5. Electromagnetism and Relativity

We've seen that Maxwell's equations have wave solutions which travel at the speed of light. But there's another place in physics where the speed of light plays a prominent role: the theory of special relativity. How does electromagnetism fit with special relativity?

Historically, the Maxwell equations were discovered before the theory of special relativity. It was thought that the light waves we derived above must be oscillations of some substance which fills all of space. This was dubbed the *aether*. The idea was that Maxwell's equations only hold in the frame in which the aether is at rest; light should then travel at speed c relative to the aether.

We now know that the concept of the aether is unnecessary baggage. Instead, Maxwell's equations hold in all inertial frames and are the first equations of physics which are consistent with the laws of special relativity. Ultimately, it was by studying the Maxwell equations that Lorentz was able to determine the form of the Lorentz transformations which subsequently laid the foundation for Einstein's vision of space and time.

Our goal in this section is to view electromagnetism through the lens of relativity. We will find that observers in different frames will disagree on what they call electric fields and what they call magnetic fields. They will observe different charge densities and different currents. But all will agree that these quantities are related by the same Maxwell equations. Moreover, there is a pay-off to this. It's only when we formulate the Maxwell equations in a way which is manifestly consistent with relativity that we see their true beauty. The slightly cumbersome vector calculus equations that we've been playing with throughout these lectures will be replaced by a much more elegant and simple-looking set of equations.

5.1 A Review of Special Relativity

We start with a very quick review of the relevant concepts of special relativity. (For more details see the lecture notes on Dynamics and Relativity). The basic postulate of relativity is that the laws of physics are the same in all inertial reference frames. The guts of the theory tell us how things look to observers who are moving relative to each other.

The first observer sits in an inertial frame $\mathcal{S}$ with spacetime coordinates (ct, x, y, z) the second observer sits in an inertial frame $\mathcal{S}'$ with spacetime coordinates (ct', x', y', z').

If we take $\mathcal{S}'$ to be moving with speed v in the x-direction relative to $\mathcal{S}$ then the coordinate systems are related by the Lorentz boost

$$x' = \gamma\left(x - \frac{v}{c}\,ct\right) \quad \text{and} \quad ct' = \gamma\left(ct - \frac{v}{c}x\right) \tag{5.1}$$

while $y' = y$ and $z' = z$. Here c is the speed of light which has the value,

$$c = 299792458\ ms^{-1}$$

Meanwhile γ is the ubiquitous factor

$$\gamma = \sqrt{\frac{1}{1 - v^2/c^2}} \tag{5.2}$$

The Lorentz transformation (5.1) encodes within it all of the fun ideas of time dilation and length contraction that we saw in our first course on relativity.

5.1.1 Four-Vectors

It's extremely useful to package these spacetime coordinates in 4-vectors, with indices running from $\mu = 0$ to $\mu = 3$

$$X^\mu = (ct, x, y, z) \qquad \mu = 0, 1, 2, 3$$

Note that the index is a superscript rather than subscript. This will be important shortly. A general Lorentz transformation is a linear map from X to X' of the form

$$(X')^\mu = \Lambda^\mu{}_\nu X^\nu$$

Here Λ is a 4×4 matrix which obeys the matrix equation

$$\Lambda^T \eta \Lambda = \eta \qquad \Leftrightarrow \qquad \Lambda^\rho{}_\mu \eta_{\rho\sigma} \Lambda^\sigma{}_\nu = \eta_{\mu\nu} \tag{5.3}$$

with $\eta_{\mu\nu}$ the *Minkowski metric*

$$\eta_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)$$

The solutions to (5.3) fall into two classes. The first class is simply rotations. Given a 3×3 rotation matrix R obeying $R^T R = 1$, we can construct a Lorentz transformation Λ obeying (5.3) by embedding R in the spatial part,

$$\Lambda^\mu{}_\nu = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & \mathrm{R} & \\ 0 & & & \end{array}\right) \tag{5.4}$$

These transformations describe how to relate the coordinates of two observers who are rotated with respect to each other.

The other class of solutions to (5.3) are the Lorentz boosts. These are the transformations appropriate for observers moving relative to each other. The Lorentz transformation (5.1) is equivalent to

$$\Lambda^{\mu}{}_{\nu} = \left(\begin{array}{cc|cc} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \tag{5.5}$$

There are similar solutions associated to boosts along the y and z axes.

The beauty of 4-vectors is that it's extremely easy to write down invariant quantities. These are things which all observers, no matter which their reference frame, can agree on. To construct these we take the inner product of two 4-vectors. The trick is that this inner product uses the Minkowski metric and so comes with some minus signs. For example, the square of the distance from the origin to some point in spacetime labelled by X is

$$X \cdot X = X^{\mu}\eta_{\mu\nu}X^{\nu} = c^2t^2 - x^2 - y^2 - z^2$$

which is the invariant interval. Similarly, if we're given two four-vectors X and Y then the inner product $X \cdot Y = X^{\mu}\eta_{\mu\nu}Y^{\nu}$ is also a Lorentz invariant.

5.1.2 Proper Time

The key to building relativistic theories of Nature is to find the variables that have nice properties under Lorentz transformations. The 4-vectors X, labelling spacetime points, are a good start. But we need more. Here we review how the other kinematical variables of velocity, momentum and acceleration fit into 4-vectors.

Suppose that, in some frame, the particle traces out a worldline. The clever trick is to find a way to parameterise this path in a way that all observers agree upon. The natural choice is the *proper time* τ, the duration of time experienced by the particle itself. If you're sitting in some frame, watching some particle move with an old-fashioned Newtonian 3-velocity $\mathbf{u}(t)$, then it's simple to show that the relationship between your time t and the proper time of the particle τ is given by

$$\frac{dt}{d\tau} = \gamma(\mathbf{u})$$

The proper time allows us to define the 4-velocity and the 4-momentum. Suppose that the particle traces out a path $X(\tau)$ in some frame. Then the 4-velocity is

$$U = \frac{dX}{d\tau} = \gamma \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix}$$

Similarly, the 4-momentum is $P = mU$ where m is the rest mass of the particle. We write

$$P = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix} \tag{5.6}$$

where $E = m\gamma c^2$ is the energy of the particle and $\mathbf{p} = \gamma m\mathbf{u}$ is the 3-momentum in special relativity.

The importance of U and P is that they too are 4-vectors. Because all observers agree on τ, the transformation law of U and P are inherited from X. This means that under a Lorentz transformation, they too change as $U \to \Lambda U$ and $P \to \Lambda P$. And it means that inner products of U and P are guaranteed to be Lorentz invariant.

5.1.3 Indices Up, Indices Down

Before we move on, we do need to introduce one extra notational novelty. The minus signs in the Minkowski metric η means that it's useful to introduce a slight twist to the usual summation convention of repeated indices. For all the 4-vectors that we introduced above, we always place the spacetime index $\mu = 0, 1, 2, 3$ as a superscript (i.e. up) rather than a subscript.

$$X^\mu = \begin{pmatrix} ct \\ \mathbf{x} \end{pmatrix}$$

This is because the same object with an index down, X_μ, will mean something subtly different. We define

$$X_\mu = \begin{pmatrix} ct \\ -\mathbf{x} \end{pmatrix}$$

With this convention, the Minkowski inner product can be written using the usual convention of summing over repeated indices as

$$X^\mu X_\mu = c^2t^2 - \mathbf{x} \cdot \mathbf{x}$$

In contrast, $X^\mu X^\mu = c^2t^2 + \mathbf{x}^2$ is a dumb thing to write in the context of special relativity since it looks very different to observers in different inertial frames. In fact, we will shortly declare it illegal to write things like $X^\mu X^\mu$.

There is a natural way to think of X_μ in terms of X^μ using the Minkowski metric $\eta_{\mu\nu} = \text{diag}(+1, -1, -1, -1)$. The following equation is trivially true:

$$X_\mu = \eta_{\mu\nu} X^\nu$$

This means that we can think of the Minkowski metric as allowing us to lower indices. To raise indices back up, we need the inverse of $\eta_{\mu\nu}$ which, fortunately, is the same matrix: $\eta^{\mu\nu} = \text{diag}(+1, -1, -1, -1)$ which means we have $\eta^{\mu\rho}\eta_{\rho\nu} = \delta^\mu{}_\nu$ and we can write

$$X^\nu = \eta^{\nu\mu} X_\mu$$

From now on, we're going to retain this distinction between all upper and lower indices. All the four-vectors that we've met so far have upper indices. But all can be lowered in the same way. For example, we have

$$U_\mu = \gamma \begin{pmatrix} c \\ -\mathbf{u} \end{pmatrix} \tag{5.7}$$

This trick of distinguishing between indices up and indices down provides a simple formalism to ensure that all objects have nice transformation properties under the Lorentz group. We insist that, just as in the usual summation convention, repeated indices only ever appear in pairs. But now we further insist that pairs always appear with one index up and the other down. The result will be an object which is invariant under Lorentz transformations.

5.1.4 Vectors, Covectors and Tensors

In future courses, you will learn that there is somewhat deeper mathematics lying behind distinguishing X^μ and X_μ: formally, these objects live in different spaces (sometimes called dual spaces). We'll continue to refer to X^μ as vectors, but to distinguish them, we'll call X_μ *covectors*. (In slightly fancier language, the components of the vector X^μ are sometimes said to be *contravariant* while the components of the covector X_μ are said to be *covariant*).

For now, the primary difference between a vector and covector is how they transform under rotations and boosts. We know that, under a Lorentz transformation, any 4-vector changes as

$$X^\mu \to X'^\mu = \Lambda^\mu{}_\nu X^\nu \tag{5.8}$$

From this, we see that a covector should transform as

$$\begin{aligned} X_\mu \to X'_\mu &= \eta_{\mu\rho} X'^{\rho} \\ &= \eta_{\mu\rho} \Lambda^\rho{}_\sigma X^\sigma \\ &= \eta_{\mu\rho} \Lambda^\rho{}_\sigma \eta^{\sigma\nu} X_\nu \end{aligned}$$

Using our rule for raising and lowering indices, now applied to the Lorentz transformation Λ, we can also write this as

$$X_\mu \to \Lambda_\mu{}^\nu X_\nu$$

where our notation is now getting dangerously subtle: you have to stare to see whether the upper or lower index on the Lorentz transformation comes first.

There is a sense in which $\Lambda_\mu{}^\nu$ can be thought of a the components of the inverse matrix Λ^{-1}. To see this, we go back to the definition of the Lorentz transformation (5.3), and start to use our new rules for raising and lowering indices

$$\begin{aligned} \Lambda^\rho{}_\mu \eta_{\rho\sigma} \Lambda^\sigma{}_\nu = \eta_{\mu\nu} \quad &\Rightarrow \quad \Lambda^\rho{}_\mu \Lambda_{\rho\nu} = \eta_{\mu\nu} \\ &\Rightarrow \quad \Lambda^\rho{}_\mu \Lambda_\rho{}^\sigma = \delta^\sigma_\mu \\ &\Rightarrow \quad \Lambda_\rho{}^\sigma \Lambda^\rho{}_\mu = \delta^\sigma_\mu \end{aligned}$$

In the last line above, we've simply reversed the order of the two terms on the left. (When written in index notation, these are just the entries of the matrix so there's no problem with commuting them). Now we compare this to the formula for the inverse of a matrix,

$$(\Lambda^{-1})^\sigma{}_\rho \Lambda^\rho{}_\mu = \delta^\sigma_\mu \quad \Rightarrow \quad (\Lambda^{-1})^\sigma{}_\rho = \Lambda_\rho{}^\sigma \tag{5.9}$$

Note that you need to be careful where you place the indices in equations like this. The result (5.9) is analogous to the statement that the inverse of a rotation matrix is the transpose matrix. For general Lorentz transformations, we learn that the inverse is sort of the transpose where "sort of" means that there are minus signs from raising and lowering. The placement of indices in (5.9) tells us where those minus signs go.

The upshot of (5.9) is that if we want to abandon index notation all together then vectors transform as $X \to \Lambda X$ while covectors – which, for the purpose of this sentence, we'll call $\tilde{X}$ – transform as $\tilde{X} \to \Lambda^{-1}\tilde{X}$. However, in what follows, we have no intention of abandoning index notation. Instead, we will embrace it. It will be our friend and our guide in showing that the Maxwell equations are consistent with special relativity.

A particularly useful example of a covector is the *four-derivative*. This is the relativistic generalisation of ∇, defined by

$$\partial_\mu = \frac{\partial}{\partial X^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}\ , \nabla \right)$$

Notice that the superscript on the spacetime 4-vector X^μ has migrated to a subscript on the derivative ∂_μ. For this to make notational sense, we should check that ∂_μ does indeed transform as covector. This is a simple application of the chain rule. Under a Lorentz transformation, $X^\mu \to X'^\mu = \Lambda^\mu{}_\nu X^\nu$, so we have

$$\partial_\mu = \frac{\partial}{\partial X^\mu} \to \frac{\partial}{\partial X'^\mu} = \frac{\partial X^\nu}{\partial X'^\mu}\frac{\partial}{\partial X^\nu} = (\Lambda^{-1})^\nu{}_\mu \partial_\nu = \Lambda_\mu{}^\nu \partial_\nu$$

which is indeed the transformation of a co-vector.

Tensors

Vectors and covectors are the simplest examples of objects which have nice transformation properties under the Lorentz group. But there are many more examples. The most general object can have a bunch of upper indices and a bunch of lower indices, $T^{\mu_1\ldots\mu_n}{}_{\nu_1\ldots\nu_m}$. These objects are also called *tensors* of type (n, m). In order to qualify as a tensor, they must transform under a Lorentz transformation as

$$T'^{\mu_1\ldots\mu_n}{}_{\nu_1\ldots\nu_m} = \Lambda^{\mu_1}{}_{\rho_1}\ldots\Lambda^{\mu_n}{}_{\rho_n}\Lambda_{\nu_1}{}^{\sigma_1}\ldots\Lambda_{\nu_m}{}^{\sigma_m}T^{\rho_1\ldots\rho_n}{}_{\sigma_1\ldots\sigma_m} \tag{5.10}$$

You can always use the Minkowski metric to raise and lower indices on tensors, changing the type of tensor but keeping the total number of indices $n + m$ fixed.

Tensors of this kind are the building blocks of all our theories. This is because if you build equations only out of tensors which transform in this manner then, as long as the $\mu, \nu, \ldots$ indices match up on both sides of the equation, you're guaranteed to have an equation that looks the same in all inertial frames. Such equations are said to be *covariant*. You'll see more of this kind of thing in courses on *General Relativity* and *Differential Geometry*.

In some sense, this index notation is too good. Remember all those wonderful things that you first learned about in special relativity: time dilation and length contraction and twins and spaceships so on. You'll never have to worry about those again. From now on, you can guarantee that you're working with a theory consistent with relativity by ensuring two simple things

- That you only deal with tensors.
- That the indices match up on both sides of the equation.

It's sad, but true. It's all part of growing up and not having fun anymore.

5.2 Conserved Currents

We started these lectures by discussing the charge density $\rho(\mathbf{x}, t)$, the current density $\mathbf{J}(\mathbf{x}, t)$ and their relation through the continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0$$

which tells us that charge is locally conserved.

The continuity equation is already fully consistent with relativity. To see this, we first need to appreciate that the charge and current densities sit nicely together in a 4-vector,

$$J^\mu = \begin{pmatrix} \rho c \\ \mathbf{J} \end{pmatrix}$$

Of course, placing objects in a four-vector has consequence: it tells us how these objects look to different observers. Let's quickly convince ourselves that it makes sense that charge density and current do indeed transform in this way. We can start by considering a situation where there are only static charges with density ρ_0 and no current. So $J^\mu = (\rho_0, 0)$. Now, in a frame that is boosted by velocity $\mathbf{v}$, the current will appear as $J'^\mu = \Lambda^\mu{}_\nu J^\nu$ with the Lorentz transformation given by (5.5). The new charge density and current are then

$$\rho' = \gamma \rho_0 \quad , \quad \mathbf{J}' = -\gamma \rho \mathbf{v}$$

The first of these equations tells us that different observers see different charge densities. This is because of Lorentz contraction: charge *density* means charge per unit volume. And the volume gets squeezed because lengths parallel to the motion undergo Lorentz contraction. That's the reason for the factor of γ in the observed charge density. Meanwhile, the second of these equations is just the relativistic extension of the formula $\mathbf{J} = \rho \mathbf{v}$ that we first saw in the introduction. (The extra minus sign is because $\mathbf{v}$ here denotes the velocity of the boosted observer; the charge is therefore moving with relative velocity $-\mathbf{v}$).

In our new, relativistic, notation, the continuity equation takes the particularly simple form

$$\partial_\mu J^\mu = 0 \tag{5.11}$$

This equation is Lorentz invariant. This follows simply because the indices are contracted in the right way: one up, and one down.

5.2.1 Magnetism and Relativity

We've learned something unsurprising: boosted charge gives rise to a current. But, combined with our previous knowledge, this tells us something new and important: boosted electric fields must give rise to magnetic fields. The rest of this chapter will be devoted to understanding the details of how this happens. But first, we're going to look at a simple example where we can re-derive the magnetic force purely from the Coulomb force and a dose of Lorentz contraction.

To start, consider a bunch of positive charges $+q$ moving along a line with speed $+v$ and a bunch of negative charges $-q$ moving in the opposite direction with speed $-v$ as shown in the figure. If there is equal density, n, of positive and negative charges then the charge density vanishes while the current is

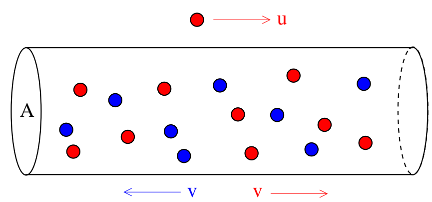

Figure 47:

$$I = 2nAqv$$

where A is the cross-sectional area of the wire. Now consider a test particle, also carrying charge q, which is moving parallel to the wire with some speed u. It doesn't feel any electric force because the wire is neutral, but we know it experiences a magnetic force. Here we will show how to find an expression for this force without ever invoking the phenomenon of magnetism.

The trick is to move to the rest frame of the test particle. This means we have to boost by speed u. The usual addition formula tells us that the velocities of the positive and negative charges now differ, given by

$$v_{\pm} = \frac{v \mp u}{1 \mp uv/c^2}$$

But with the boost comes a Lorentz contraction which means that the charge density changes. Moreover, because the velocities of positive and negative charges are now different, this will mean that, viewed from the rest frame of our particle, the wire is no longer neutral. Let's see how this works. First, we'll introduce n_0, the density of charges when the particles in the wire are at rest. Then the density of the $+q$ charges in the original frame is

$$\rho = qn = \gamma(v)qn_0$$

The charge density of the $-q$ particles is the same, but with opposite sign, so that in the original frame the wire is neutral. However, in our new frame, the charge densities are

$$\rho_\pm = qn_\pm = q\gamma(v_\pm)n_0 = \left(1 \mp \frac{uv}{c^2}\right)\gamma(u)\gamma(v)\,qn_0$$

where you've got to do a little bit of algebra to get to the last result. Since $v_- > v_+$, we have $n_- > n_+$ and the wire carries negative charge. The overall net charge density in the new frame is

$$\rho' = qn' = q(n_+ - n_-) = -\frac{2uv}{c^2}\,\gamma(u)\,qn$$

But we know that a line of electric charge creates an electric field; we calculated it in (2.6); it is

$$E(r) = -\frac{2uv}{c^2}\,\frac{\gamma(u)\,qnA}{2\pi\epsilon_0 r}\hat{\mathbf{r}}$$

where r is the radial direction away from the wire. This means that, in its rest frame, the particle experiences a force

$$F' = -u\gamma(u)\,\frac{nAq^2v}{\pi\epsilon_0 c^2 r}$$

where the minus sign tells us that the force is towards the wire for $u > 0$. But if there's a force in one frame, there must also be a force in another. Transforming back to where we came from, we conclude that even when the wire is neutral there has to be a force

$$F = \frac{F'}{\gamma(u)} = -u\,\frac{nq^2Av}{\pi\epsilon_0 c^2 r} = -uq\,\frac{\mu_0 I}{2\pi r} \tag{5.12}$$

But this precisely agrees with the Lorentz force law, with the magnetic field given by the expression (3.5) that we computed for a straight wire. Notice that if $u > 0$ then the test particle – which has charge q – is moving in the same direction as the particles in the wire which have charge q and the force is attractive. If $u < 0$ then it moves in the opposite direction and the force is repulsive.

This analysis provides an explicit demonstration of how an electric force in one frame of reference is interpreted as a magnetic force in another. There's also something rather surprising about the result. We're used to thinking of length contraction as an exotic result which is only important when we approach the speed of light. Yet the electrons in a wire crawl along. They take around an hour to travel a meter! Nonetheless, we can easily detect the magnetic force between two wires which, as we've seen above, can be directly attributed to the length contraction in the electron density.

The discussion above needs a minor alteration for actual wires. In the rest frame of the wire the positive charges – which are ions, atoms stripped of some of their electrons – are stationary while the electrons move. Following the explanation above, you might think that there is an imbalance of charge density already in this frame. But that's not correct. The current is due to some battery feeding electrons into the wire and taking them out the other end. And this is done in such a way that the wire is neutral in the rest frame, with the electron density exactly compensating the ion density. In contrast, if we moved to a frame in which the ions and electrons had equal and opposite speeds, the wire would appear charged. Although the starting point is slightly different, the end result remains.

5.3 Gauge Potentials and the Electromagnetic Tensor

Under Lorentz transformations, electric and magnetic fields will transform into each other. In this section, we want to understand more precisely how this happens. At first sight, it looks as if it's going to be tricky. So far the objects which have nice transformation properties under Lorentz transformations are 4-vectors. But here we've got two 3-vectors, $\mathbf{E}$ and $\mathbf{B}$. How do we make those transform into each other?

5.3.1 Gauge Invariance and Relativity

To get an idea for how this happens, we first turn to some objects that we met previously: the scalar and vector potentials ϕ and $\mathbf{A}$. Recall that we introduced these to solve some of the equations of electrostatics and magnetostatics,

$$\begin{aligned} \nabla \times \mathbf{E} = 0 \quad &\Rightarrow \quad \mathbf{E} = -\nabla\phi \\ \nabla \cdot \mathbf{B} = 0 \quad &\Rightarrow \quad \mathbf{B} = \nabla \times \mathbf{A} \end{aligned}$$

However, in general these expressions can't be correct. We know that when $\mathbf{B}$ and $\mathbf{E}$ change with time, the two source-free Maxwell equations are

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \quad \text{and} \quad \nabla \cdot \mathbf{B} = 0$$

Nonetheless, it's still possible to use the scalar and vector potentials to solve both of these equations. The solutions are

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A}$$

where now $\phi = \phi(\mathbf{x}, t)$ and $\mathbf{A} = \mathbf{A}(\mathbf{x}, t)$.

Just as we saw before, there is no unique choice of ϕ and $\mathbf{A}$. We can always shift $\mathbf{A} \to \mathbf{A} + \nabla\chi$ and $\mathbf{B}$ remains unchanged. However, now this requires a compensating shift of ϕ.

$$\phi \to \phi - \frac{\partial \chi}{\partial t} \quad \text{and} \quad \mathbf{A} \to \mathbf{A} + \nabla\chi \tag{5.13}$$

with $\chi = \chi(\mathbf{x}, t)$. These are *gauge transformations*. They reproduce our earlier gauge transformation for $\mathbf{A}$, while also encompassing constant shifts in ϕ.

How does this help with our attempt to reformulate electromagnetism in a way compatible with special relativity? Well, now we have a scalar, and a 3-vector: these are ripe to place in a 4-vector. We define

$$A^\mu = \begin{pmatrix} \phi/c \\ \mathbf{A} \end{pmatrix}$$

Or, equivalently, $A_\mu = (\phi/c, -\mathbf{A})$. In this language, the gauge transformations (5.13) take a particularly nice form,

$$A_\mu \to A_\mu - \partial_\mu \chi \tag{5.14}$$

where χ is any function of space and time

5.3.2 The Electromagnetic Tensor

We now have all the ingredients necessary to determine how the electric and magnetic fields transform. From the 4-derivative $\partial_\mu = (\partial/\partial(ct), \nabla)$ and the 4-vector $A_\mu = (\phi/c, -\mathbf{A})$, we can form the anti-symmetric tensor

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

This is constructed to be invariant under gauge transformations (5.14). We have

$$F_{\mu\nu} \to F_{\mu\nu} + \partial_\mu\partial_\nu\chi - \partial_\nu\partial_\mu\chi = F_{\mu\nu}$$

This already suggests that the components involve the $\mathbf{E}$ and $\mathbf{B}$ fields. To check that this is indeed the case, we can do a few small computations,

$$F_{01} = \frac{1}{c}\frac{\partial(-A_x)}{\partial t} - \frac{\partial(\phi/c)}{\partial x} = \frac{E_x}{c}$$

and

$$F_{12} = \frac{\partial(-A_y)}{\partial x} - \frac{\partial(-A_x)}{\partial y} = -B_z$$

Similar computations for all other entries give us a matrix of electric and magnetic fields,

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix} \tag{5.15}$$

This, then, is the answer to our original question. You can make a Lorentz covariant object consisting of two 3-vectors by arranging them in an anti-symmetric tensor. $F_{\mu\nu}$ is called the *electromagnetic tensor*. Equivalently, we can raise both indices using the Minkowski metric to get

$$F^{\mu\nu} = \eta^{\mu\rho}\eta^{\nu\sigma}F_{\rho\sigma} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}$$

Both $F_{\mu\nu}$ and $F^{\mu\nu}$ are tensors. They are tensors because they're constructed out of objects, A_μ, ∂_μ and $\eta_{\mu\nu}$, which themselves transform nicely under the Lorentz group. This means that the field strength must transform as

$$F'^{\mu\nu} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F^{\rho\sigma} \tag{5.16}$$

Alternatively, if you want to get rid of the indices, this reads $F' = \Lambda F \Lambda^T$. The observer in a new frame sees electric and magnetic fields $\mathbf{E}'$ and $\mathbf{B}'$ that differ from the original observer. The two are related by (5.16). Let's look at what this means in a couple of illustrative examples.

Rotations

To compute the transformation (5.16), it's probably simplest to just do the sums that are implicit in the repeated ρ and σ labels. Alternatively, if you want to revert to matrix multiplication then this is the same as $F' = \Lambda F \Lambda^T$. Either way, we get the same result. For a rotation, the 3×3 matrix R is embedded in the lower-right hand block of Λ as shown in (5.4). A quick calculation shows that the transformation of the electric and magnetic fields in (5.16) is as expected,

$$\mathbf{E}' = R\,\mathbf{E} \quad \text{and} \quad \mathbf{B}' = R\,\mathbf{B}$$

Boosts

Things are more interesting for boosts. Let's consider a boost v in the x-direction, with Λ given by (5.5). Again, you need to do a few short calculations. For example, we have

$$\begin{aligned}-\frac{E'_x}{c} = F'^{\,01} &= \Lambda^0{}_\rho \Lambda^1{}_\sigma F^{\rho\sigma} \\ &= \Lambda^0{}_0 \Lambda^1{}_1 F^{01} + \Lambda^0{}_1 \Lambda^1{}_0 F^{10} \\ &= \frac{\gamma^2 v^2}{c^2}\frac{E_x}{c} - \gamma^2 \frac{E_x}{c} = -\frac{E_x}{c}\end{aligned}$$

and

$$\begin{aligned}-\frac{E'_y}{c} = F'^{\,02} &= \Lambda^0{}_\rho \Lambda^2{}_\sigma F^{\rho\sigma} \\ &= \Lambda^0{}_0 \Lambda^2{}_2 F^{02} + \Lambda^0{}_1 \Lambda^2{}_2 F^{12} \\ &= -\gamma\frac{E_y}{c} + \frac{\gamma v}{c} B_z = -\frac{\gamma}{c}(E_y - vB_z)\end{aligned}$$

and

$$\begin{aligned}-B'_z = F'^{\,12} &= \Lambda^1{}_\rho \Lambda^2{}_\sigma F^{\rho\sigma} \\ &= \Lambda^1{}_0 \Lambda^2{}_2 F^{02} + \Lambda^1{}_1 \Lambda^2{}_2 F^{12} \\ &= \frac{\gamma v}{c^2} E_y - \gamma B_z = -\gamma(B_z - vE_y/c^2)\end{aligned}$$

The final result for the transformation of the electric field after a boost in the x-direction is

$$\begin{aligned}E'_x &= E_x \\ E'_y &= \gamma(E_y - vB_z) \\ E'_z &= \gamma(E_z + vB_y)\end{aligned} \tag{5.17}$$

and, for the magnetic field,

$$\begin{aligned}B'_x &= B_x \\ B'_y &= \gamma\left(B_y + \frac{v}{c^2}E_z\right) \\ B'_z &= \gamma\left(B_z - \frac{v}{c^2}E_y\right)\end{aligned} \tag{5.18}$$

As we anticipated above, what appears to be a magnetic field to one observer looks like an electric field to another, and vice versa.

Note that in the limit $v \ll c$, we have $\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}$ and $\mathbf{B}' = \mathbf{B}$. This can be thought of as the Galilean boost of electric and magnetic fields. We recognise $\mathbf{E}+\mathbf{v}\times\mathbf{B}$ as the combination that appears in the Lorentz force law. We'll return to this force in Section 5.4.1 where we'll see how it's compatible with special relativity.

5.3.3 An Example: A Boosted Line Charge

In Section 2.1.3, we computed the electric field due to a line with uniform charge density η per unit length. If we take the line to lie along the x-axis, we have (2.6)

$$\mathbf{E} = \frac{\eta}{2\pi\epsilon_0(y^2+z^2)}\begin{pmatrix}0\\ y\\ z\end{pmatrix} \tag{5.19}$$

Meanwhile, the magnetic field vanishes for static electric charges: $\mathbf{B} = 0$. Let's see what this looks like from the perspective of an observer moving with speed v in the x-direction, parallel to the wire. In the moving frame the electric and magnetic fields are given by (5.17) and (5.18). These read

$$\begin{aligned}
\mathbf{E}' &= \frac{\eta\gamma}{2\pi\epsilon_0(y^2+z^2)}\begin{pmatrix}0\\ y\\ z\end{pmatrix} = \frac{\eta\gamma}{2\pi\epsilon_0(y'^2+z'^2)}\begin{pmatrix}0\\ y'\\ z'\end{pmatrix}\\
\mathbf{B}' &= \frac{\eta\gamma v}{2\pi\epsilon_0 c^2(y^2+z^2)}\begin{pmatrix}0\\ z\\ -y\end{pmatrix} = \frac{\eta\gamma v}{2\pi\epsilon_0 c^2(y'^2+z'^2)}\begin{pmatrix}0\\ z'\\ -y'\end{pmatrix}
\end{aligned} \tag{5.20}$$

In the second equality, we've rewritten the expression in terms of the coordinates of $\mathcal{S}'$ which, because the boost is in the x-direction, are trivial: $y = y'$ and $z = z'$.

From the perspective of an observer in frame $\mathcal{S}'$, the charge density in the wire is $\eta' = \gamma\eta$, where the factor of γ comes from Lorentz contraction. This can be seen in the expression above for the electric field. Since the charge density is now moving, the observer in frame $\mathcal{S}'$ sees a current $I' = -\gamma\eta v$. Then we can rewrite (5.20) as

$$\mathbf{B}' = \frac{\mu_0 I'}{2\pi\sqrt{y'^2+z'^2}}\hat{\boldsymbol{\varphi}}' \tag{5.21}$$

But this is something that we've seen before. It's the magnetic field due to a current in a wire (3.5). We computed this in Section 3.1.1 using Ampére's law. But here we've re-derived the same result without ever mentioning Ampére's law! Instead, our starting point (5.19) needed Gauss' law and we then used only the Lorentz transformation of electric and magnetic fields. We can only conclude that, under a Lorentz transformation, Gauss' law must be related to Ampére's law. Indeed, we'll shortly see explicitly that this is the case. For now, it's worth repeating the lesson that we learned in Section 5.2.1: the magnetic field can be viewed as a relativistic effect.

5.3.4 Another Example: A Boosted Point Charge

Consider a point charge Q, stationary in an inertial frame $\mathcal{S}$. We know that it's electric field is given by

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_0 r^2}\,\hat{\mathbf{r}} = \frac{Q}{4\pi\epsilon_0[x^2+y^2+z^2]^{3/2}}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

while its magnetic field vanishes. Now let's look at this same particle from the frame $\mathcal{S}'$, moving with velocity $\mathbf{v} = (v,0,0)$ with respect to $\mathcal{S}$. The Lorentz boost which relates the two is given by (5.5) and so the new electric field are given by (5.17),

$$\mathbf{E}' = \frac{Q}{4\pi\epsilon_0[x^2+y^2+z^2]^{3/2}}\begin{pmatrix} x \\ \gamma y \\ \gamma z \end{pmatrix}$$

But this is still expressed in terms of the original coordinates. We should now rewrite this in terms of the coordinates of $\mathcal{S}'$, which are $x' = \gamma(x - vt)$ and $y' = y$ and $z' = z$. Inverting these, we have

$$\mathbf{E}' = \frac{Q\gamma}{4\pi\epsilon_0[\gamma^2(x'+vt')^2 + y'^2 + z'^2]^{3/2}}\begin{pmatrix} x'+vt' \\ y' \\ z' \end{pmatrix} \tag{5.22}$$

In the frame $\mathcal{S}'$, the particle sits at $\mathbf{x}' = (-vt', 0, 0)$, so we see that the electric field emanates from the position of the charge, as it should. For now, let's look at the electric field when $t' = 0$ so that the particle sits at the origin in the new frame. The electric field points outwards radially, along the direction

$$\mathbf{r}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

However, the electric field is not isotropic. This arises from the denominator of (5.22) which is not proportional to r'^3 because there's an extra factor of γ^2 in front of the x' component. Instead, at $t' = 0$, the denominator involves the combination

$$\begin{aligned}
\gamma^2 x'^2 + y'^2 + z'^2 &= (\gamma^2 - 1)x'^2 + \mathbf{r}'^2 \\
&= \frac{v^2\gamma^2}{c^2}x'^2 + \mathbf{r}'^2 \\
&= \left(\frac{v^2\gamma^2}{c^2}\cos^2\theta + 1\right)\mathbf{r}'^2 \\
&= \gamma^2\left(1 - \frac{v^2}{c^2}\sin^2\theta\right)\mathbf{r}'^2
\end{aligned}$$

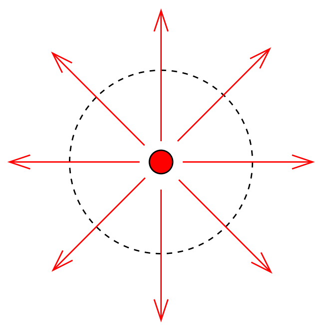

Figure 48: The isotropic field lines of a static charge

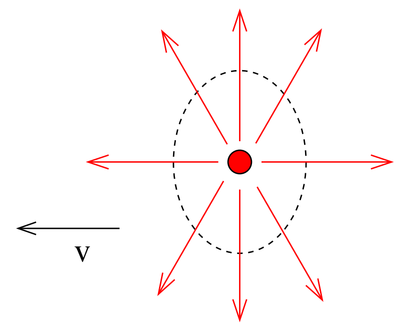

Figure 49: The squeezed field lines of a moving charge

where the θ is the angle between $\mathbf{r}'$ and the x'-axis and, in the last line, we've just used some simple trig and the definition of $\gamma^2 = 1/(1 - v^2/c^2)$. This means that we can write the electric field in frame S' as

$$\mathbf{E}' = \frac{1}{\gamma^2(1 - v^2 \sin^2\theta/c^2)^{3/2}} \frac{Q}{4\pi\epsilon_0 r'^2} \hat{\mathbf{r}}'$$

The pre-factor is responsible for the fact that the electric field is not isotropic. We see that it reduces the electric field along the x'-axis (i.e when $\theta = 0$) and increases the field along the perpendicular y' and z' axes (i.e. when $\theta = \pi/2$). This can be thought of as a consequence of Lorentz contraction, squeezing the electric field lines in the direction of travel.

The moving particle also gives rise to a magnetic field. This is easily computed using the Lorentz transformations (5.18). It is

$$\mathbf{B} = \frac{\mu_0 Q\gamma v}{4\pi[\gamma^2(x' + vt')^2 + y'^2 + z'^2]^{3/2}} \begin{pmatrix} 0 \\ z' \\ -y' \end{pmatrix}$$

5.3.5 Lorentz Scalars

We can now ask a familiar question: is there any combination of the electric and magnetic fields that all observers agree upon? Now we have the power of index notation at our disposal, this is easy to answer. We just need to write down an object that doesn't have any floating μ or ν indices. Unfortunately, we don't get to use the obvious choice of $\eta_{\mu\nu}F^{\mu\nu}$ because this vanishes on account of the anti-symmetry of $F^{\mu\nu}$. The simplest thing we can write down is

$$\frac{1}{2}F_{\mu\nu}F^{\mu\nu} = -\frac{\mathbf{E}^2}{c^2} + \mathbf{B}^2 \tag{5.23}$$

Note the relative minus sign between $\mathbf{E}$ and $\mathbf{B}$, mirroring a similar minus sign in the spacetime interval.

However, this isn't the only Lorentz scalar that we can construct from $\mathbf{E}$ and $\mathbf{B}$. There is another, somewhat more subtle, object. To build this, we need to appreciate that Minkowski spacetime comes equipped with another natural tensor object, beyond the familiar metric $\eta_{\mu\nu}$. This is the fully anti-symmetric object known as the *alternating tensor*,

$$\epsilon^{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{if } \mu\nu\rho\sigma \text{ is an even permutation of } 0123 \\ -1 & \text{if } \mu\nu\rho\sigma \text{ is an odd permutation of } 0123 \end{cases}$$

while $\epsilon^{\mu\nu\rho\sigma} = 0$ if there are any repeated indices.

To see why this is a natural object in Minkowski space, let's look at how it changes under Lorentz transformations. The usual tensor transformation is

$$\epsilon'^{\mu\nu\rho\sigma} = \Lambda^\mu{}_\kappa \Lambda^\nu{}_\lambda \Lambda^\rho{}_\alpha \Lambda^\sigma{}_\beta \epsilon^{\kappa\lambda\alpha\beta}$$

It's simple to check that $\epsilon'^{\mu\nu\rho\sigma}$ is also full anti-symmetric; it inherits this property from $\epsilon^{\kappa\lambda\alpha\beta}$ on the right-hand side. But this means that $\epsilon'^{\mu\nu\rho\sigma}$ must be proportional to $\epsilon^{\mu\nu\rho\sigma}$. We only need to determine the constant of proportionality. To do this, we can look at

$$\epsilon'^{0123} = \Lambda^0{}_\kappa \Lambda^1{}_\lambda \Lambda^2{}_\alpha \Lambda^3{}_\beta \epsilon^{\kappa\lambda\alpha\beta} = \det(\Lambda)$$

Now any Lorentz transformations have $\det(\Lambda) = \pm 1$. Those with $\det(\Lambda) = 1$ make up the "proper Lorentz group" $SO(1,3)$. (This was covered in the *Dynamics and Relativity* notes). These proper Lorentz transformations do not include reflections or time reversal. We learn that the alternating tensor $\epsilon^{\mu\nu\rho\sigma}$ is invariant under proper Lorentz transformations. What it's really telling us is that Minkowski space comes with an oriented orthonormal basis. By lowering indices with the Minkowski metric, we can also construct the tensor $\epsilon_{\mu\nu\rho\sigma}$ which has $\epsilon_{0123} = -1$.

The alternating tensor allows us to construct a second tensor field, sometimes called the *dual electromagnetic tensor* (although "dual" is perhaps the most overused word in physics),

$$\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z/c & -E_y/c \\ B_y & -E_z/c & 0 & E_x/c \\ B_z & E_y/c & -E_x/c & 0 \end{pmatrix} \tag{5.24}$$

$\tilde{F}^{\mu\nu}$ is sometimes also written as ${}^\star F^{\mu\nu}$. We see that this is looks just like $F^{\mu\nu}$ but with the electric and magnetic fields swapped around. Actually, looking closely you'll see that there's a minus sign difference as well: $\tilde{F}^{\mu\nu}$ arises from $F^{\mu\nu}$ by the substitution $\mathbf{E} \to c\mathbf{B}$ and $\mathbf{B} \to -\mathbf{E}/c$.

The statement that $\tilde{F}^{\mu\nu}$ is a tensor means that it too has nice properties under Lorentz transformations,

$$\tilde{F}'^{\,\mu\nu} = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}\tilde{F}^{\rho\sigma}$$

and we can use this to build new Lorentz invariant quantities. Taking the obvious square of $\tilde{F}$ doesn't give us anything new, since

$$\tilde{F}^{\mu\nu}\tilde{F}_{\mu\nu} = -F^{\mu\nu}F_{\mu\nu}$$

But by contracting $\tilde{F}$ with the original F we do find a new Lorentz invariant

$$\frac{1}{4}\tilde{F}^{\mu\nu}F_{\mu\nu} = -\frac{1}{c}\,\mathbf{E}\cdot\mathbf{B} \tag{5.25}$$

This tells us that the inner-product of $\mathbf{E}$ and $\mathbf{B}$ is the same viewed in all frames.

5.4 Maxwell Equations

We now have the machinery to write the Maxwell equations in a way which is manifestly compatible with special relativity. They take a particularly simple form:

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu \quad \text{and} \quad \partial_\mu \tilde{F}^{\mu\nu} = 0 \tag{5.26}$$

Pretty aren't they!

The Maxwell equations are not *invariant* under Lorentz transformations. This is because there is the dangling ν index on both sides. However, because the equations are built out of objects which transform nicely – $F^{\mu\nu}$, $\tilde{F}^{\mu\nu}$, J^μ and ∂_μ – the equations themselves also transform nicely. For example, we will see shortly that Gauss' law transforms into Ampére's law under a Lorentz boost, something we anticipated in Section 5.3.3. We say that the equations are *covariant* under Lorentz transformations.

This means that an observer in a different frame will mix everything up: space and time, charges and currents, and electric and magnetic fields. Although observers disagree on what these things are, they all agree on how they fit together. This is what it means for an equation to be covariant: the ingredients change, but the relationship between them stays the same. All observers agree that, in their frame, the electric and magnetic fields are governed by the same Maxwell equations.

Given the objects $F^{\mu\nu}$, $\tilde{F}^{\mu\nu}$, J^μ and ∂_μ, the Maxwell equations are not the only thing you could write down compatible with Lorentz invariance. But they are by far the simplest . Any other equation would be non-linear in F or $\tilde{F}$ or contain more derivative terms or some such thing. Of course, simplicity is no guarantee that equations are correct. For this we need experiment. But surprisingly often in physics we find that the simplest equations are also the right ones.

Unpacking the Maxwell Equations

Let's now check that the Maxwell equations (5.26) in relativistic form do indeed coincide with the vector calculus equations that we've been studying in this course. We just need to expand the different parts of the equation. The components of the first Maxwell equation give

$$\begin{aligned} \partial_i F^{i0} = \mu_0 J^0 \quad &\Rightarrow \quad \nabla\cdot\mathbf{E} = \frac{\rho}{\epsilon_0} \\ \partial_\mu F^{\mu i} = \mu_0 J^i \quad &\Rightarrow \quad -\frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} + \nabla\times\mathbf{B} = \mu_0 \mathbf{J} \end{aligned}$$

In the first equation, which arises from $\nu = 0$, we sum only over spatial indices $i = 1, 2, 3$ because $F^{00} = 0$. Meanwhile the components of the second Maxwell equation give

$$\begin{aligned} \partial_i \tilde{F}^{i0} = 0 \quad &\Rightarrow \quad \nabla\cdot\mathbf{B} = 0 \\ \partial_\mu \tilde{F}^{\mu i} = 0 \quad &\Rightarrow \quad \frac{\partial \mathbf{B}}{\partial t} + \nabla\times\mathbf{E} = 0 \end{aligned}$$

These, of course, are the familiar equations that we've all grown to love over this course.

Here a few further, simple comments about the advantages of writing the Maxwell equations in relativistic form. First, the Maxwell equations imply that current is conserved. This follows because $F^{\mu\nu}$ is anti-symmetric, so $\partial_\mu\partial_\nu F^{\mu\nu} = 0$ automatically, simply because $\partial_\mu\partial_\nu$ is symmetric. The first of the Maxwell equations (5.26) then requires that the continuity equation holds

$$\partial_\mu J^\mu = 0$$

This is the same calculation that we did in vector notation in Section 4.2.1. Note that it's marginally easier in the relativistic framework.

The second Maxwell equation can be written in a number of different ways. It is equivalent to:

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \quad \Leftrightarrow \quad \epsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma} = 0 \quad \Leftrightarrow \quad \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} + \partial_\mu F_{\nu\rho} = 0 \qquad (5.27)$$

where the last of these equalities follows because the equation is constructed so that it is fully anti-symmetric with respect to exchanging any of the indices ρ, μ and ν. (Just expand out for a few examples to see this).

The gauge potential A_μ was originally introduced to solve the two Maxwell equations which are contained in $\partial_\mu \tilde{F}^{\mu\nu} = 0$. Again, this is marginally easier to see in relativistic notation. If we write $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ then

$$\partial_\mu \tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}\partial_\mu F_{\rho\sigma} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}\partial_\mu(\partial_\rho A_\sigma - \partial_\sigma A_\rho) = 0$$

where the final equality holds because of the symmetry of the two derivatives, combined with the anti-symmetry of the ϵ-tensor. This means that we could equally well write the Maxwell equations as

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu \quad \text{where} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{5.28}$$

The first of these coincides with the first equation in (5.26); the second is an alternative way of writing the second equation in (5.26). In more advanced formulations of electromagnetism (for example, in the Lagrangian formulation), this is the form in which the Maxwell equations arise.

5.4.1 The Lorentz Force Law

There's one last aspect of electromagnetism that we need to show is compatible with relativity: the Lorentz force law. In the Newtonian world, the equation of motion for a particle moving with velocity $\mathbf{u}$ and momentum $\mathbf{p} = m\mathbf{u}$ is

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \tag{5.29}$$

We want to write this equation in 4-vector notation in a way that makes it clear how all the objects change under Lorentz transformations.

By now it should be intuitively clear how this is going to work. A moving particle experiences the magnetic force. But if we boost to its rest frame, there is no magnetic force. Instead, the magnetic field transforms into an electric field and we find the same force, now interpreted as an electric force.

The relativistic version of (5.29) involves the 4-momentum P^μ, defined in (5.6), the proper time τ, reviewed in Section 5.1.2, and our new friend the electromagnetic tensor $F^{\mu\nu}$. The electromagnetic force acting on a point particle of charge q can then be written as

$$\frac{dP^\mu}{d\tau} = q\, F^{\mu\nu} U_\nu \tag{5.30}$$

where the 4-velocity is

$$U^\mu = \frac{dX^\mu}{d\tau} = \gamma \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \tag{5.31}$$

and the 4-momentum is $P = mU$. Again, we see that the relativistic form of the equation (5.30) is somewhat prettier than the original equation (5.29).

Unpacking the Lorentz Force Law

Let's check to see that the relativistic equation (5.30) is giving us the right physics. It is, of course, four equations: one for each $\mu = 0, 1, 2, 3$. It's simple to multiply out the right-hand side, remembering that U_μ comes with an extra minus sign in the spatial components relative to (5.31). We find that the $\mu = 1, 2, 3$ components of (5.30) arrange themselves into a familiar vector equation,

$$\frac{d\mathbf{p}}{d\tau} = q\gamma(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \quad \Rightarrow \quad \frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \tag{5.32}$$

where we've used the relationship $dt/d\tau = \gamma$. We find that we recover the Lorentz force law. Actually, there's a slight difference from the usual Newtonian force law (5.29), although the difference is buried in our notation. In the Newtonian setting, the momentum is $\mathbf{p} = m\mathbf{u}$. However, in the relativistic setting above, the momentum is $\mathbf{p} = m\gamma\mathbf{u}$. Needless to say, the relativistic version is correct, although the difference only shows up at high speeds.

The relativistic formulation of the Lorentz force (5.30) also contains an extra equation coming from $\mu = 0$. This reads

$$\frac{dP^0}{d\tau} = \frac{q}{c}\gamma\, \mathbf{E} \cdot \mathbf{u} \tag{5.33}$$

Recall that the temporal component of the four-momentum is the energy $P^0 = E/c$. Here the energy is $E = m\gamma c^2$ which includes both the rest-mass of the particle and its kinetic energy. The extra equation in (5.30) is simply telling us that the kinetic energy increases when work is done by an electric field

$$\frac{d(\text{Energy})}{dt} = q\mathbf{E} \cdot \mathbf{u}$$

where I've written energy as a word rather than as E to avoid confusing it with the electric field $\mathbf{E}$.

5.4.2 Motion in Constant Fields

We already know how electric and magnetic fields act on particles in a Newtonian world. Electric fields accelerate particles in straight lines; magnetic fields make particles go in circles. Here we're going to redo this analysis in the relativistic framework. The Lorentz force law remains the same. The only difference is that momentum is now $\mathbf{p} = m\gamma\mathbf{u}$. We'll see how this changes things.

Constant Electric Field

Consider a vanishing magnetic field and constant electric field $\mathbf{E} = (E, 0, 0)$. (Note that E here denotes electric field, not energy!). The equation of motion (5.32) for a charged particle with velocity $\mathbf{u} = (u, 0, 0)$ is

$$m\frac{d(\gamma u)}{dt} = qE \quad \Rightarrow \quad m\gamma u = qEt$$

where we've implicitly assumed that the particle starts from rest at $t = 0$. Rearranging, we get

$$u = \frac{dx}{dt} = \frac{qEt}{\sqrt{m^2 + q^2E^2t^2/c^2}}$$

Reassuringly, the speed never exceeds the speed of light. Instead, $u \to c$ as $t \to \infty$ as one would expect. It's simple to integrate this once more. If the particle starts from the origin, we have

$$x = \frac{mc^2}{qE}\left(\sqrt{1 + \frac{q^2E^2t^2}{m^2c^2}} - 1\right)$$

For early times, when the speeds are not too high, this reduces to

$$mx \approx \frac{1}{2}qEt^2 + \ldots$$

which is the usual non-relativistic result for particles undergoing constant acceleration in a straight line.

Constant Magnetic Field

Now let's turn the electric field off and look at the case of constant magnetic field $\mathbf{B} = (0, 0, B)$. In the non-relativistic world, we know that particles turn circles with frequency $\omega = qB/m$. Let's see how relativity changes things.

We start by looking at the zeroth component of the force equation (5.33) which, in the absence of an electric field, reads

$$\frac{dP^0}{d\tau} = 0$$

This tells us that magnetic fields do no work. We knew this from our course on Newtonian physics, but it remains true in the relativistic context. So we know that energy, $E = m\gamma c^2$, is constant. But this tells us that the speed (i.e. the magnitude of the velocity) remains constant. In other words, the velocity, and hence the position, once again turn circles. The equation of motion is now

$$m\frac{d(\gamma \mathbf{u})}{dt} = q\mathbf{u} \times \mathbf{B}$$

Since γ is constant, the equation takes the same form as in the non-relativistic case and the solutions are circles (or helices if the particle also moves in the z-direction). The only difference is that the frequency with which the particle moves in a circle now depends on how fast the particle is moving,

$$\omega = \frac{qB}{m\gamma}$$

If you wanted, you could interpret this as due to the relativistic increase in the mass of a moving particle. Naturally, for small speeds $\gamma \approx 1$ and we reproduce the more familiar cyclotron frequency $\omega \approx qB/m$.

So far we have looked at situations in which $\mathbf{E} = 0$ and in which $\mathbf{B} = 0$. But we've seen that $\mathbf{E}\cdot\mathbf{B} = 0$ and $\mathbf{E}^2 - \mathbf{B}^2$ are both Lorentz invariant quantities. This means that the solutions we've described above can be boosted to apply to any situation where $\mathbf{E} \cdot \mathbf{B} = 0$ and $\mathbf{E}^2 - \mathbf{B}^2$ is either > 0 or < 0. In the general situation, both electric and magnetic fields are turned on so $\mathbf{E} \cdot \mathbf{B} \neq 0$ and we have three possibilities to consider depending on whether $\mathbf{E}^2 - \mathbf{B}^2$ is > 0 or < 0 or $= 0$.

5.5 ...and Action

The principle of least action provides an elegant and powerful way to think about the classical mechanics of particles. In this section we will see that the action principle can also be used to describe classical fields.

5.5.1 Non-Relativistic Particles

The principle of least action was described in some detail in the lectures on Classical Dynamics. For a particle moving along a trajectory $\mathbf{x}(t)$, subject to the potential $V(\mathbf{x})$, the action is given by

$$S[\mathbf{x}(t)] = \int_{t_1}^{t_2} dt\ \left[\frac{1}{2}m\dot{\mathbf{x}}^2 - V(\mathbf{x})\right] \tag{5.34}$$

We fix the position of the particle at time t_1 and t_2. The principle of least action says that when the particle moves between these two points, it takes a path that extremises the value of the action.

It is simple to show that the principle of least action is equivalent to Newtonian equation of motion. We vary the path, $\mathbf{x}(t) \to \mathbf{x}(t) + \delta\mathbf{x}(t)$, subject to the requirement that $\delta\mathbf{x}(t_1) = \delta\mathbf{x}(t_2) = 0$ so that the end points are fixed. The change in the action is then

$$\begin{aligned}\delta S &= \int_{t_1}^{t_2} dt\ \Big[m\dot{\mathbf{x}}\cdot\delta\dot{\mathbf{x}} - \nabla V\cdot\delta\mathbf{x}\Big] \\ &= \int_{t_1}^{t_2} dt\ \Big[-m\ddot{\mathbf{x}} - \nabla V\Big]\cdot\delta\mathbf{x} + \Big[m\dot{\mathbf{x}}\cdot\delta\mathbf{x}\Big]_{t_1}^{t_2}\end{aligned}$$

where the second line follows after integration by parts. The boundary term vanishes because the end points are fixed. The path $\mathbf{x}(t)$ extremises the action if $\delta S = 0$ for all variations $\delta\mathbf{x}(t)$. This holds only if if the

$$m\ddot{\mathbf{x}} = -\nabla V$$

which we recognise as the Newtonian equation of motion.

For this course, we're interested in writing down the action for a particle of charge q interacting with electric and magnetic fields. This is written in terms of the potential $\phi(\mathbf{x})$ and the vector potential $\mathbf{A}(\mathbf{x})$. It is

$$S[\mathbf{x}(t)] = \int_{t_1}^{t_2} dt\ \left[\frac{1}{2}m\dot{\mathbf{x}}^2 - q\phi(\mathbf{x}) + q\dot{\mathbf{x}}\cdot\mathbf{A}(\mathbf{x})\right] \tag{5.35}$$

We will now show that this reproduces the Lorentz force law. The electric term involving ϕ is just of the usual potential energy type and the fact it gives the right equation of motion follows immediately from the definition of the electric field $\mathbf{E} = -\nabla\phi$. Meanwhile, we have a short calculation to do for the the magnetic force. It is a calculation that is best done in index notation

$$\begin{aligned}
\delta \int_{t_1}^{t_2} dt \;\; \left[\dot{x}^i A^i(\mathbf{x})\right] &= \int_{t_1}^{t_2} dt \;\; \left[\delta \dot{x}^i A^i(x) + \dot{x}^i \delta A^i(x)\right] \\
&= \int_{t_1}^{t_2} dt \;\; \left[-\delta x^i \frac{\partial A^i}{\partial t} + \dot{x}^i \frac{\partial A^i}{\partial x^j} \delta x^j\right] \\
&= \int_{t_1}^{t_2} dt \;\; \left[-\delta x^i \frac{\partial A^i}{\partial x^j} \dot{x}^j + \dot{x}^i \frac{\partial A^i}{\partial x^j} \delta x^j\right] \\
&= \int_{t_1}^{t_2} dt \;\; \left[-\frac{\partial A^i}{\partial x^j} + \frac{\partial A^j}{\partial x^i}\right] \dot{x}^j \delta x^i \\
&= \int_{t_1}^{t_2} dt \; \epsilon_{ijk} \dot{x}^i \delta x^j B^k = \int_{t_1}^{t_2} dt \; (\dot{\mathbf{x}} \times \mathbf{B}) \cdot \delta \mathbf{x}
\end{aligned}$$

where, in the second line, we've integrated by parts and thrown away the boundary term and, in the third line, we've relabelled the indices in the second term. In the final line, we've used the definition of the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$. The net result is that varying the action (5.35) indeed reproduces the Lorentz force law

$$m\ddot{\mathbf{x}} = q(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B})$$

There's something interesting about the action (5.35). The potentials ϕ and $\mathbf{A}$ have been our constant companions throughout these lectures but, until now, they've only played an auxiliary role. They were useful in helping us solve the Maxwell equations. But they weren't necessary. At any stage, we could have worked just with $\mathbf{E}$ and $\mathbf{B}$ and not worried about the underlying potentials. That's no longer true when we turn to the Lagrangian formulation. There's no Lagrangian formulation of electromagnetism that involves only $\mathbf{E}$ and $\mathbf{B}$. Instead, you're obliged to use the potentials ϕ and $\mathbf{A}$. This is true for both the point particle action (5.35) and for the action that we'll meet shortly that leads to the Maxwell equations.

Whenever some mathematical object is written in terms of ϕ and $\mathbf{A}$, some minor alarm bells should start to ring. This is because the are not unique functions, but are defined only up to gauge transformations (5.13).

$$\phi \to \phi - \frac{\partial \chi}{\partial t} \quad \text{ and } \quad \mathbf{A} \to \mathbf{A} + \nabla \chi$$

with $\chi = \chi(\mathbf{x}, t)$. Anything physical should not depend on the choice of χ. This is true for the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$. Happily, it is also true for the action (5.35). Under a gauge transformation, this shifts as

$$S \to S + q \int_{t_1}^{t_2} dt \;\; \left[\frac{\partial \chi}{\partial t} + \dot{\mathbf{x}} \cdot \nabla \chi\right] = S + q \int_{t_1}^{t_2} dt \; \frac{d\chi}{dt}$$

We see that the change of the action is a total time derivative. But we know from the lectures on Classical Dynamics that adding a total derivative to the action doesn't change the physics.

5.5.2 Relativistic Particles

Our first task is to write down an action for a relativistic particle. As we'll see, there are two ways to do this; the first is simpler, but the second is better.

A simple action for a relativistic particle is

$$S[\mathbf{x}(t)] = -mc^2 \int dt\, \sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}} \tag{5.36}$$

First note that if we Taylor expand the action, we get back the familiar Newtonian action (5.34) for a free particle. More importantly, the canonical momentum associated to the action is

$$\mathbf{p} = \frac{\partial S}{\partial \dot{\mathbf{x}}} = m\gamma\dot{\mathbf{x}}$$

where $\gamma = (1 - \dot{\mathbf{x}}^2/c^2)^{-1/2}$ is the usual relativistic gamma factor. This then gives us the right equation of motion for a free relativistic particle,

$$\frac{d\mathbf{p}}{dt} = 0$$

It's straightforward to couple this particle to electric and magnetic fields: we just include the same terms that we saw in (5.35),

$$S[\mathbf{x}(t)] = \int dt\, \left(-mc^2\sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}} - q\phi + q\dot{\mathbf{x}}\cdot\mathbf{A} \right) \tag{5.37}$$

Although this action gives the right relativistic equations of motion, there's something more than a little unsatisfactory about it. This is because it gives the equations of a motion is a very particular reference frame, with a very particular choice of time coordinate t. And that's not really in the spirit of special relativity. Indeed, the essence of Minkowski space is that time and space sit on a very similar footing, with Lorentz transformations rotating the two. How can we write down an action that puts time and space on the same footing and manifestly exhibits invariance under Lorentz transformations?

The Covariant Action

As we will now see, the construction of such an action needs a new ingredient. To start with, we'll follow our nose. It's clear that if we want an action with manifest Lorentz invariance then we should work with the four-vector $X^\mu = (ct, \mathbf{x})$. The worldline of a particle is then some parameterised curve

$$X^\mu(\sigma)$$

with σ is a label that tells us where we sit on the curve. This four-vector will be our degree of freedom in constructing an action.

It's worth pausing to stress just how different our current situation is from the original form of the relativistic action (5.36). For (5.36), we constructed an action based on the path $\mathbf{x}(t)$, where $\mathbf{x}$ is the degree of freedom and the time t is used to parameterise the curve. But now we're going to construct an action based on $X^\mu(\sigma)$, which means that we've promoted time to a dynamical degree of freedom, sitting alongside $\mathbf{x}$. That's going to need some explaining. After all, the number of degrees of freedom is one of the crudest ways we have to describe a system and usually if we add an extra degree of freedom, we're going to be describing something rather different. But here we're not aiming at describing the same physical system as (5.36) – a relativistic particle – just with the symmetries manifest.

Relatedly, we have now introduced a different parameter σ that describes where we sit on the worldline $X^\mu(\sigma)$. What choice of σ should we take? For now we'll just let σ be any parameterisation that we like. We'll soon see that, in fact, this is more or less the right answer!

If we're working with $X^\mu(\sigma)$ as our degree of freedom, it's straightforward to construct an action that exhibits Lorentz invariance. The one that works turns out to be

$$S[X^\mu(\sigma)] = -mc \int_{\sigma_1}^{\sigma_2} d\sigma \; \sqrt{\eta_{\mu\nu} \frac{dX^\mu}{d\sigma} \frac{dX^\nu}{d\sigma}} \tag{5.38}$$

The coefficients in front ensure that the action has dimensions $[S] = \text{Energy} \times \text{Time}$ as it should. We see immediately that this action is invariant under Lorentz transformations $X^\mu \to \Lambda^\mu{}_\nu X^\nu$ that we saw earlier in (5.8). This follows just because the integrand is a tensor with the μ, ν indices contracted correctly. For this reason, (5.38) is known as the *covariant action*.

The action S is actually closely related to something familiar from the world of special relativity: it is proportional to the proper time experienced by the particle. Recall that a particle moving along a worldline $X^\mu(\sigma)$, experience a proper time

$$\tau(\sigma) = \frac{1}{c}\int_\sigma^0 d\sigma' \,\sqrt{\eta_{\mu\nu}\frac{dX^\mu}{d\sigma'}\frac{dX^\nu}{d\sigma'}} \tag{5.39}$$

In special relativity, the proper time is maximised by a particle that does *not* accelerate. This fact is famous from the twin paradox where the dull stay-at-home twin ages fastest. Here it sits nicely with the fact that the proper time is identified with the action and hence is extremised on solutions to the equations of motion.

In addition to Lorentz invariance, the action (5.38) has a second symmetry of a very different kind, and this is the key to understanding the issues that we raised above. This second symmetry is *reparameterisation invariance*. Suppose that we pick a different parameterisation of the path, $\tilde{\sigma}$, related to the first parameterisation by a monotonic function $\tilde{\sigma}(\sigma)$. Then we could equally as well construct an action $\tilde{S}$ using this new parameter, given by

$$\tilde{S} = -mc\int_{\tilde{\sigma}_1}^{\tilde{\sigma}_2} d\tilde{\sigma}\,\sqrt{\eta_{\mu\nu}\frac{dX^\mu}{d\tilde{\sigma}}\frac{dX^\nu}{d\tilde{\sigma}}}$$

We might worry that this different parameterisation will give different equations of motion. Happily this is not the case because the two actions are, in fact, identical

$$\tilde{S} = -mc\int_{\sigma_1}^{\sigma_2} d\sigma\,\frac{d\tilde{\sigma}}{d\sigma}\sqrt{\eta_{\mu\nu}\frac{dX^\mu}{d\sigma}\frac{dX^\nu}{d\sigma}\left(\frac{d\sigma}{d\tilde{\sigma}}\right)^2} = S$$

We see that the action takes the same form regardless of our choice of parameterisation. Although we've called this a "symmetry", it's not a symmetry in the same sense as Lorentz transformations. In particular, reparameterisation does not generate new solutions from old ones. Instead, it is a redundancy in the way we describe the system. It is similar to the gauge "symmetry" of Maxwell and Yang-Mills theory which, despite the name, is also a redundancy rather than a symmetry.

Reparameterisation invariance has a number of consequences. The first is that it explains why the action (5.38) has only three degrees of freedom, even though it is a function of four variables $X^\mu(\sigma)$. This is because one of the degrees of freedom X^μ is not physical. Suppose that you solve the equation of motion to find a trajectory $X^\mu(\sigma)$. In most dynamical systems, each of these four functions would tell you something about the physical trajectory. But, for us, reparameterisation invariance means that there is no actual information in the value of σ. To find the physical path, we should eliminate σ to find the relationship between the X^μ. And this kills one degree of freedom.

We can see this most clearly by making a cunning choice for the parameter σ that parameterises the worldline. Suppose that we choose σ to coincide with the time t for some intertial observer: $\sigma = t$. Then $dX^0/d\sigma = c$ and the action (5.38) then becomes

$$S = -mc^2 \int_{t_1}^{t_2} dt \sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}}$$

where here $\dot{\mathbf{x}} = d\mathbf{x}/dt$. But this is the action (5.36) that we started this section with. So our two actions (5.38) and (5.36) are indeed equivalent, but each has different advantages. The action (5.36) makes it clear that we are dealing with a system with three degrees of freedom $\mathbf{x}$, but Lorentz invariance is hidden. Meanwhile the action (5.38) has manifest Lorentz invariance, but at the cost of introducing more degrees of freedom than are physical. But, as we've seen above, the reparameterisation invariance of the action allows us to remove the time degree of freedom and return to (5.36).

There's yet another manifestation of reparameterisation invariance. To see this, we compute the canonical momentum associated to X^μ,

$$P_\mu = \frac{\partial L}{\partial \dot{X}^\mu} = -mc \frac{1}{\sqrt{\dot{X}^\nu \dot{X}_\nu}} \dot{X}_\mu$$

where here $\dot{X}^\mu = \partial X^\mu / \partial \sigma$. You can check that P^μ above coincides with the four-momentum $P^\mu = m dX^\mu/d\tau$ that we defined previously in (5.6). (This follows from the fact that the proper time τ defined by (5.39) $d\tau/d\sigma = L/mc^2$ with L the Lagrangian.) It's a familiar result from special relativity that these momenta are not all independent, but obey

$$P^\mu P_\mu = m^2 c^2 \tag{5.40}$$

While this result is familiar in special relativity, it's rather surprising from the perspective of Lagrangian mechanics. This novel feature can be traced to the existence of reparameterisation invariance, meaning that there was a redundancy in our original description. Indeed, whenever theories have such a redundancy there will be some constraint analogous to (5.40). As one final comment, note that if we expand out (5.40), we have

$$(P^0)^2 = \mathbf{p}^2 + m^2 c^2$$

In particular, we see that we must have $P^0 \neq 0$. This is important. There's nothing that tells us that we must have $\mathbf{p} \neq 0$. The particle is quite able to just sit still in space if it wants. But $P^0 \neq 0$ tells us that the particle is obliged to move in the time direction. Physically, this again reflects the fact that the action (5.38) has only three degrees of freedom, not four. Physiologically, this is why you get old.

Finally, we can couple the covariant action (5.38) to electromagnetism. We do this by introducing the gauge field four-vector $A^\mu = (\phi/c, \mathbf{A})$ and extend the action (5.38) to

$$S[X^\mu(\sigma)] = \int_{\sigma_1}^{\sigma_2} d\sigma \left[-mc\sqrt{\eta_{\mu\nu}\frac{dX^\mu}{d\sigma}\frac{dX^\nu}{d\sigma}} - qA_\mu(X)\frac{dX^\mu}{d\sigma}\right] \tag{5.41}$$

If we again pick the worldline parameter σ to coincide with the time of some inertial observer, $\sigma = t$, then we again find that this action coincides with our earlier result (5.37).

5.5.3 The Maxwell Action

Our next goal is to write down an action principle for the Maxwell equations. Again we need a change of perspective which, this time, is just the usual shift from thinking about particles to thinking about fields. The action associates a number S to every field configuration $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{B}(\mathbf{x}, t)$. We will show that the action that reproduces the Maxwell equation takes the beautifully compact form

$$S[A_\mu(\mathbf{x}, t)] = -\frac{1}{4\mu_0 c}\int d^4x\; F_{\mu\nu}F^{\mu\nu} \tag{5.42}$$

Before we compute the equations of motion, here are a number of comments.

- The action is Lorentz invariant. This is true both of the integrand $F_{\mu\nu}F^{\mu\nu}$ and the measure $d^4x = cdtd^3x$. Under a Lorentz transformation (5.8), the measure picks up a Jacobian factor $\det\Lambda = 1$.

- For a non-relativistic particle, the action takes the form of "kinetic energy minus potential energy". But there is a similar interpretation of the Maxwell action (5.42). Expanding out the integrand using (5.23), we have

  $$S = \int dtd^3x\; \left(\frac{\epsilon_0}{2}\mathbf{E}^2 - \frac{1}{2\mu_0}\mathbf{B}^2\right)$$

  Comparing to the energy stored in electric and magnetic fields that we derived in (4.3), we see that $\mathbf{E}^2$ is like the kinetic energy, while $\mathbf{B}^2$ is like the potential energy.

- As we can see, the action depends on the electric field $\mathbf{E}(\mathbf{x}, t)$ and magnetic field $\mathbf{B}(\mathbf{x}, t)$. Nonetheless, the action should be viewed as a functional of the underlying gauge field $A_\mu(\mathbf{x}, t)$, albeit one that is invariant under gauge transformations $A_\mu \to A_\mu - \partial_\mu\chi$. This mirrors what we saw for the action for the Lorentz force law (5.35) where we were also obliged to introduce the scalar and vector potentials. The need to view the Maxwell action (5.42) as functional of the gauge potential A_μ is reflected in the fact that we should vary with respect to A_μ, rather than $\mathbf{E}$ or $\mathbf{B}$, when deriving the equations of motion. This is what we do next.

We vary the action by considering a neighbouring configuration $A_\mu + \delta A_\mu$. Using the definition of the electromagnetic tensor is $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, the change in the action is

$$\begin{aligned}\delta S &= -\frac{1}{4\mu_0}\int d^4x\ 2\left(\partial_\mu \delta A_\nu - \partial_\nu \delta A_\mu\right) F^{\mu\nu} \\ &= -\frac{1}{\mu_0}\int d^4x\ F^{\mu\nu}\partial_\mu \delta A_\nu \\ &= \frac{1}{\mu_0}\int d^4x\ (\partial_\mu F^{\mu\nu})\delta A_\nu\end{aligned}$$

where, as usual, we have discarded the total derivative term after integrating by parts. We see that the principle of least action, $\delta S = 0$, gives the vacuum Maxwell equations

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

Note that we only get half the Maxwell equations from the variation of the action. The other half, $\partial_\mu \tilde{F}^{\mu\nu} = 0$ follow immediately from working with the gauge potential A_μ.

The action (5.42) gives the vacuum Maxwell equations. If we have some fixed current J^μ, we can modify the action to read

$$S[A_\mu] = \frac{1}{c}\int d^4x\ \left(-\frac{1}{4\mu_0}F^{\mu\nu}F_{\mu\nu} - A_\mu J^\mu\right) \tag{5.43}$$

Repeating the steps above, we now get the Maxwell equation (5.28),

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\mu$$

The current J^μ in (5.43) couples directly to the gauge potential A_μ. This introduces a level of jeopardy, because the action should be invariant under gauge transformations $A_\mu \to A_\mu + \partial_\mu \chi$. Under such a gauge transformation, the action shifts as

$$S \to S + \frac{1}{c}\int d^4x\ (\partial_\mu \chi)J^\mu = S - \frac{1}{c}\int d^4x\ \chi(\partial_\mu J^\mu)$$

We see that the action is invariant only if the current is conserved, meaning $\partial_\mu J^\mu = 0$. But this, of course, is the expected property of the electric current. We see that the action principle introduces a nice interplay between gauge invariance and current conservation.

We can combine our Maxwell action (5.42) with the action for a relativistic point particle (5.41). We then have

$$S[A_\mu, X^\mu] = -\frac{1}{4\mu_0 c}\int d^4x\ F_{\mu\nu}F^{\mu\nu} + \int d\sigma\ \left[-mc\sqrt{\frac{dX^\mu}{d\sigma}\frac{dX_\mu}{d\sigma}} - qA_\mu(X)\frac{dX^\mu}{d\sigma}\right]$$

Comparing the last term to that in (5.43), we see that the current from a relativistic particle takes the form

$$J^\mu = qc\int d\sigma\ \frac{dX^\mu}{d\sigma}\,\delta^4(x - X(\sigma))$$

The Theta Term

As we saw previously, there is one other Lorentz invariant term that we can construct from the electric and magnetic fields. This is (5.25)

$$\frac{1}{4}\tilde{F}^{\mu\nu}F_{\mu\nu} = -\frac{1}{c}\mathbf{E}\cdot\mathbf{B}$$

We might wonder what would happen if we were to add this term to the Maxwell action (5.42). To answer this, we need to think about what the term $\tilde{F}_{\mu\nu}F^{\mu\nu}$ looks like when written in terms of the gauge potential A_μ. We have

$$\tilde{F}^{\mu\nu}F_{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}F_{\mu\nu} = \epsilon^{\mu\nu\rho\sigma}(\partial_\rho A_\sigma)F_{\mu\nu} = \epsilon^{\mu\nu\rho\sigma}\partial_\rho(A_\sigma F_{\mu\nu})$$

where the last equality holds because the derivatives in $F_{\mu\nu}$ are anti-symmetrised with ∂_ρ. The upshot is that this term is a total derivative and total derivatives don't affect the equations of motion. So adding such a term doesn't do anything.

In fact, that last statement is only partially true. Adding total derivatives to the action doesn't change the classical equation of motion. But it can change the quantum theory in subtle and interesting ways. That's also true here, where the term $\tilde{F}^{\mu\nu}F_{\mu\nu}$ is known as the *theta term*. (Named, unhelpfully, after the coefficient that sits in front of it which is usually called θ.) The theta term has an interesting role to play in, among other places, the story of topological insulators. You can read more about this in the lectures on Gauge Theory.

5.6 More on Energy and Momentum

The electric and magnetic fields carry both energy and momentum. The purpose of this section is to further explore their properties.

5.6.1 Energy and Momentum Conservation

The energy density stored in the electric and magnetic fields is (4.3),

$$\mathcal{E} = \frac{\epsilon_0}{2}\mathbf{E}^2 + \frac{1}{2\mu_0}\mathbf{B}^2 \tag{5.44}$$

The importance of energy lies in the fact that it's conserved. Because we're dealing with an energy density, it must be conserved locally which means that there must be an underlying continuity equation. This is the essence of Poynting's theorem that we derived in Section 4.4. This follows by taking the time derivative and using the Maxwell equations

$$\begin{aligned}\frac{\partial \mathcal{E}}{\partial t} &= \epsilon_0 \mathbf{E}\cdot\frac{\partial \mathbf{E}}{\partial t} + \frac{1}{\mu_0}\mathbf{B}\cdot\frac{\partial \mathbf{B}}{\partial t} \\ &= \frac{1}{\mu_0}\mathbf{E}\cdot\nabla\times\mathbf{B} - \mathbf{E}\cdot\mathbf{J} - \frac{1}{\mu_0}\mathbf{B}\cdot\nabla\times\mathbf{E}\end{aligned}$$

which we can write as

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla\cdot\mathbf{S} = -\mathbf{E}\cdot\mathbf{J} \quad \text{with} \quad \mathbf{S} = \mathbf{E}\times\mathbf{B} \tag{5.45}$$

Here $\mathbf{S}$ is the *Poynting vector* that we introduced previously in Section 4.4. It has the interpretation of the energy current. In the absence of any external electric current, so $\mathbf{J} = 0$, (5.45) tells us that energy in the electromagnetic field is conserved. However, if there are electric currents $\mathbf{J} \neq 0$ around, then the electric field does work on them, extracting energy from the field. That's the meaning of the right-hand side of (5.45).

The derivation above shows that the Poynting vector $\mathbf{S}$ can be viewed as the flow of energy carried by the electromagnetic field. But it also has a second, closely related interpretation: it is the *momentum* in the electromagnetic field. More precisely, the electromagnetic momentum density is

$$\boldsymbol{\mathcal{P}} = \frac{1}{c^2}\mathbf{S} = \epsilon_0\mathbf{E}\times\mathbf{B} \tag{5.46}$$

Momentum is also conserved, and that means that there must be a second continuity equation involving the time derivative of $\boldsymbol{\mathcal{P}}$ .And there is. We have

$$\begin{aligned}\frac{\partial \boldsymbol{\mathcal{P}}}{\partial t} &= \epsilon_0\left(\frac{\partial \mathbf{E}}{\partial t}\times\mathbf{B} + \mathbf{E}\times\frac{\partial \mathbf{B}}{\partial t}\right) \\ &= \frac{1}{\mu_0}(\nabla\times\mathbf{B})\times\mathbf{B} - \mathbf{J}\times\mathbf{B} - \epsilon_0\mathbf{E}\times(\nabla\times\mathbf{E})\end{aligned}$$

We use the vector identity

$$(\nabla \times \mathbf{B}) \times \mathbf{B} = (\mathbf{B} \cdot \nabla)\mathbf{B} - \frac{1}{2}\nabla \mathbf{B}^2$$

with a similar expression for $\mathbf{E}$. At this point, it's helpful to revert to index notation. We have

$$\begin{aligned}\frac{\partial \mathcal{P}_i}{\partial t} &= \frac{1}{\mu_0}\left(B_j\partial_j B_i - \frac{1}{2}\partial_i \mathbf{B}^2\right) + \epsilon_0\left(E_j\partial_j E_i - \frac{1}{2}\partial_i \mathbf{E}^2\right) - \epsilon_{ijk}J_j B_k \\ &= \partial_j\left[\frac{1}{\mu_0}\left(B_iB_j - \frac{1}{2}\delta_{ij}\mathbf{B}^2\right) + \epsilon_0\left(E_iE_j - \frac{1}{2}\delta_{ij}\mathbf{E}^2\right)\right] \\ &\qquad + \frac{1}{\mu_0}B_i\partial_j B_j - \epsilon_0 E_i \partial_j E_j - \epsilon_{ijk}J_jB_k\end{aligned}$$

The first term in square brackets is a total derivative. That's just what we want for a continuity equation. Meanwhile, we replace the $\nabla \cdot \mathbf{B}$ and $\nabla \cdot \mathbf{E}$ terms on the final line by the appropriate Maxwell equation. The end result is three continuity equations, one for the momentum in each different direction

$$\frac{\partial \mathcal{P}_i}{\partial t} + \partial_j \sigma_{ij} = -(\rho \mathbf{E} - \mathbf{J} \times \mathbf{B})_i \tag{5.47}$$

where σ_{ij} is the collection of terms in the previous square bracket

$$\sigma_{ij} = \epsilon_0\left(\frac{1}{2}\delta_{ij}\mathbf{E}^2 - E_iE_j\right) + \frac{1}{\mu_0}\left(\frac{1}{2}\delta_{ij}\mathbf{B}^2 - B_iB_j\right) \tag{5.48}$$

This is known as the *Maxwell stress tensor*. Note that it is symmetric. We'll come back to this shortly. We also met a stress-tensor σ_{ij} in our lectures on Fluid Mechanics: they are conceptually the same object.

In the absence of any charges or currents, the right-hand side of (5.47) vanishes and we learn that the vector $\boldsymbol{\mathcal{P}}$ is conserved. But we recognise the right-hand of (5.47) as the force density on charges and currents. If the currents are mobile electrons, then this force will increase their momentum and so we expect a corresponding decrease of the momentum in the electromagnetic field. That's indeed what we see.

As we've seen, the momentum density $\boldsymbol{\mathcal{P}}$ and the energy flux $\mathbf{S}$ are proportional: $\boldsymbol{\mathcal{P}} = \mathbf{S}/c^2$. There are two ways to see why the factor of c^2 is needed. The first is that it ensures that the right-hand side of (5.47) is the force experienced by charges and currents, so that (5.47) can be viewed as a field theoretic generalisation of "$F = ma$". The second is to invoke some quantum mechanical intuition, where the energy and momentum of photons are related by $p = E/c$. That accounts for one factor of c. The other arises because the energy flux is Ec, so the momentum is $p = (Ec)/c^2$.

5.6.2 The Energy-Momentum Tensor

There is an interesting interplay between field theory and relativity. This is illustrated by the fact that the energy density can actually be viewed as the zeroth component of two different four vectors!

The first of these four vectors follows because we're dealing with a field theory. This means that energy (or, more precisely, energy density) is conserved locally and sits in a current $J^\mu = (\mathcal{E}, \mathbf{S}/c)$, which obeys $\partial_\mu J^\mu = 0$ as seen in (5.45). (Recall that $\partial_0 = \frac{1}{c}\partial_t$ which is why there's that extra factor of c in the energy flux.)

But, in relativistic particle mechanics, the energy sits in a four vector with momentum (5.6). This suggests that we can also form the four vector $(\mathcal{E}/c, \mathcal{P})$. What's going on?

In fact, the energy density naturally sits not in a vector, but in a rank 2 tensor. This is known as the *stress-energy tensor*, or sometimes as the *energy-momentum tensor*, and sometimes, rather lazily and confusingly, just as the stress-tensor. It takes the form

$$T^{\mu\nu} = \begin{pmatrix} \mathcal{E} & c\mathcal{P}_i \\ S_i/c & \sigma_{ij} \end{pmatrix}$$

This includes both four-vectors above, one as a row vector and the other as a column vector. Because of the relation (5.46) between the energy flux $\mathbf{S}$ and the momentum density $\boldsymbol{\mathcal{P}}$, the energy-momentum tensor is actually symmetric

$$T^{\mu\nu} = T^{\nu\mu}$$

We'll return to this observation shortly.

Above, we argued that $T_{\mu\nu}$ should be a tensor on the grounds that the energy density $\mathcal{E}$ can be viewed as the zeroth component of two different four-vectors. Putting in the various definitions for the energy density $\mathcal{E}$ (5.44), the Poynting vector (5.45), the momentum $\boldsymbol{\mathcal{P}}$ (5.46), and the stress tensor σ_{ij} (5.48), you can check that the energy-momentum tensor can be constructed from the electromagnetic tensor $F^{\mu\nu}$ defined in (5.15). We have

$$T^{\mu\nu} = -\frac{1}{\mu_0}\left(F^{\mu\rho}F^{\nu}{}_{\rho} - \frac{1}{4}\eta^{\mu\nu}F^{\rho\sigma}F_{\rho\sigma}\right) \tag{5.49}$$

For example, the T^{00} component is

$$T^{00} = -\frac{1}{\mu_0}\left(F^{0\rho}F^{0}{}_{\rho} - \frac{1}{4}F^{\rho\sigma}F_{\rho\sigma}\right) = -\frac{1}{\mu_0}\left(-\mathbf{E}^2/c^2 - \frac{1}{2}(-\mathbf{E}^2/c^2 + \mathbf{B}^2)\right) = \mathcal{E}$$

This shows that $T^{\mu\nu}$ is indeed a tensor as advertised, meaning that it has the appropriate transformation law. Under a Lorentz transformation Λ, we have

$$T^{\mu\nu} \to \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}T^{\rho\sigma} \tag{5.50}$$

The column vectors are constructed out of conserved currents for energy and momentum respectively. This means that, in vacuum, the energy-momentum tensor obeys

$$\partial_\mu T^{\mu\nu} = 0$$

for each $\nu = 0, 1, 2, 3$. This captures the conservation of energy and momentum. Of course, because $T^{\mu\nu} = T^{\nu\mu}$, we also have $\partial_\nu T^{\mu\nu} = 0$.

If we turn on background electric charges ρ and electric currents $\mathbf{J}$ then, as we have seen, $T^{\mu\nu}$ is not conserved as the electromagnetic fields do work. From (5.49), we have

$$\begin{aligned} \mu_0 \partial_\mu T^{\mu}{}_{\nu} &= -(\partial_\mu F^{\mu\rho})F_{\nu\rho} - F^{\mu\rho}\partial_\mu F_{\nu\rho} + \frac{1}{2}F^{\rho\sigma}\partial_\nu F_{\rho\sigma} \\ &= -\mu_0 J^\rho F^{\nu}{}_{\rho} - \frac{1}{2}F^{\rho\sigma}\left(\partial_\rho F_{\nu\sigma} - \partial_\sigma F_{\nu\rho} - \partial_\nu F_{\rho\sigma}\right) \end{aligned}$$

To get to the last line, we've used Maxwell's equations in the form $\partial_\mu F^{\mu\nu} = \mu_0 J^\nu$ and engaged in some relabelling of dummy indices. (The $F^{\mu\rho}\partial_\mu F_{\nu\rho}$ term in the first line is split into two, with the dummy indices relabelled differently in the two cases.) But the final term in brackets vanishes, a fact that is equivalent to the other set of Maxwell equations $\partial_\mu \tilde{F}^{\mu\nu} = 0$ as we previously noted in (5.27). The upshot is that, in the presence of charges and currents, we have

$$\partial_\mu T^{\mu\nu} = -F^{\nu\rho}J_\rho$$

This combines our previous equations (5.45) and (5.47) into tensor form.

All relativistic field theories have an energy-momentum tensor. This plays a special role in a number of contexts, not least in General Relativity where $T^{\mu\nu}$ sits on the right-hand side of the Einstein equations and sources the gravitational field, in much the same way as J^μ sources the electromagnetic field in these lectures.

The energy-momentum tensor (5.49) has one further property that is special to Maxwell theory: it is traceless

$$T^{\mu}{}_{\mu} = 0$$

This follows from (5.49) because $\eta^{\mu\nu}\eta_{\mu\nu} = 4$. Although we won't show it here, it turns out that the energy-momentum tensor is traceless because of a special symmetry of Maxwell theory known as *conformal symmetry*.

If we have a gas of photons that are homogeneous, then the energy-momentum tensor necessarily takes the form

$$T^{\mu\nu} = \mathrm{diag}(\mathcal{E}, P, P, P)$$

Here the diagonal entries P of the stress-tensor have the interpretation of pressure. The tracelessness of the energy-momentum tensor tells us that the energy density and pressure are related by $\mathcal{E} = P/3$. This fact plays an important role in Cosmology.

5.6.3 Angular Momentum

In classical mechanics, there are three important conserved quantities: energy, momentum and angular momentum. But for our electromagnetic fields, we have described only the first two. We now rectify this.

In fact, we'll see that it will be fruitful to think more generally about *any* field theory that has a conserved energy density $\mathcal{E}$ and conserved momentum density $\boldsymbol{\mathcal{P}}$, such that the following two continuity equations hold:

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \mathbf{S} = 0 \quad \text{and} \quad \frac{\partial \mathcal{P}_i}{\partial t} + \partial_j \sigma_{ij} = 0$$

We know that in Maxwell theory the energy flux $\mathbf{S}$ is proportional to the momentum $\boldsymbol{\mathcal{P}}$. This, ultimately, was responsible for the symmetry $T^{\mu\nu} = T^{\nu\mu}$. In what follows, we won't assume any relation between $\mathbf{S}$ and $\boldsymbol{\mathcal{P}}$. Instead, we will see that this is a requirement of the conservation of angular momentum, together with Lorentz symmetry.

Following our nose from classical mechanics, we expect that the angular momentum density of the field is

$$\mathbf{L}(\mathbf{x}) = \mathbf{x} \times \boldsymbol{\mathcal{P}}(\mathbf{x})$$

We can ask if this is conserved. Differentiating by time only acts on $\boldsymbol{\mathcal{P}}$, not on the vector $\mathbf{x}$ above which simply tells us the point in space that we're looking at. We then have

$$\frac{\partial L_i}{\partial t} = \epsilon_{ijk} x_j \frac{\partial \mathcal{P}_k}{\partial t} = -\epsilon_{ijk} x_j \partial_l \sigma_{kl} = -\epsilon_{ijk} \partial_l \left(x_j \sigma_{kl} \right) + \epsilon_{ijk} \sigma_{kj}$$

We see that we get a continuity equation for angular momentum

$$\frac{\partial L_i}{\partial t} + \partial_l \left(\epsilon_{ijk} x_j \sigma_{kl} \right) = 0 \tag{5.51}$$

only if the stress tensor is symmetric: $\sigma_{ij} = \sigma_{ji}$.

The stress tensor σ_{ij} also plays a role in Fluid Mechanics. In that context, we gave a slightly awkward argument that σ_{ij} should be symmetric by showing that something bad would happen if it wasn't. That something bad was that a finite torque would give rise an infinite angular velocity. That's closely related to the much simpler derivation above that shows we have conservation of angular momentum if and only if σ_{ij} is symmetric.

The discussion above holds for any field theory with rotational invariance. However, if we have a Lorentz invariant theory like electromagnetism then it tells us that the energy-momentum tensor must also be symmetric. This follows from the Lorentz transformation law (5.50). If T^{ij} is symmetric in one frame then it is symmetric in all frames if and only $T^{i0} = T^{0i}$. This relates the energy flux and momentum as in (5.46).

Just as the energy density and momentum density sit nicely in a Lorentz invariant tensor $T^{\mu\nu}$, so too does the angular momentum density. However, this time it's a 3-tensor,

$$S^{\mu\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho}$$

By construction, $S^{\mu\rho\sigma} = -S^{\mu\sigma\rho}$. This tensor is conserved provided the $T^{\mu\nu} = T^{\nu\mu}$, a fact that follows from the same kind of calculation we did for the angular momentum, but now in Lorentz covariant form

$$\partial_\mu S^{\mu\rho\sigma} = T^{\rho\sigma} + x^\rho \partial_\mu T^{\mu\sigma} - T^{\sigma\rho} - x^\sigma \partial_\mu T^{\mu\rho} = 0$$

where we've used both the symmetry of $T^{\mu\nu}$ and the fact that it's conserved, so $\partial_\mu T^{\mu\nu} = 0$.

The components of $S^{\mu\rho\sigma}$ include the angular momentum, which can be found lurking in $S^{0ij} = c\epsilon^{ijk} L^k$. The equation $\partial_\mu S^{\mu ij} = 0$ is then just conservation of angular momentum of the field that we saw previously in (5.51). But that means there are also three more conserved quantities in this tensor, namely $\partial_\mu S^{\mu 0i} = 0$ for $i = 1, 2, 3$. What are these?! It's simple to find the answer by expanding out

$$S^{00i} = -(x^i \mathcal{E} - S^i t)$$

In fact, this has a rather straightforward meaning when $S^i = 0$: it is just the "centre of mass", or more precisely the centre of energy, of the field configuration. When $S^i \neq 0$, there is an additional drift term. The fact that this is conserved is rather like a field-theoretic version of Newton's first law which says that, in the absence of any force, a particle will continue at a constant speed. We see that after all these relativistic gymnastics, we come to something familiar, albeit in unfamiliar language.