'}$. These are related by converting using Lorentz transformation matrices:

$$F^{\mu'\nu'} = F^{\alpha\beta}\Lambda^{\mu'}{}_{\alpha}\Lambda^{\nu'}{}_{\beta} \ . \tag{11.28}$$

Let's work through this using the Lorentz transformation matrix

$$\Lambda^{\mu'}{}_{\alpha} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \ . \tag{11.29}$$

In other words, we take $\mathcal{O}'$ to be moving with $\mathbf{v} = v\mathbf{e}_x$ relative to $\mathcal{O}$. Let's use this calculate how the components of the Faraday tensor translate between frames. Start by working through the transformation for the $(0'1')$ component:

$$\begin{aligned} F^{0'1'} &= \Lambda^{0'}{}_{0}\Lambda^{1'}{}_{1}F^{01} + \Lambda^{0'}{}_{1}\Lambda^{1'}{}_{0}F^{10} \\ &= \left(\gamma^2 - \gamma^2\beta^2\right) F^{01} \\ &= F^{01} \ . \end{aligned} \tag{11.30}$$

On the first line, we expanded the transformation rule to write out all the non-zero terms that contribute to $F^{0'1'}$. This amounts to all the lambda matrix elements that have $0'$ on the first index, and all the matrix elements that have $1'$ on the first index. A total of 4 such elements exist: $\Lambda^{0'}{}_{0} = \gamma$, $\Lambda^{0'}{}_{1} = -\gamma\beta$, $\Lambda^{1'}{}_{0} = -\gamma\beta$, and $\Lambda^{1'}{}_{1} = \gamma$; all the others ones with $0'$ or $1'$ in the first position are zero. We then used antisymmetry, and then used the fact that $\gamma = 1/\sqrt{1-\beta^2}$ to clean this expression up. Translating back into electric and magnetic field components, this tells us

$$E^{x'} = E^{x} \ . \tag{11.31}$$

Move on to the $(0'2')$ component:

$$\begin{aligned} F^{0'2'} &= \Lambda^{0'}{}_{0}\Lambda^{2'}{}_{2}F^{02} + \Lambda^{0'}{}_{1}\Lambda^{2'}{}_{2}F^{12} \\ &= \gamma F^{02} - \gamma\beta F^{12} \ . \end{aligned} \tag{11.32}$$

We cannot simplify this any further, so we now translate back into electric and magnetic field components:

$$E^{y'} = \gamma(E^y - vB^z) \ . \tag{11.33}$$

Next the $(0'3')$ component:

$$\begin{aligned} F^{0'3'} &= \Lambda^{0'}{}_{0}\Lambda^{3'}{}_{3}F^{03} + \Lambda^{0'}{}_{1}\Lambda^{3'}{}_{3}F^{13} \\ &= \gamma F^{03} - \gamma\beta F^{13} \ . \end{aligned} \tag{11.34}$$

This becomes

$$E^{z'} = \gamma(E^z + vB^y) \ . \tag{11.35}$$

Doing a similar exercise for the components of the Faraday tensor which map to the magnetic fields, we find

$$B^{x'} = B^x \ , \qquad B^{y'} = \gamma(B^y + vE^z/c^2) \ , \qquad B^{z'} = \gamma(B^z - vE^y/c^2) \ . \tag{11.36}$$

By repeating this analysis for frames moving with $\mathbf{v} = v\mathbf{e}_y$ and $\mathbf{v} = v\mathbf{e}_z$, it's not too difficult to work out the general rule for transforming between frames. For *completely general* $\mathbf{v}$, we have

$$\mathbf{E}'_\parallel = \mathbf{E}_\parallel , \qquad \mathbf{E}'_\perp = \gamma\left(\mathbf{E}_\perp + \mathbf{v} \times \mathbf{B}_\perp\right) \; ; \tag{11.37}$$

$$\mathbf{B}'_\parallel = \mathbf{B}_\parallel , \qquad \mathbf{B}'_\perp = \gamma\left(\mathbf{B}_\perp - \mathbf{v} \times \mathbf{E}_\perp/c^2\right) \; . \tag{11.38}$$

Here, $\mathbf{E}_\parallel$ denotes the component of $\mathbf{E}$ that is parallel $\mathbf{v}$. Let $\mathbf{e}_v \equiv \mathbf{v}/v$ denote the unit vector along the velocity vector; then, $\mathbf{E}_\parallel = (\mathbf{E}\cdot\mathbf{e}_v)\mathbf{e}_v$. The other component, $\mathbf{E}_\perp = \mathbf{E} - \mathbf{E}_\parallel$, denotes the part of $\mathbf{E}$ that is orthogonal to $\mathbf{v}$. The magnetic field vectors $\mathbf{B}_\parallel$ and $\mathbf{B}_\perp$ are defined likewise.

When I first was presented with the transformation laws (11.37) and (11.38), I was utterly baffled. Though I understood the derivation (which I learned from Purcell's E&M textbook), the rule we find for transforming these fields looks *nothing* like any of the Lorentz transformation rules I learned for other quantities! It was only after learning about tensors, understanding that $\mathbf{E}$ and $\mathbf{B}$ were best thought of us components of a rank-2 antisymmetric tensor, and spending some time developing fluency with operations like Eq. (11.28) that I started to become comfortable with these rules.