8.07 Class Notes Fall 2011

Comments and questions to jbelcher@mit.edu

Table of Contents
(*LIVE LINKS*)

13 Special Relativity

13.1 Learning Objectives

We discuss the Principle of Relativity and the conundrum facing the late 1800's physicist.

13.2 Co-moving frames

13.2.1 Setting up the unbarred coordinate system

We first discuss how space-time events are measured in *different, co-moving coordinate frames*, and how to relate the measurements made in one frame to the measurements made in another, co-moving frame. Let first describe how we set up a set of observers in a given coordinate system, and how we record events in that system.

To construct a set of observers for a given coordinate system, we do the following. We get together a large number of people at some place ("the origin") far back in the past. They all are given identical clocks and rulers, and we make sure that the clocks are synchronized and that the rulers are of the same length by direct comparison. They agree upon a set of Cartesian coordinate directions in space, and some central authority assigns a position in space for each observer, using distance along these coordinate directions to specify positions. Each observer is also given a lab book. They then start out from the origin and take up their assigned positions in space, using their standard rulers to measure the distance as they go. They do this arbitrarily slowly, so that their clocks have an arbitrarily small difference due to time dilation (we consider this in a bit). At the end of this process, they are all at rest with respect to one another, observing what is happening right where they are, and nowhere else.

Once they are in position, we are can now record how an event or sequence of events "actually" happens. Assume that we have enough observers that they are densely spread in space. Then something happens, for example a particle moves through space. Each observer only records in her lab book the events which happen right at her feet. Thus one observer might say, "I saw the particle at my feet at my time t according to my clock, and I am located at position $\mathbf{r}$ ". Because the event happens right at that observer's feet, there is no worry about the time it takes the information to propagate from the event itself to that observer (they are all infinitesimally small). So we have an true measure of when the event occurs. All the observers involved faithfully record everything that happens right at their feet, and nowhere else, for the duration of the event or sequence of events.

Then, after the event is over, all the observers return to the origin, and all the lab books are collected. The event is then analyzed, by reconstructing what happened at every point in space, at every time, by someone who was right there when it happened. When the reconstruction is finished, we have a description of the event or sequence of events as they actually happened, the same description an omnipotent all-seeing all-knowing deity would give.

For example, we have a record of the position of a point particle as a function of coordinate time t, say $\mathbf{X}(t)$. From that record of position, we can calculate the particle's velocity at time t, by simply calculating the vector displacement $\Delta\mathbf{r}\,\Delta\mathbf{r} = \mathbf{X}(t+\Delta t) - \mathbf{X}(t)$ in spatial coordinates from t to $t + \Delta t$, and setting $\mathbf{v}(\text{t})=\Delta\mathbf{r}/\Delta\text{t}$. And so on. This is the way that we measure what "actually" happened in this system--we do not rely on having to stand at one point and "watch" what happens to something far from us. We use an infinite number of observers who record only what happens right where they are, and thus do not have to rely on information propagating to them at finite speeds.

Now, we need only add one requirement--we want this system to be an "inertial" coordinate system. It will be an inertial system if a particle in motion, left to itself, remains in motion at the same velocity. That is, if no forces act on the particle, we observe that the particle moves at constant velocity in this system. Then our system is inertial.

13.2.2 Setting up a co-moving inertial frame--the barred coordinate system

Now, we set up another inertial system, with another infinite set of observers, using exactly the same procedure as above, and using the same brand of rulers and clocks. However, in this new system (the barred system), the set of observers, when they are in place and ready to record a set of space-time events, are all observed (in the manner described above) by our observers in the first (unbarred) system to move at a constant velocity $\mathbf{v}$. For convenience, we assume that

$$\mathbf{v} = \text{v}\,\hat{\mathbf{x}} \ . \qquad (13.2.1)$$

We now observe an event in space-time, or a sequence of events, in both coordinate systems. After the observations are made, the observers in both frames all return to some agreed upon point in space, and all the lab books are collected, for observers in both frames. The event is then analyzed, in both frames, by reconstructing what happened at every point in space, at every time, in both frames, by someone who was right there when it happened. When the reconstruction is finished, we have a description of the event or sequence of events as they happened, in both frames. Now the question is the following. How are the coordinates recorded for an event in space-time in the unbarred frame, (x, y, z, t) related to the coordinates recorded for *that same* space-time event in the barred frame, $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$?

13.2.3 Selecting a common space-time origin for the co-moving frames

To find that relationship, let us first redefine the origins in space-time for the barred and unbarred frames, so that they coincide. To do that, let us pick a particular event in space-time, say the birth of a future MIT student at Massachusetts General

Hospital. Suppose the observer in the unbarred frame who was right there at the moment of birth is located at (x_b, y_b, z_b, t_b) in space-time. We use these coordinates to redefine this event as the origin of coordinates for the unbarred system--that is, we go back to our unbarred notebooks and we subtract from all our space-time observations at (x, y, z, t) the coordinates of this particular space-time event, that is, we re-compute locations in space and time as $(x - x_b, y - y_b, z - z_b, t - t_b)$.

Similarly, suppose the observer in the barred frame who was right there at the moment of birth of our potential student is located at $(\bar{x}_b, \bar{y}_b, \bar{z}_b, \bar{t}_b)$ .. We use these coordinates to redefine this event as the origin of coordinates for the barred system--that is, we go back to our barred notebooks and we subtract from all our space-time observations at $(\bar{x}_b, \bar{y}_b, \bar{z}_b, \bar{t}_b)$ the coordinates of this particular space-time event, that is, we re-compute locations in space and time as $(\bar{x} - \bar{x}_b, \bar{y} - \bar{y}_b, \bar{z} - \bar{z}_b, \bar{t} - \bar{t}_b)$. Our co-moving coordinate systems after this process now have the same point in space-time as their common origin, which is the birth of our future MIT student. We will always assume that this process has been carried out, so that the spatial origins of our two co-moving coordinate systems coincide at $t = \bar{t} = 0$.

13.2.4 The Galilean transformation

Now, what is the prescription or mapping that takes us from the coordinates of a given space-time event in the unbarred frame to the coordinates of that same space-time event in the barred frame? The Galilean mapping, or transformation (which turns out to be incorrect), when the barred frame is moving with velocity $\mathbf{v} = \mathrm{v}\,\hat{\mathbf{x}}$ with respect to the unbarred frame, is as follows:

$$\begin{aligned} \bar{t} &= t \\ \bar{x} &= x - \mathrm{v}t \\ \bar{y} &= y \\ \bar{z} &= z \end{aligned} \tag{13.2.2}$$

Note that we have built into this transformation the condition that the spatial origins coincide at $t = \bar{t} = 0$.

Let us be really clear about what this mapping means, by selecting a specific example. Suppose that the velocity of the observers in the barred frame as measured in the unbarred frame is $\mathbf{v} = 1\text{ meter/year }\hat{\mathbf{x}}$, and that the unbarred frame is at rest with respect to MGH. Suppose that at the age of 10, our future student has a minor accident and is wheeled into the same room at MGH in which he was born exactly 10 years earlier. The coordinates of this "return" event in the unbarred system, which we assume is at rest with respect to MGH, are (0,0,0,10 years). The coordinates of this event in the barred system are, according to out mapping rules given in (13.2.2), (-v x 10 years, 0,0,10 years) = (-10 meters, 0,0,10 years). What does this mean? It means that the observer in the barred frame who observes the student being wheeled into the same room, observes this at a time $\bar{t} = 10$ *years* at a position located $\bar{x} = -10$ meters down the $\bar{x}$ axis. Note that the *origin* of the barred system at this time is located at this time a distance +10 meters up the x axis. Just as we expect.

13.3 Gravitational interactions invariant under Galilean transformation

The Principle of Relativity has been around for a long time, long before Einstein, and was first set out in the context of mechanics. Quite simply, the Principle of Relativity says that there is no physical measurement we can make that can determine the absolute speed of the coordinate system in which we are making the measurement. An equivalent statement is that the form of physical laws must be the same in all co-moving frames. Newton's most striking success, the equations describing planetary motion, are a good example of the Principle of Relativity in mechanics. We review that example, so as to get some idea of the context in which the equations of electromagnetism emerged many years later.

Consider two particles interacting gravitationally, as seen in two different co-moving frames, where the prescription for going from the coordinates of one event in space-time to that same event as seen in the co-moving frame is given by (13.2.2). We want to show in the context of Newtonian mechanics and the Galilean transformation, that we cannot make any measurements of the interaction of these particles that will determine the relative velocity of the two co-moving frames.

Suppose the trajectory of particle 1 with mass m_1 is $\mathbf{X}_1(t)$, and the trajectory of particle 2 with mass m_2 is $\mathbf{X}_2(t)$. We take as given that in the unbarred frame, the equations of motion describing the gravitational interaction of these two particles are

$$m_1 \frac{d^2}{dt^2}\mathbf{X}_1(t) = -\frac{G\, m_1 m_2}{\left|\mathbf{X}_1(t) - \mathbf{X}_2(t)\right|^2} \frac{\mathbf{X}_1(t) - \mathbf{X}_2(t)}{\left|\mathbf{X}_1(t) - \mathbf{X}_2(t)\right|} \tag{13.3.1}$$

$$m_2 \frac{d^2}{dt^2}\mathbf{X}_2(t) = -\frac{G\, m_1 m_2}{\left|\mathbf{X}_2(t) - \mathbf{X}_1(t)\right|^2} \frac{\mathbf{X}_2(t) - \mathbf{X}_1(t)}{\left|\mathbf{X}_2(t) - \mathbf{X}_1(t)\right|} \tag{13.3.2}$$

Now we ask the following questions. Given the equations above and the Galilean transformation, can we determine the equations of motion for the particles as they would appear in the barred frame? The answer to this equations is yes, as we demonstrate.

First of all, what is the trajectory of the two particles as seen in the barred frame? Pick a given time *t* in the unbarred frame. At that time, particle 1 is at $\mathbf{X}_1(t)$. The particle being at position $\mathbf{X}_1(t)$ at time *t* is an event in space-time. What are the coordinates of that space-time event in the barred frame? Using (13.2.2), the coordinates are $\bar{t} = t$ and $\bar{\mathbf{X}}_1(t) = \mathbf{X}_1(t) - \mathbf{v}\,t$. That is, we have that the trajectory of particle 1 as seen in the barred frame is given by

$$\bar{\mathbf{X}}_1(\bar{t}) = \mathbf{X}_1(t) - \mathbf{v}\,t \tag{13.3.3}$$

or

$$\mathbf{X}_1(t) = \bar{\mathbf{X}}_1(\bar{t}) + \mathbf{v}\,t \tag{13.3.4}$$

Using these equations, and similar equations for particle 2, we can easily see that

$$\begin{aligned} \mathbf{X}_1(t) - \mathbf{X}_2(t) &= \bar{\mathbf{X}}_1(\bar{t}) - \bar{\mathbf{X}}_2(\bar{t}) \\ \frac{d}{d\,t}\mathbf{X}_1(t) &= \frac{d}{d\,\bar{t}}\bar{\mathbf{X}}_1(\bar{t}) + \mathbf{v} \\ \frac{d^2}{d\,t^2}\mathbf{X}_1(t) &= \frac{d^2}{d\,\bar{t}^2}\bar{\mathbf{X}}_1(\bar{t}) \end{aligned} \tag{13.3.5}$$

In equations (13.3.5), we have all that we need to find the equation of motion for particle 1 in the barred frame, assuming that (13.3.1) is true. That is, we easily have that if (13.3.1) is true and if the Galilean transformations (13.2.2) hold, then in the barred frame,

$$m_1 \frac{d^2}{d\,\bar{t}^2}\bar{\mathbf{X}}_1(\bar{t}) = -\frac{G\,m_1\,m_2}{\left|\bar{\mathbf{X}}_1(\bar{t}) - \bar{\mathbf{X}}_2(\bar{t})\right|^2}\,\frac{\bar{\mathbf{X}}_1(\bar{t}) - \bar{\mathbf{X}}_2(\bar{t})}{\left|\bar{\mathbf{X}}_1(\bar{t}) - \bar{\mathbf{X}}_2(\bar{t})\right|} \tag{13.3.6}$$

with a similar equation for particle 2. Thus the equations in the barred system have exactly the same form as the equations in the unbarred system. This means that we cannot do any experiment in the barred system that will tell us the relative velocity between the two systems. Let's be really precise about what we mean by this statement.

13.4 What does it mean for mathematical equations to have the same form in co-moving Systems?

Here is one way to state the Principle of Relativity. You are put inside a closed metal box that is at rest in the barred frame and therefore moving at constant velocity **v** in the unbarred frame. You cannot look outside of the metal box, or interact with objects outside of the metal box. For example, you cannot look out of the box at some observer at rest in the unbarred frame through a window in the box. Or, you cannot stick your hand outside the box and touch something that is moving by you, at rest in the unbarred frame. However, you are allowed to have any measuring equipment you wish inside of the box. In fact, you are given the entire contents of Junior Lab, and you can do any physical experiment you want inside the box, to arbitrary precision. Then the Principle of Relativity states that there is no experiment you can do inside the box that will determine your velocity **v** with respect to the unbarred frame, or with respect to any other inertial frame, for that matter.

Mathematically, what this means is that the equations that describe the laws of physics must have the same form in different co-moving frames. In the above example of two co-moving frames, that means that the equations of motion in the two frames must contain no terms which refer to the relative velocity between the frames. If the equations of motion did contain such terms, then the motion of the gravitationally interacting particles in your box would reflect that difference, and that would be an observable difference between an experiment performed inside your box, in the barred frame, and the same experiment performed in the unbarred frame. But since the equations (13.3.1) and (13.3.6) describing the gravitational interaction in the two systems contain no such terms (they have the same *form*), any experiment you perform will yield results independent of the relative velocity. Therefore any such experiment will tell you nothing about your velocity with respect to the unbarred frame.

13.5 Sound waves under Galilean transformations

Instead of gravitational interaction, let us turn to an example where there ***is*** a preferred frame. Consider the equations describing the propagation of sound waves in air. For sound waves, there is in fact a preferred frame in which the equations assume a particularly simple form--the rest frame of the air. This is because we are talking about a fluid--a set of particles interacting frequently via collisions, which therefore have a common motion. Consider variations in time and the x direction only. In the rest frame of the air, the equations describing the velocity of a element of the air $\mathbf{w}(x,t) = \hat{\mathbf{x}}\, w(x,t)$ at (x,t) with mass density $\rho_{mass}(x,t)$ at (x,t) and gas pressure $p(x,t)$ are the momentum equation

$$\rho_{mass}(x,t)\frac{\partial}{\partial t}\mathbf{w}(x,t) = -\nabla p(x,t) = -\hat{\mathbf{x}}\frac{\partial}{\partial x}\, p(x,t) = -\hat{\mathbf{x}}\frac{\partial p}{\partial \rho}\frac{\partial \rho_{mass}(x,t)}{\partial x} \tag{13.5.1}$$

where we have assumed that there is a unique relation $p(\rho_{mass})$ between the mass density and the pressure, and the conservation of mass equation,

$$\frac{\partial}{\partial t}\rho_{mass} + \nabla\cdot\left(\rho_{mass}\mathbf{w}\right) = 0 \tag{13.5.2}$$

We define the "speed of sound" to be $s^2 = \dfrac{\partial p}{\partial \rho}$. Further more, we drop second order terms in (13.5.1) and (13.5.2). That is, if $\rho_{mass}(x,t) = \rho^{o}{}_{mass} + \delta\,\rho_{mass}(x,t)$, and if w is already considered first order small, then $\rho_{mass}(x,t)\dfrac{\partial}{\partial t}\mathbf{w}(x,t) \cong \rho^{o}{}_{mass}\dfrac{\partial}{\partial t}\mathbf{w}(x,t)$ to first order in small quantities. With these approximations, equations (13.5.1) and (13.5.2) become

$$\rho^{o}{}_{mass}\frac{\partial}{\partial t}w(x,t)=s^{2}\frac{\partial}{\partial x}\delta\,\rho_{mass}(x,t) \tag{13.5.3}$$

$$\frac{\partial}{\partial t}\delta\,\rho_{mass}(x,t)+\rho^{o}{}_{mass}\frac{\partial}{\partial x}w(x,t)=0 \tag{13.5.4}$$

and with a little manipulation, we can find an equation for w(x,t) that is

$$(\frac{\partial^2}{\partial x^2}-\frac{1}{s^2}\frac{\partial^2}{\partial t^2})\,w(x,t)=0 \tag{13.5.5}$$

This is just the wave equation in the rest frame of the air, which tells us that in this frame, sound waves propagate at the speed s.

Now, the obvious question is, what does equation (13.5.5) look like in a co-moving frame, assuming that the Galilean transformation (13.2.2) holds. In particular, how does the operator $(\frac{\partial^2}{\partial x^2}-\frac{1}{s^2}\frac{\partial^2}{\partial t^2})$ transforms under (13.2.2)? Well, suppose we have a scalar function $G(x,t)$. Using the chain rule for partial derivatives, with $G(x,t)=G(x(\bar{x},\bar{t}),t(\bar{x},\bar{t}))=\bar{G}(\bar{x},\bar{t})$, where we have put a bar on *G* because the functional form of $\bar{G}(\bar{x},\bar{t})$ is different from that of $G(x,t)$, we have

$$\frac{\partial G}{\partial t}=\frac{\partial\bar{G}}{\partial\bar{t}}\frac{\partial\bar{t}}{\partial t}+\frac{\partial\bar{G}}{\partial\bar{x}}\frac{\partial\bar{x}}{\partial t}=\frac{\partial\bar{G}}{\partial\bar{t}}-\mathrm{v}\frac{\partial\bar{G}}{\partial\bar{x}}=\left[\frac{\partial}{\partial\bar{t}}-\mathrm{v}\frac{\partial}{\partial\bar{x}}\right]\bar{G} \tag{13.5.6}$$

where we have used (13.2.2) to conclude that $\frac{\partial\bar{x}}{\partial t}=-\mathrm{v}$. Also, we have

$$\frac{\partial G}{\partial x}=\frac{\partial\bar{G}}{\partial\bar{t}}\frac{\partial\bar{t}}{\partial x}+\frac{\partial\bar{G}}{\partial\bar{x}}\frac{\partial\bar{x}}{\partial x}=\frac{\partial\bar{G}}{\partial\bar{x}} \tag{13.5.7}$$

Now, we similarly have that

$$\frac{\partial}{\partial t}\frac{\partial G}{\partial t}=\left[\frac{\partial}{\partial\bar{t}}-\mathrm{v}\frac{\partial}{\partial\bar{x}}\right]\left[\frac{\partial}{\partial\bar{t}}-\mathrm{v}\frac{\partial}{\partial\bar{x}}\right]\bar{G}=\left[\frac{\partial^2}{\partial\bar{t}^2}-2\mathrm{v}\frac{\partial^2}{\partial\bar{t}\,\partial\bar{x}}+\mathrm{v}^2\frac{\partial^2}{\partial\bar{x}^2}\right]\bar{G} \tag{13.5.8}$$

$$\frac{\partial^2 G}{\partial x^2}=\frac{\partial^2\bar{G}}{\partial\bar{x}^2} \tag{13.5.9}$$

so that the operator $(\frac{\partial^2}{\partial x^2}-\frac{1}{s^2}\frac{\partial^2}{\partial t^2})$ becomes in the barred coordinates

$$(\frac{\partial^2}{\partial x^2}{}^2-\frac{1}{s^2}\frac{\partial^2}{\partial t^2})G=\left[(1-\frac{\mathrm{v}^2}{s^2})\frac{\partial^2}{\partial \bar{x}^2}+2\frac{\mathrm{v}}{s^2}\frac{\partial^2}{\partial \bar{t}\,\partial \bar{x}}-\frac{1}{s^2}\frac{\partial^2}{\partial \bar{t}^2}\right]\bar{G} \qquad (13.5.10)$$

This equation certainly does *not* have the same form in the co-moving frame under the Galilean transformation. In fact it has two extra terms, which make a lot of sense. Suppose that we are looking for solutions to the wave equation in the barred frame, that is, a solution which makes (13.5.10) zero. Try a solution of the form

$$\bar{G}(\bar{x},\bar{t})=e^{i(\bar{k}\bar{x}-\bar{\varpi}\bar{t})} \qquad (13.5.11)$$

If this is to be a solution the "wave" equation in the barred frame, then $\bar{\varpi}$ and $\bar{k}$ must satisfy

$$(1-\frac{\mathrm{v}^2}{s^2})\bar{k}^2+2\frac{\mathrm{v}}{s^2}\bar{k}\bar{\varpi}-\frac{1}{s^2}\bar{\varpi}^2=0 \qquad (13.5.12)$$

With a little manipulation, this can be written as

$$\left(\bar{\varpi}-\mathrm{v}\bar{k}\right)^2-s^2\bar{k}^2=0 \qquad (13.5.13)$$

or

$$\frac{\bar{\varpi}}{\bar{k}}=\mathrm{v}\pm s \qquad (13.5.14)$$

Remember, the ratio $\frac{\bar{\varpi}}{\bar{k}}$ is the speed at which you see this pattern propagate in the barred frame in the $\bar{x}$ direction, and it is not s. This is exactly the behavior our everyday experience predicts. In frames other than the rest frame of the air, we see a sound wave move at different speeds than s, and the difference is just what you would expect, $\mathbf{v}=\mathrm{v}\hat{\mathbf{x}}$. So in this case, the equations describing the physical laws do have a particularly simple form in one frame--the rest frame of the fluid. If we measure the speed of sound in different directions in any other frame, we can indeed determine the velocity of that frame with respect to the frame in which the air is at rest.

But it is clear why that is a preferred frame, and it is clear that this does not violate the Principle of Relativity. What we are measuring if we do this is the relative velocity between the barred frame and the rest frame of the air. But this is like sticking our hand out of the moving metal box and feeling the air blowing past, or touching a table at rest in the unbarred frame and feeling the frictional force--we are not allowed to do that. If we are confined to the inside of the box, measuring the speed of sound inside the box will not give us the speed with which we are moving with respect to the unbarred frame, because our air inside the box is carried along with the box.

13.6 The dilemma of the late 1800's physicist

Let's just pose in a simple way the dilemma that many physicists faced in the late 1800's and early 1900's. It was well known of course that Maxwell's Equations yielded a wave equation with a propagation speed of the speed of light, c. The physics and mathematics of sound waves was also well known, so that everyone was aware that under Galilean transformations, the wave equation had the form (13.5.5) only in one frame--the rest frame of the medium in which the wave propagates. In other frames, the equations would be different--just as our wave equation in the barred frame (13.5.10) is different from (13.5.5). Which led most physicists to two conclusions. First, that Maxwell's Equations as we have written them down must only be correct in the rest frame of the medium in which light propagates (that medium was thought to be the ether). Therefore, if we want to write them down in other frames, they must have a form that is a different form from the form we have been studying.

Second, and more importantly, just as for a sound wave in air, it was thought that one could measure our speed with respect to the ether by just measuring the speed of light in different directions. That is, in your metal box with the junior lab equipment, you could measure the speed of the box in an absolute sense by measuring the speed of a light beam in different directions, without every looking out the window or interacting with the world outside the box.

Of course, the problem was that this experiment was done in the late 1800's by Michelson and Morley, and there was no discernible difference in the speed of light in different directions, even though the Earth moves around the Sun at 25 km/sec, and the Sun moves around the Galaxy at 200 km/sec, and so on. This experiment validated the Principle of Relativity, but no one could understand how this could be. There were lots of different ways to try to get around this conundrum (the mistaken belief that any wave motion must have a preferred frame) such as the "ether drag", and the Lorentz-Fitzgerald contraction, etc., but nothing that hung together until Einstein came along. Here were his choices

	Choice 1	**Choice 2**	**Choice 3**
Newton's Laws	OK	OK	Need to be modified
Galilean Transformations	OK	OK	Need to be modified
Principle of Relativity	OK for mechanics & E&M	OK for mechanics, not for E&M	OK for mechanics & E&M
Maxwell's Equations	Need to be modified in some way	OK	OK, same form in every inertial frame

He of course choose the last column, i.e., that Maxwell's Equations have the same form in every frame, and therefore satisfy the Principle of Relativity. But in choosing this alternative, not only do we need to find an alternative to the Galilean transformations, but we need to modify Newton's Laws to make sure that they preserve the Principle of Relativity under whatever transformation we decide is the right one.

13.7 The transformation of space and time

The Principle of Relativity says that the laws of physics should be the same in all inertial frames. In mathematical terms, this translates into the statement that the form of the equations of physics should be the same in all co-moving frames. We have shown above that the laws of gravitational interaction as set down by Newton are the same in co-moving frames if the Galilean transformation holds. In fact, this transformation is incorrect, although it is a good approximation for $V << c$.

What is the correct transformation? We deduce in this section the correct transformation laws for space and time by requiring that Maxwell's Equations have the same form in both our barred and unbarred frames. For the moment, we let this be a purely mathematical exercise, and ignore the physics. *Griffiths* takes exactly the opposite tack, concentrating first on the physics, and less on the actual structure of Maxwell's Equations. These approaches complement each other, and you should read and understand both.

First, before asking what the correct transformation is, it is clear that Maxwell's Equations do not remain the same in form under Galilean transformations, because the wave equation for light does not remain the same under Galilean transformations, as we demonstrated above. But we know that the Galilean transformations must be valid for speeds small compared to the speed of light, from experience. Therefore, we try to find a transformation that is close to the same form as the Galilean transformation. We try the form

$$\begin{aligned} \bar{t} &= a_{00}\, t + a_{01}\, x \\ \bar{x} &= a_{10}\, t + a_{11}\, x \\ \bar{y} &= y \\ \bar{z} &= z \end{aligned} \tag{13.7.1}$$

where the four unknown coefficients here can be functions of the relative velocity $\mathbf{v} = \mathrm{v}\,\hat{\mathbf{x}}$. There is in fact a relation between a_{10} and a_{11} which must hold, namely that

$$\frac{a_{10}}{a_{11}} = -\mathrm{v} \tag{13.7.2}$$

Why must this be true? Because the origin of the barred frame is moving at velocity $\mathbf{v} = \mathrm{v}\hat{\mathbf{x}}$ with respect to the unbarred frame. Since the origin or the barred frame is at $\bar{x} = a_{10}\, t + a_{11}\, x = 0$, those points (x,t) in the unbarred frame which map into the origin of the barred frame satisfy $x = -\dfrac{a_{10}}{a_{11}}t$, but they must also satisfy $x = \mathrm{v}t$, therefore (13.7.2) must hold.

Now, with this assumed form for the transformation, we can easily derive, in the same manner as (13.5.6) above, that

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \bar{t}}\frac{\partial \bar{t}}{\partial t} + \frac{\partial}{\partial \bar{x}}\frac{\partial \bar{x}}{\partial t} = a_{00}\frac{\partial}{\partial \bar{t}} + a_{10}\frac{\partial}{\partial \bar{x}} \qquad (13.7.3)$$

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial \bar{t}}\frac{\partial \bar{t}}{\partial x} + \frac{\partial}{\partial \bar{x}}\frac{\partial \bar{x}}{\partial x} = a_{01}\frac{\partial}{\partial \bar{t}} + a_{11}\frac{\partial}{\partial \bar{x}} \qquad (13.7.4)$$

$$\frac{\partial}{\partial y} = \frac{\partial}{\partial \bar{y}} \qquad\qquad \frac{\partial}{\partial z} = \frac{\partial}{\partial \bar{z}} \qquad (13.7.5)$$

which means that under this transformation

$$\frac{\partial^2}{\partial t^2} = \left[a_{00}\frac{\partial}{\partial \bar{t}} + a_{10}\frac{\partial}{\partial \bar{x}}\right]\left[a_{00}\frac{\partial}{\partial \bar{t}} + a_{10}\frac{\partial}{\partial \bar{x}}\right] = a_{00}{}^2\frac{\partial^2}{\partial \bar{t}^2} + 2a_{00}a_{10}\frac{\partial^2}{\partial \bar{x}\partial \bar{t}} + a_{10}{}^2\frac{\partial^2}{\partial \bar{t}^2}$$
(13.7.6)

$$\frac{\partial^2}{\partial x^2} = \left[a_{01}\frac{\partial}{\partial \bar{t}} + a_{11}\frac{\partial}{\partial \bar{x}}\right]\left[a_{01}\frac{\partial}{\partial \bar{t}} + a_{11}\frac{\partial}{\partial \bar{x}}\right] = a_{01}{}^2\frac{\partial^2}{\partial \bar{t}^2} + 2a_{01}a_{11}\frac{\partial^2}{\partial \bar{x}\partial \bar{t}} + a_{11}{}^2\frac{\partial^2}{\partial \bar{x}^2}$$
(13.7.7)

so that the wave equation becomes in the barred system is

$$(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}) =$$
$$\left[(a_{11}{}^2 - \frac{1}{c^2}a_{10}{}^2)\frac{\partial^2}{\partial \bar{x}^2} + \frac{\partial^2}{\partial \bar{y}^2} + \frac{\partial^2}{\partial \bar{z}^2} + 2(a_{01}a_{11} - \frac{1}{c^2}a_{10}a_{00})\frac{\partial^2}{\partial \bar{x}\partial \bar{t}} - \frac{1}{c^2}(a_{00}{}^2 - c^2a_{01}{}^2)\frac{\partial^2}{\partial \bar{t}^2}\right]$$
(13.7.8)

Now, if we want the form of this equation in the barred system to be unchanged, then we clearly must have

$$(a_{11}{}^2 - \frac{1}{c^2}a_{10}{}^2) = 1 \quad (a_{01}a_{11} - \frac{1}{c^2}a_{10}a_{00}) = 0 \quad (a_{00}{}^2 - c^2a_{01}{}^2) = 1 \qquad (13.7.9)$$

which together with (13.7.2) gives us

$$a_{11} = a_{00} = \frac{1}{\sqrt{1-\frac{\mathrm{v}^2}{c^2}}} = \frac{1}{\sqrt{1-\beta^2}} = \gamma \qquad a_{10} = -\gamma \mathrm{v} \qquad a_{01} = -\frac{\gamma \mathrm{v}}{c^2} \tag{13.7.10}$$

and our equations (13.7.1) become

$$\begin{aligned} \bar{t} &= \gamma\left(t - \frac{\mathrm{v}}{c^2}x\right) \\ \bar{x} &= \gamma\left(-\mathrm{v}t + x\right) \\ \bar{y} &= y \\ \bar{z} &= z \end{aligned} \tag{13.7.11}$$

These equations define the Lorentz transformations[3]. They reduce to the Galilean transformations for $\mathrm{v}/c \ll 1$. If we define the coordinate $x^0 \equiv ct$, then this transformation can be written in the matrix form

$$\begin{pmatrix} \bar{x}^0 \\ \bar{x}^1 \\ \bar{x}^2 \\ \bar{x}^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{13.7.12}$$

or

$$x'^{\mu} = \sum_{\nu=0}^{3} \Lambda^{\mu}{}_{\nu}\, x^{\nu} \tag{13.7.13}$$

---

[3] There is an excellent collection of papers in a book "The Principle of Relativity" by Lorentz, Einstein, Minkowski, and Weyl, Dover, 1952, including Lorentz's original paper where he derives this transformation, another paper by Lorentz discussing the Michelson-Morley experiment and his contraction hypothesis, and Einstein's original 1905 paper, among others.

14 Transformation of Sources and Fields

14.1 Learning Objectives

Having derived the way that space and time transform, we now derive the way that the potentials, fields and sources transform. Again, our only guide in this is the requirement that the form of Maxwell's equations be the same in co-moving frames.

14.2 How do ρ and $\mathbf{J}$ transform?

Now that we know how space and time transform, let us inquire about how the fields the sources ρ and **J** transform, again approaching this from the mathematical requirement that the form of Maxwell's Equations be the same in different inertial frames. With (13.7.12), (13.7.3) and (13.7.4) become

$$\frac{\partial}{\partial t} = \gamma \frac{\partial}{\partial \overline{t}} - \gamma \mathrm{v} \frac{\partial}{\partial \overline{x}} \tag{14.2.1}$$

$$\frac{\partial}{\partial x} = -\gamma \frac{\mathrm{v}}{c^2} \frac{\partial}{\partial \overline{t}} + \gamma \frac{\partial}{\partial \overline{x}} \tag{14.2.2}$$

With these relations, charge conservation (4.3.1) ($\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0$) is

$$\begin{aligned} &\frac{\partial}{\partial t}\rho + \frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z} = \\ &\left(\gamma \frac{\partial}{\partial \overline{t}} - \gamma \mathrm{v} \frac{\partial}{\partial \overline{x}} \right)\rho + \left(-\gamma \frac{\mathrm{v}}{c^2} \frac{\partial}{\partial \overline{t}} + \gamma \frac{\partial}{\partial \overline{x}} \right) J_x + \frac{\partial J_y}{\partial \overline{y}} + \frac{\partial J_z}{\partial \overline{z}} \\ &= \frac{\partial}{\partial \overline{t}} \gamma \left(\rho - \frac{\mathrm{v}}{c^2} J \right) + \frac{\partial}{\partial \overline{x}} \gamma \left(J_x - \mathrm{v}\rho \right) + \frac{\partial J_y}{\partial \overline{y}} + \frac{\partial J_z}{\partial \overline{z}} = 0 \end{aligned} \tag{14.2.3}$$

But if the conservation of charge is to hold in the barred frame, we expect that in that frame we will also have $\frac{\partial}{\partial \overline{t}} \overline{\rho} + \overline{\nabla} \cdot \overline{\mathbf{J}} = 0$. If we look at the last line of equation (14.2.3), this means that we must have

$$\overline{\rho} = \gamma \left(\rho - \frac{\mathrm{v}}{c^2} J_x \right) \qquad \overline{J}_x = \gamma \left(J_x - \mathrm{v}\rho \right) \qquad \overline{J}_y = J_y \qquad \overline{J}_z = J_z \tag{14.2.4}$$

14.3 How do E and B transform?

So we know how the sources transform. How about the fields? With (14.2.1) and (14.2.2), equation (4.2.1) ($\nabla \cdot \mathbf{E} = \rho / \varepsilon_o$) becomes

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \left(-\gamma \frac{\mathrm{v}}{c^2} \frac{\partial}{\partial \bar{t}} + \gamma \frac{\partial}{\partial \bar{x}} \right) E_x + \frac{\partial E_y}{\partial \bar{y}} + \frac{\partial E_z}{\partial \bar{z}} = \frac{\rho}{\varepsilon_o} \tag{14.3.1}$$

and the *x*-component of equation (4.2.3) becomes

$$\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z} = \mu_o J_x + \frac{1}{c^2} \frac{\partial E_x}{\partial t} = \mu_o J_x + \frac{1}{c^2} \left(\gamma \frac{\partial}{\partial \bar{t}} - \gamma \mathrm{v} \frac{\partial}{\partial \bar{x}} \right) E_x \tag{14.3.2}$$

If we solve equation (14.3.2) for $\gamma \frac{\partial E_x}{\partial \bar{t}}$,we obtain

$$\gamma \frac{\partial}{\partial \bar{t}} E_x = c^2 \left(\frac{\partial B_z}{\partial \bar{y}} - \frac{\partial B_y}{\partial \bar{z}} \right) - c^2 \mu_o J_x + \left(\gamma \mathrm{v} \frac{\partial E_x}{\partial \bar{x}} \right) \tag{14.3.3}$$

Inserting (14.3.3) into (14.3.1) gives

$$-\frac{\mathrm{v}}{c^2} \left[c^2 \left(\frac{\partial B_z}{\partial \bar{y}} - \frac{\partial B_y}{\partial \bar{z}} \right) - c^2 \mu_o J_x + \left(\gamma \mathrm{v} \frac{\partial E_x}{\partial \bar{x}} \right) \right] + \gamma \frac{\partial E_x}{\partial \bar{x}} + \frac{\partial E_y}{\partial \bar{y}} + \frac{\partial E_z}{\partial \bar{z}} = \frac{\rho}{\varepsilon_o} \tag{14.3.4}$$

which with a little rearrangement can be written as

$$\frac{1}{\gamma} \frac{\partial E_x}{\partial \bar{x}} + \frac{\partial}{\partial \bar{y}} \left(E_y - \mathrm{v} B_z \right) + \frac{\partial}{\partial \bar{z}} \left(E_z + \mathrm{v} B_y \right) = \frac{\rho}{\varepsilon_o} - \frac{\mathrm{v}}{c^2} \frac{J_x}{\varepsilon_o} \tag{14.3.5}$$

or

$$\frac{\partial E_x}{\partial \bar{x}} + \frac{\partial}{\partial \bar{y}} \left[\gamma \left(E_y - \mathrm{v} B_z \right) \right] + \frac{\partial}{\partial \bar{z}} \left[\gamma \left(E_z + \mathrm{v} B_y \right) \right] = \frac{\gamma}{\varepsilon_o} \left(\rho - \frac{\mathrm{v}}{c^2} J_x \right) \tag{14.3.6}$$

but we know that if Maxwell's Equations have the same form in the barred system, then we must have

$$\frac{\partial \bar{E}_x}{\partial \bar{x}} + \frac{\partial \bar{E}_y}{\partial \bar{y}} + \frac{\partial \bar{E}_z}{\partial \bar{z}} = \frac{\bar{\rho}}{\varepsilon_o} \tag{14.3.7}$$

which means that if (14.3.6) holds, the electric field components in the barred frame must be related to those in the unbarred frame by

$$\bar{E}_x = E_x \qquad \bar{E}_y = \gamma \left(E_y - \mathrm{v} B_z \right) \qquad \bar{E}_z = \gamma \left(E_z + \mathrm{v} B_y \right) \tag{14.3.8}$$

and that the charge density in the barred frame must be related to quantities in the unbarred frame via

$$\bar{\rho} = \gamma\left(\rho - \frac{\mathrm{v}}{c^2} J_x\right) \tag{14.3.9}$$

But (14.3.9) is nothing new, it is just the first equation in (14.2.4).

How does the magnetic field transform? Well, consider (4.2.4) ($\nabla \cdot \mathbf{B} = 0$). Using (14.2.2), we have

$$\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = \left(-\gamma \frac{\mathrm{v}}{c^2}\frac{\partial}{\partial \bar{t}} + \gamma \frac{\partial}{\partial \bar{x}}\right) B_x + \frac{\partial B_y}{\partial \bar{y}} + \frac{\partial B_z}{\partial \bar{z}} = 0 \tag{14.3.10}$$

and with the x-component of (4.2.2) ($\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$) and (14.2.2), we have

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -\frac{\partial B_x}{\partial t} = -\left(\gamma \frac{\partial}{\partial \bar{t}} - \gamma \mathrm{v} \frac{\partial}{\partial \bar{x}}\right) B_x \tag{14.3.11}$$

and if we solve (14.3.11) for $\gamma \frac{\partial B_x}{\partial \bar{t}}$,we obtain

$$\gamma \frac{\partial B_x}{\partial \bar{t}} = \gamma \mathrm{v} \frac{\partial B_x}{\partial \bar{x}} + \frac{\partial E_z}{\partial \bar{y}} - \frac{\partial E_y}{\partial \bar{z}} \tag{14.3.12}$$

If we insert (14.3.12) into (14.3.10), we obtain

$$-\frac{\mathrm{v}}{c^2}\left[\gamma \mathrm{v} \frac{\partial B_x}{\partial \bar{x}} + \frac{\partial E_z}{\partial \bar{y}} - \frac{\partial E_y}{\partial \bar{z}}\right] + \gamma \frac{\partial B_x}{\partial \bar{x}} + \frac{\partial B_y}{\partial \bar{y}} + \frac{\partial B_z}{\partial \bar{z}} = 0 \tag{14.3.13}$$

or

$$\frac{\partial B_x}{\partial \bar{x}} + \frac{\partial}{\partial \bar{y}}\left[\gamma\left(B_y + \frac{\mathrm{v}}{c^2} E_z\right)\right] + \frac{\partial}{\partial \bar{z}}\left[\gamma (B_z - \frac{\mathrm{v}}{c^2} E_y)\right] = 0 \tag{14.3.14}$$

As above, this means that since $\bar{\nabla} \cdot \bar{\mathbf{B}} = 0$ in the barred frame, we must have

$$\bar{B}_x = B_x \qquad \bar{B}_y = \gamma\left(B_y + \frac{\mathrm{v}}{c^2} E_z\right) \qquad \bar{B}_z = \gamma (B_z - \frac{\mathrm{v}}{c^2} E_y) \tag{14.3.15}$$

14.4 How do the potentials transform?

What about the potentials? Well, if we define the four-vector current by

$$J^{\mu} = (c\rho, J_x, J_y, J_z) \tag{14.4.1}$$

then (14.2.4) tells us that this four-vector transforms the same way that x^{μ} does. Moreover, if we define the four-vector potential by

$$A^{\mu} = (\frac{\phi}{c}, A_x, A_y, A_z) \tag{14.4.2}$$

then we know that

$$(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2})\, A^{\mu} = -\mu_o J^{\mu} \tag{14.4.3}$$

Therefore, since the differential operator in (14.4.3) does not change form, and J^{μ} transforms as x^{μ} , then A^{μ} must also transform the same way.

We have therefore derived the transformation properties of space and time, and of all the electromagnetic quantities that appear in Maxwell's Equations, simply by assuming that Maxwell's Equations must have the same form in co-moving systems. In particular, the way we have derived the transformation properties of the fields is that used by Einstein in his original paper.

15 Manifest Covariance

15.1 Learning Objectives

We look at how to write Maxwell’s Equations in “manifestly covariant” form. This means that at a glance we can tell that Maxwell’s Equations have the same form in all co-moving frames.

15.2 Contra-variant and covariant vectors

To write Maxwell's Equations in a "manifestly covariant" form simply means that we write them in a way such that at a glance they can be seen to be covariant--it is "manifest". By covariant, we mean that they have the same form in all inertial frames. We already know that they are covariant of course--we showed that in the previous section, but we want to demonstrate this in a more elegant way.

To do this, we need to define contra-variant and co-variant four vectors. A four vector is contra-variant if it transforms like x^{μ}, that is, if

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) \qquad (15.2.1)$$

then

$$\begin{pmatrix} \bar{x}^0 \\ \bar{x}^1 \\ \bar{x}^2 \\ \bar{x}^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \qquad (15.2.2)$$

or

$$\bar{x}^\mu = \sum_{\nu=0}^{3} \Lambda^\mu{}_\nu \, x^\nu \qquad \Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (15.2.3)$$

As an example of a contra-variant four vector, consider J^μ

$$J^\mu = (c\rho, J_x, J_y, J_z) \qquad (15.2.4)$$

Rewriting equation (14.2.4) slightly, we have

$$c\bar{\rho} = \gamma\left(c\rho - \frac{\text{v}}{c} J_x\right) \qquad \bar{J}_x = \gamma\left(J_x - \frac{\text{v}}{c} c\rho\right) \qquad \bar{J}_y = J_y \qquad \bar{J}_z = J_z \qquad (15.2.5)$$

and that therefore J^μ transforms like a four vector.

Any set of four things that transform in this manner we call a *contra-variant four vector*, and we denote such vectors by using a superscript for the index μ, which runs from 0 to 3, denoting the time and the three spatial components, in that order.

In contrast, we define a *covariant four vector* as any set of four things that transform as

$$\begin{pmatrix} \bar{S}_0 \\ \bar{S}_1 \\ \bar{S}_2 \\ \bar{S}_3 \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} \qquad (15.2.6)$$

and we denote such vectors by using a subscript for the index μ, which runs from 0 to 3. For example, consider the vector x_μ , defined by

$$x_\mu = (x_0, x_1, x_2, x_3) = (-ct, x, y, z) \tag{15.2.7}$$

(that is, all we have done is to change the sign of the time component). Then this set of four things transforms according to (15.2.6), and not according to (15.2.2).

Thus given a contra-variant four vector, we can always define a covariant counterpart of that vector by simply changing the sign of the time component. We always have that S^μ and S_μ are related by

$$\begin{pmatrix} S^0 \\ S^1 \\ S^2 \\ S^3 \end{pmatrix} = \begin{pmatrix} -S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} \tag{15.2.8}$$

Another example of a covariant vector is the differential operator ∂_μ , defined by

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = (\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}) = (\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}) \tag{15.2.9}$$

If we look at equations (14.2.1) and (14.2.2), and solve them for $\frac{\partial}{\partial \overline{t}}$ and $\frac{\partial}{\partial \overline{x}}$, we have

$$\frac{\partial}{\partial \overline{t}} = \gamma\left(\frac{\partial}{\partial t} + \mathrm{v}\frac{\partial}{\partial x}\right) \quad \text{and} \quad \frac{\partial}{\partial \overline{x}} = \gamma\left(\frac{\mathrm{v}}{c^2}\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right) \tag{15.2.10}$$

or in matrix form

$$\begin{pmatrix} \dfrac{\partial}{\partial \overline{x}^0} \\ \dfrac{\partial}{\partial \overline{x}^1} \\ \dfrac{\partial}{\partial \overline{x}^2} \\ \dfrac{\partial}{\partial \overline{x}^3} \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \dfrac{\partial}{\partial x^0} \\ \dfrac{\partial}{\partial x^1} \\ \dfrac{\partial}{\partial x^2} \\ \dfrac{\partial}{\partial x^3} \end{pmatrix} \tag{15.2.11}$$

which means that this is a covariant vector. In contrast, if we define the differential operator ∂^μ by

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = (\frac{\partial}{\partial x_0}, \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}) = (-\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}) \qquad (15.2.12)$$

then it transforms as a contra-variant vector.

15.3 The invariant length of a four vector, and the four "dot product"

Just as a three vector **A** has a length squared $A^2 = A_x{}^2 + A_y{}^2 + A_z{}^2$ that is invariant under ordinary spatial rotations, a four vector (contra-variant or covariant) has a "length" that is invariant under Lorentz transformations. The invariant "length" squared of a contra-variant four vector S^μ is $-\left(S^0\right)^2 + \left(S^1\right)^2 + \left(S^2\right)^2 + \left(S^3\right)^2$. The invariant "length" squared of a covariant four vector S_μ is $\left[-\left(S_0\right)^2 + \left(S_1\right)^2 + \left(S_2\right)^2 + \left(S_3\right)^2\right]$**,** which in light of (15.2.8), is exactly the as the length squared of the corresponding contra-variant vector. In fact, we have what corresponds to a four "dot product" of a four vector with itself,

$$S^\mu S_\mu = \sum_{\mu=0}^{3} S^\mu S_\mu = \left[-\left(S^0\right)^2 + \left(S^1\right)^2 + \left(S^2\right)^2 + \left(S^3\right)^2\right] = \left[-\left(S_0\right)^2 + \left(S_1\right)^2 + \left(S_2\right)^2 + \left(S_3\right)^2\right]$$
(15.3.1)

which does not change from system to system. Given any contra-variant vector a^μ and covariant vector b_μ, the four dot product is defined by

$$a^\mu b_\mu = a_\mu b^\mu = -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3 = -a_0 b_0 + a_1 b_1 + a_2 b_2 + a_3 b_3 \qquad (15.3.2)$$

Again, this four dot product yields the same value no matter what frame it is calculated in--it is a "Lorentz invariant", as can be shown directly by plugging in the transformations properties of the vectors. Note that in equation (15.3.2), we are using the following convention.

Whenever we have a contra-variant index and a covariant index repeated, there is an implied summation over that index from 0 to 3

Note that if we look at the differential operator ∂^μ and its covariant counterpart ∂_μ, we have for the four dot product $\partial^\mu \partial_\mu$ that

$$\partial^\mu \partial_\mu = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \qquad (15.3.3)$$

which since it is the form of a four dot product, must be the same in every Lorentz frame. Of course, we know that already, since we explicitly constructed the Lorentz transformation to guarantee that this is true.

15.4 Second rank four tensors

Just as in three dimensions, we can define second rank four tensors. Remember the way we defined a second rank tensor in three dimensions. If a three vector **A** transformed under spatial rotations like (see Problem 1-2 of Assignment 1)

$$\bar{A}_i = R_{ij} A_j \qquad (15.4.1)$$

then we defined the a second rank three tensor T_{ij} as any nine things which transformed as

$$\bar{T}_{ij} = R_{im} R_{jn} T_{mn} \qquad (15.4.2)$$

The easiest way to construct an object that transforms as a second rank three tensor is of course to take any two three vectors **A** and **B** and form a second rank tensor T_{ij} by setting $T_{ij} = A_i B_j$. This set of nine objects clearly transforms as (15.4.2) demands.

As we have seen over and over again since we first defined second rank three tensors, their main utility is that

> If $\overset{\leftrightarrow}{\mathbf{T}}$ is a second rank (three) tensor and **C** is any (three) vector, the dot product of **C** with $\overset{\leftrightarrow}{\mathbf{T}}$ "from the left" is a vector, $\mathbf{C}\cdot\overset{\leftrightarrow}{\mathbf{T}}$, and is given by $(\mathbf{C}\cdot\overset{\leftrightarrow}{\mathbf{T}})_j = C_i T_{ij}$. The dot product of **C** with $\overset{\leftrightarrow}{\mathbf{T}}$ "from the right" is a vector, $\overset{\leftrightarrow}{\mathbf{T}}\cdot\mathbf{C}$, and is given by $(\overset{\leftrightarrow}{\mathbf{T}}\cdot\mathbf{C})_j = T_{ji} C_i$. If $\overset{\leftrightarrow}{\mathbf{T}}$ is a symmetric these are the same vector.

We define second rank four tensors in a very analogous fashion. We define a second rank contra-variant four tensor as any set of sixteen objects $H^{\lambda\sigma}$ which transform as

$$\bar{H}^{\mu\nu} = \Lambda^\mu{}_\lambda \Lambda^\nu{}_\sigma H^{\lambda\sigma} \qquad (15.4.3)$$

The easiest way to construct an object that transforms as a second rank contra-variant four tensor is to take any two contra-variant four vectors A^λ and B^σ and form a second rank tensor $H^{\lambda\sigma}$ by setting $H^{\lambda\sigma} = A^\lambda B^\sigma$. This set of sixteen objects clearly transforms as (15.4.3) demands.

Just as in the three tensor case, the main utility of second rank four tensors is statements like

If $H^{\lambda\sigma}$ is a second rank contra-variant four tensor and C_λ is any covariant four vector, then the four dot product of C_λ with $H^{\lambda\sigma}$, $C_\lambda H^{\lambda\sigma}$ is a contra-variant four vector with contra-variant index σ. Again, we can define the four dot product from the left or the right, but for symmetric second rank four tensors, the result is the same, and we will only encounter symmetric tensors in electromagnetism.

15.5 The field tensor $F^{\lambda\sigma}$ and the transformation of E and B

We can define the four-vector potential A^μ by (see (14.4.2))

$$A^\mu = \left(\frac{\phi}{c}, A_x, A_y, A_z \right) \tag{15.5.1}$$

Consider the second rank contra-variant four tensor defined by

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{15.5.2}$$

What is this tensor? Well, if we write out the components of this tensor, we have for F^{01}

$$F^{01} = \partial^0 A^1 - \partial^1 A^0 = -\frac{1}{c}\frac{\partial}{\partial t} A_x - \frac{\partial}{\partial x}\frac{V}{c} = \frac{E_x}{c} \tag{15.5.3}$$

where we have used the fact that $\mathbf{E} = -\nabla\phi - \dfrac{\partial \mathbf{A}}{\partial t}$. What about components like F^{12}?

$$F^{12} = \partial^1 A^2 - \partial^2 A^1 = \frac{\partial}{\partial x} A_y - \frac{\partial}{\partial y} A_x = B_z \tag{15.5.4}$$

where we have used $\mathbf{B} = \nabla \times \mathbf{A}$. Proceeding in this way, we find that

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix} \tag{15.5.5}$$

Thus we see that the electric and magnetic fields transform as the components of a second rank four tensor, that is, in the manner described by given by (15.4.3). If we write this equation in the following form

$$\bar{F}^{\mu\nu} = \Lambda^{\mu}{}_{\lambda}\, \Lambda^{\nu}{}_{\sigma}\, F^{\lambda\sigma} = \sum_{\sigma}\sum_{\lambda} \Lambda^{\mu}{}_{\lambda} F^{\lambda\sigma} \left(\Lambda^{transpose}\right)^{\sigma}{}_{\nu} \tag{15.5.6}$$

where $\Lambda^{transpose}$ is the transpose of the matrix given in (15.2.3) , then this looks like matrix multiplication. In fact, we have

$$\bar{F}^{\mu\nu} = \begin{pmatrix} 0 & \bar{E}_X/c & \bar{E}_y/c & \bar{E}_z/c \\ -\bar{E}_x/c & 0 & \bar{B}_z & -\bar{B}_y \\ -\bar{E}_y/c & -\bar{B}_z & 0 & \bar{B}_x \\ -\bar{E}_Z/c & \bar{B}_y & -\bar{B}_x & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix} \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\gamma\beta E_x/c & \gamma E_x/c & E_y/c & E_z/c \\ -\gamma E_x/c & \gamma\beta E_x/c & B_z & -B_y \\ -\gamma E_y/c + \gamma\beta B_z & \gamma\beta E_y/c - \gamma B_z & 0 & B_x \\ -\gamma E_z/c - \gamma\beta B_y & \gamma\beta E_z/c + \gamma B_y & -B_x & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \gamma^2\left(\beta E_x/c - \beta E_x/c\right) & \left(\gamma^2 - \gamma^2\beta^2\right)E_x/c & \gamma\left(E_y/c - \gamma\beta B_z\right) & \gamma\left(E_z/c + \gamma\beta B_y\right) \\ -\left(\gamma^2 - \gamma^2\beta^2\right)E_x/c & \gamma^2\left(\beta E_x/c - \beta E_x/c\right) & -\gamma\beta E_y/c + \gamma B_z & -\gamma\beta E_z/c - \gamma B_y \\ -\gamma E_y/c + \gamma\beta B_z & \gamma\beta E_y/c - \gamma B_z & 0 & B_x \\ -\gamma E_z/c - \gamma\beta B_y & \gamma\beta E_z/c + \gamma B_y & -B_x & 0 \end{pmatrix}$$

$$\bar{F}^{\mu\nu} = \begin{pmatrix} 0 & E_x/c & \gamma\left(E_y - \mathrm{v}B_z\right)/c & \gamma\left(E_z + \mathrm{v}B_y\right)/c \\ -E_x/c & 0 & +\gamma\left(B_z - \mathrm{v}E_y/c^2\right) & -\gamma\left(B_y + \mathrm{v}E_z/c^2\right) \\ -\gamma\left(E_y - \mathrm{v}B_z\right)/c_z & -\gamma\left(B_z - \mathrm{v}E_y/c^2\right) & 0 & B_x \\ -\gamma\left(E_z + \mathrm{v}B_y\right)/c & \gamma\left(B_y + \mathrm{v}E_z/c^2\right) & -B_x & 0 \end{pmatrix} \tag{15.5.7}$$

and if we just pick off the components of the first and last matrices in (15.5.7) we have

$$\bar{E}_x = E_x \qquad \bar{E}_y = \gamma\left(E_y - \mathrm{v}\,B_z\right) \qquad \bar{E}_z = \gamma\left(E_z + \mathrm{v}\,B_y\right) \tag{15.5.8}$$

$$\overline{B}_x = B_x \qquad \overline{B}_y = \gamma\left(B_y + \frac{\mathrm{v}}{c^2}E_z\right) \qquad \overline{B}_z = \gamma(B_z - \frac{\mathrm{v}}{c^2}E_y) \tag{15.5.9}$$

which is the same as we obtained before using Einstein's approach. These equations can also be written as

$$\overline{E}_{\parallel} = E_{\parallel} \qquad \overline{\mathbf{E}}_{\perp} = \gamma\left(\mathbf{E}_{\perp} + \mathbf{v}\times\mathbf{B}\right) \tag{15.5.10}$$

$$\overline{B}_{\parallel} = B_{\parallel} \qquad \overline{\mathbf{B}}_{\perp} = \gamma\left(\mathbf{B}_{\perp} - \frac{1}{c^2}\mathbf{v}\times\mathbf{E}\right) \tag{15.5.11}$$

where parallel and perpendicular refer to the direction of the relative velocity $\mathbf{v} = \mathrm{v}\,\hat{\mathbf{x}}$

We can also define the covariant second rank four tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, which has the form in terms of **E** and **B**

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ +E_x/c & 0 & B_z & -B_y \\ +E_y/c & -B_z & 0 & B_x \\ +E_z/c & B_y & -B_x & 0 \end{pmatrix} \tag{15.5.12}$$

Just as in three space, we can define the totally anti-symmetric forth rank four tensor $\varepsilon^{\mu\nu\lambda\sigma}$

$$\varepsilon^{\mu\nu\lambda\sigma} = \begin{cases} +1 \text{ if } \mu\nu\lambda\sigma \text{ is an even permutation of 0123} \\ -1 \text{ if } \mu\nu\lambda\sigma \text{ is an odd permutation of 0123} \\ \qquad 0 \text{ otherwise} \end{cases} \tag{15.5.13}$$

The *dual tensor* $G^{\mu\nu}$ to the field tensor $F^{\mu\nu}$ is defined by $G^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\lambda\sigma}F_{\lambda\sigma}$. This has the form (cf. page 501 of *Griffiths*)

$$G^{\mu\nu} = \begin{pmatrix} 0 & B_x & B_y & B_z \\ -B_x & 0 & -E_z/c & E_y/c \\ -B_y & E_z/c & 0 & -E_x/c \\ -B_z & -E_y/c & E_x/c & 0 \end{pmatrix} \tag{15.5.14}$$

15.6 The manifestly covariant form of Maxwell's Equations

First, consider the four divergence of $F^{\mu\nu}$, that is $\partial_\mu F^{\mu\nu}$. We have from (15.2.9) and (15.5.5) that

$$\partial_\mu F^{\mu\nu} = (\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}) \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix} \tag{15.6.1}$$

or

$$\partial_\mu F^{\mu\nu} = \begin{pmatrix} -\frac{\nabla\cdot\mathbf{E}}{c} \\ \frac{1}{c^2}\frac{\partial E_x}{\partial t} - \frac{\partial B_z}{\partial y} + \frac{\partial B_y}{\partial z} \\ \frac{1}{c^2}\frac{\partial E_y}{\partial t} + \frac{\partial B_z}{\partial x} - \frac{\partial B_x}{\partial z} \\ \frac{1}{c^2}\frac{\partial E_z}{\partial t} - \frac{\partial B_y}{\partial x} + \frac{\partial B_x}{\partial y} \end{pmatrix} = \begin{pmatrix} -\mu_o \varepsilon_o c \nabla\cdot\mathbf{E} \\ \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} - \nabla\,\mathbf{x}\,\mathbf{B} \end{pmatrix} = -\mu_o \begin{pmatrix} c\rho \\ \mathbf{J} \end{pmatrix} \tag{15.6.2}$$

so that we see that two or our four Maxwell's equations are contained in the following manifestly covariant equation

$$\partial_\mu F^{\mu\nu} = -\mu_o J^\nu \tag{15.6.3}$$

This is “manifestly covariant” because it is a relationship between four vectors, and the equation has this form regardless of the system, because of the way four vectors transform. We find the other two Maxwell's Equations are contained in the equation

$$\partial_\mu G^{\mu\nu} = 0 \tag{15.6.4}$$

since

$$\partial_\mu G^{\mu\nu} = (\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}) \begin{pmatrix} 0 & B_x & B_y & B_z \\ -B_x & 0 & -E_z/c & E_y/c \\ -B_y & E_z/c & 0 & -E_x/c \\ -B_z & -E_y/c & E_x/c & 0 \end{pmatrix} \tag{15.6.5}$$

or

$$\partial_\mu G^{\mu\nu} = \begin{pmatrix} -\nabla\cdot B \\ \frac{1}{c}\left(\frac{\partial B_x}{\partial t}+\frac{\partial E_z}{\partial y}-\frac{\partial E_y}{\partial z}\right) \\ \frac{1}{c}\left(\frac{\partial B_x}{\partial t}-\frac{\partial E_z}{\partial x}+\frac{\partial E_x}{\partial z}\right) \\ \frac{1}{c}\left(\frac{\partial B_x}{\partial t}+\frac{\partial E_y}{\partial x}-\frac{\partial E_x}{\partial y}\right) \end{pmatrix} = \begin{pmatrix} -\nabla\cdot\mathbf{B} \\ \frac{1}{c}\left(\frac{\partial \mathbf{B}}{\partial t}+\nabla\,\mathbf{x}\,\mathbf{E}\right) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{15.6.6}$$

What about our other equations? Well, charge conservation in four vector form is

$$\partial_\mu J^\mu = 0 \tag{15.6.7}$$

and the Lorentz gauge condition equation (6.1.8)on the four vector potential is

$$\partial_\mu A^\mu = 0 \tag{15.6.8}$$

where of course (cf. (23)) the four vector potential satisfies

$$\partial^\nu \partial_\nu \;\; A^\mu = -\mu_o J^\mu \tag{15.6.9}$$

All of these equations are manifestly covariant--that is, they have obvious transformation properties that guarantee that they will remain the same in form from one inertial frame to the next.

15.7 The conservation of energy and momentum in four vector form

We can define the contra-variant second rank four tensor $\Theta^{\mu\nu}$ by

$$\Theta^{\mu\nu} = \begin{pmatrix} \frac{1}{2}\left(\varepsilon_o E^2 + B^2/\mu_o\right) & S_x/c & S_y/c & S_z/c \\ S_x/c & -T_{xx} & -T_{xy} & -T_{xz} \\ S_y/c & -T_{yx} & -T_{yy} & -T_{yz} \\ S_z/c & -T_{zx} & -T_{zy} & -T_{zz} \end{pmatrix} \tag{15.7.1}$$

where $\mathbf{S}$ is the Poynting vector, $\mathbf{S} = \frac{\mathbf{E}\,\mathbf{x}\,\mathbf{B}}{\mu_o}$, and $\overleftrightarrow{\mathbf{T}}$ is the Maxwell stress tensor. If we take the four divergence of this four tensor, we have

$$\partial_\mu \Theta^{\mu\nu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \begin{pmatrix} \frac{1}{2}\left(\varepsilon_o E^2 + B^2/\mu_o\right) & S_x/c & S_y/c & S_z/c \\ S_x/c & -T_{xx} & -T_{xy} & -T_{xz} \\ S_y/c & -T_{yx} & -T_{yy} & -T_{yz} \\ S_z/c & -T_{zx} & -T_{zy} & -T_{zz} \end{pmatrix} \quad (15.7.2)$$

$$\partial_\mu \Theta^{\mu\nu} = \begin{pmatrix} \frac{1}{c}\frac{\partial}{\partial t}\frac{1}{2}\left(\varepsilon_o E^2 + B^2/\mu_o\right) + \frac{\nabla \cdot \mathbf{S}}{c} \\ \frac{\partial}{\partial t}\left(\varepsilon_o \mathbf{E} \,\mathbf{x}\, \mathbf{B}\right) + \nabla \cdot \left(-\vec{\mathbf{T}}\right) \end{pmatrix} = \begin{pmatrix} -\frac{\mathbf{J}\cdot\mathbf{E}}{c} \\ -\left(\rho\mathbf{E} + \mathbf{J} \,\mathbf{x}\, \mathbf{B}\right) \end{pmatrix} \quad (15.7.3)$$

where we have used the conservation of energy and momentum that we have previously derived to arrive at the last expression. However, we have

$$F^{\mu\sigma} J_\sigma = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix} \begin{pmatrix} -c\rho \\ J_x \\ J_y \\ J_z \end{pmatrix} = \begin{pmatrix} \frac{\mathbf{J}\cdot\mathbf{E}}{c} \\ \left(\rho\mathbf{E} + \mathbf{J} \,\mathbf{x}\, \mathbf{B}\right) \end{pmatrix} \quad (15.7.4)$$

It is clear that conservation of energy and momentum in four vector form is expressed by the equation

$$\partial_\mu \Theta^{\mu\nu} = -F^{\nu\sigma} J_\sigma \quad (15.7.5)$$

16 Relativistic Particle Dynamics

16.1 Learning Objectives

We now then turn to the question of relativistic particle dynamics.

16.2 Now for something completely different

So far, we have made Maxwell's Equations look a lot prettier, but we have added no new information by introducing our manifestly covariant formulation.

However, we noted above that Newton's Laws preserved the Principle of Relativity under Galiliean transformations, but they do not preserve that principle under Lorentz transformations. How do we reconcile this with Special Relativity? What we need to do is to try to modify Newton's Laws so that they transform correctly under Lorentz transformations. We do this by requiring that our equations for particle motion be manifestly covariant--that is, that they can be written in four vector form. Moreover, they must reduce to the familiar form at small velocities compared to *c*.

We can in fact find such equations, without a lot of trouble. This shows the power of the covariant formulation. It will tells us how to change Newton's Laws so that they are correct for relativistic motion, merely by requiring that the form of the equations be covariant.

Consider a single charged particle moving with mass *m* and charge *q* in given electric and magnetic fields. Let **X**(t) be the spatial position of the particle at time *t*. For a given space-time experiment, we measure **X**(t) at time *t* using our infinite set of coordinate observers, as described previously. In the non-relativistic world, we used to say that once we have made a set of measurements for the particle motion in given fields, those measurements will satisfy the differential equation

$$m\,\mathbf{a} = m\frac{d^2}{d\,t^2}\mathbf{X}(t) = q\left(\mathbf{E} + \mathbf{u}\times\mathbf{B}\right) \qquad (16.2.1)$$

where

$$\mathbf{u}(t) = \frac{d\,\mathbf{X}(t)}{d\,t} \qquad \text{and} \qquad \mathbf{a}(t) = \frac{d\,\mathbf{u}(t)}{d\,t} = \frac{d^2\,\mathbf{X}(t)}{d\,t^2} \qquad (16.2.2)$$

The vector **u** is the ordinary three space velocity--the velocity your infinite grid of observers compute from their observations of particle position **X**(t) versus coordinate time *t*. To be absolutely clear about this, consider how we would compute **u**(t) at time t_a along the trajectory of the particle **X**(t):

$$\mathbf{u}(t_b) = \lim_{t_b\to t_a}\frac{\mathbf{X}(t_b) - \mathbf{X}(t_a)}{t_b - t_a} = \lim_{\Delta t\to 0}\frac{\mathbf{X}(t_a + \Delta t) - \mathbf{X}(t_a)}{\Delta t} \qquad (16.2.3)$$

Where $\Delta t = t_b - t_a$. That is, once we have collected all our observer notebooks after a given experiment is over, we reconstruct what **X**(t) was during the experiment, and we also calculate things like **u**(t), or the ordinary three space acceleration **a**(t), by performing computations on our data like (16.2.3), as well as like

$$\mathbf{a}(t_b) = \lim_{t_b\to t_a}\frac{\mathbf{u}(t_b) - \mathbf{u}(t_a)}{t_b - t_a} = \lim_{\Delta t\to 0}\frac{\mathbf{u}(t_a + \Delta t) - \mathbf{u}(t_a)}{\Delta t} \qquad (16.2.4)$$

Now, we want to write equation (16.2.1) in a properly covariant form. The problem is that even though this equation involves basic observables in a given inertial frame, things like ordinary velocity **u** have terrible transformation properties from one inertial frame to another. Why? Because when we compute **u** in any given frame for some time interval Δt, we are differentiating with respect to the change in coordinate time Δt in that frame, and this changes from one co-moving inertial frame to another, because of the way that space and time transform.

There is however a time-like measure of the separation between two space-time events *a* and *b* on which observers in all inertial frames will agree. This is the combination

$$\begin{aligned} c^2\Delta\tau^2 &= c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = -\left(b_\mu - a_\mu\right)\left(b^\mu - a^\mu\right) \\ &= c^2\Delta\bar{t}^{\,2} - \Delta\bar{x}^{\,2} - \Delta\bar{y}^{\,2} - \Delta\bar{z}^{2} = -\left(\bar{b}_\mu - \bar{a}_\mu\right)\left(\bar{b}^{\,\mu} - \bar{a}^{\,\mu}\right) \end{aligned} \tag{16.2.5}$$

where a_μ is the four vector location of event *a* as seen in the unbarred frame, and b_μ is the four vector location of event *b* as seen in the unbarred frame, etc. The Lorentz transformation leaves this quantity invariant, so that no matter who calculates it in what inertial frame, the answer is always the same. In differential form, we have

$$d\tau^2 = dt^2 - \frac{\left(dx^2 + dy^2 + dz^2\right)}{c^2} = dt^2\left(1 - \frac{\left(dx^2 + dy^2 + dz^2\right)}{c^2 dt^2}\right) = dt^2\left(1 - \frac{\mathbf{u}^2(t)}{c^2}\right) \tag{16.2.6}$$

or

$$d\tau = dt\sqrt{\left(1 - \frac{\mathbf{u}^2(t)}{c^2}\right)} \tag{16.2.7}$$

Clearly this is the time like parameter we want to differentiate with respect to get nice transformation properties. Physically, the proper time $d\tau$ separating events *a* and *b* , assuming that event *b* occurs very close to event *a*, corresponds to the amount of time that would be measured in that inertial frame at which the particle appears to be instantaneously at rest at time t_a, that is an inertial frame moving at speed $\mathbf{v} = \mathbf{u}(t_a)$. This proper time τ is also the time that would pass if you were riding with the particle.

16.3 The four velocity and the four acceleration

We now are in a position to define the four velocity and four acceleration. The space-time trajectory of the particle $X^\mu(t)$ is given by

$$X^\mu(t) = \begin{pmatrix} ct \\ \mathbf{X}(t) \end{pmatrix} \tag{16.3.1}$$

But since we have $d\tau = dt\sqrt{\left(1 - \frac{\mathbf{u}^2(t)}{c^2}\right)}$, we can compute how *t* and τ are related--that is, we can find $t(\tau)$ or . In practice this can be quite difficult (we give a specific example later), but in principle it is clear that we can do this. So we can treat $X^\mu(t)$ as a function of τ or of *t*, that is $X^\mu(\tau) = X^\mu(t(\tau))$. We can define the four velocity η^μ by the equation:

$$\eta^{\mu}=\frac{d}{d\tau}X^{\mu}(\tau)=\frac{d}{d\tau}\begin{pmatrix}ct(\tau)\\ \mathbf{X}(t(\tau))\end{pmatrix}=\frac{1}{\sqrt{1-\frac{u^2}{c^2}}}\frac{d}{dt}\begin{pmatrix}ct\\ \mathbf{X}(t)\end{pmatrix}=\begin{pmatrix}\gamma_u c\\ \gamma_u\mathbf{u}(t)\end{pmatrix} \qquad (16.3.2)$$

Where $\gamma_u(t)=1/\sqrt{1-\frac{u(t)^2}{c^2}}$. We use the subscript on $\gamma_u(t)$ to remind us that this is not the constant gamma associated with going from one inertial frame to another via a Lorentz transformation. This gamma is a function of time, and is based on the time varying particle velocity **u**(t). Also, in the third step in equation (16.3.2), we have used the differential relation (16.2.7) to convert from the derivative with respect to proper time to the derivative with respect to coordinate time. Similarly, we can define the four acceleration Ξ^{μ} as

$$\Xi^{\mu}=\frac{d}{d\tau}\eta^{\mu}=\frac{1}{\sqrt{1-\frac{u^2(t)}{c^2}}}\frac{d}{dt}\eta^{\mu}=\frac{1}{\sqrt{1-\frac{u^2(t)}{c^2}}}\frac{d}{dt}\frac{1}{\sqrt{1-\frac{u^2(t)^2}{c^2}}}\begin{pmatrix}c\\ \mathbf{u}(t)\end{pmatrix} \qquad (16.3.3)$$

Taking the *t* derivatives in (16.3.3), we obtain

$$\Xi^{\mu}=\gamma_u{}^2\begin{pmatrix}\gamma_u{}^2\left(\frac{\mathbf{u}\cdot\mathbf{a}}{c}\right)\\ \mathbf{a}(t)+\gamma_u{}^2\mathbf{u}\left(\frac{\mathbf{u}\cdot\mathbf{a}}{c^2}\right)\end{pmatrix} \qquad (16.3.4)$$

16.4 The equation of motion

Now, given these four vectors, let us see if we can find a manifestly covariant equation that reduces to (16.2.1) for small velocities compared to c. Well, if we use equations (15.5.5) and (16.3.2), we have the suggestive result that

$$F^{\mu\sigma}\eta_{\sigma}=\begin{pmatrix}0 & E_x/c & E_y/c & E_z/c\\ -E_x/c & 0 & B_z & -B_y\\ -E_y/c & -B_z & 0 & B_x\\ -E_z/c & B_y & -B_x & 0\end{pmatrix}\begin{pmatrix}-\gamma_u c\\ \gamma_u u_x\\ \gamma_u u_y\\ \gamma_u u_z\end{pmatrix}=\begin{pmatrix}\gamma_u\frac{\mathbf{E}\cdot\mathbf{u}}{c}\\ \gamma_u\left(\mathbf{E}+\mathbf{u}\times\mathbf{B}\right)\end{pmatrix} \qquad (16.4.1)$$

In we compare (16.2.1) and (16.4.1), we see that the covariant equation

$$m\frac{d}{d\tau}\eta^{\mu}=qF^{\mu\sigma}\eta_{\sigma} \qquad (16.4.2)$$

will reduce to (16.2.1)in the limit of small velocities compared to the speed of light. Writing this out component by component, we have

$$m\frac{d}{d\tau}\begin{pmatrix}\gamma_u c\\ \gamma_u \mathbf{u}(t)\end{pmatrix} = m\gamma_u \frac{d}{dt}\begin{pmatrix}\gamma_u c\\ \gamma_u \mathbf{u}(t)\end{pmatrix} = q\begin{pmatrix}\gamma_u \dfrac{\mathbf{E}\cdot\mathbf{u}}{c}\\ \gamma_u\left(\mathbf{E}+\mathbf{u}\,\mathbf{x}\,\mathbf{B}\right)\end{pmatrix} \tag{16.4.3}$$

The time component of (16.4.3) is

$$\frac{d}{dt}m\gamma_u c^2 = q\,\mathbf{E}\cdot\mathbf{u} \tag{16.4.4}$$

and the spatial component of (16.4.3) is

$$\frac{d}{dt}m\,\gamma_u\mathbf{u}(t) = q\left(\mathbf{E}+\mathbf{u}\times\mathbf{B}\right) \tag{16.4.5}$$

This is really quite amazing. We have not only found a dynamic equation (16.4.2) that reduces to Newton's form at low speeds, but we have also ended up with something in (16.4.4) that is totally different than anything we have seen before. The right side of equation (16.4.4) is clearly the rate at which work is being done on our charge by the electric field, so the left hand side of (16.4.4) must be the time rate of change of the energy of the particle. Therefore we must have

$$Energy\ of\ \ particle = m\gamma_u\,c^2 = \frac{m\,c^2}{\sqrt{1-\dfrac{u^2}{c^2}}} \approx m\,c^2 + \frac{1}{2}m\,u^2 + \ldots \ \text{ for } \ u \ll c \tag{16.4.6}$$

Something totally different.

16.5 An example of relativistic motion

In Eq. (16.4.3), we found that

$$m\frac{d}{d\tau}\begin{pmatrix}\gamma_u c\\ \gamma_u \mathbf{u}(t)\end{pmatrix} = m\gamma_u \frac{d}{dt}\begin{pmatrix}\gamma_u c\\ \gamma_u \mathbf{u}(t)\end{pmatrix} = q\begin{pmatrix}\gamma_u \dfrac{\mathbf{E}\cdot\mathbf{u}}{c}\\ \gamma_u\left(\mathbf{E}+\mathbf{u}\,\mathbf{x}\,\mathbf{B}\right)\end{pmatrix} \tag{16.5.1}$$

Let's look at a specific example of particle motion using (16.5.1). Specifically, let's consider the motion of a charge in a constant electric field $\mathbf{E} = E_o\,\hat{\mathbf{x}}$, with $\mathbf{B} = 0$. Assume that at $t = 0$, the charge is at rest at the origin. We want to find its subsequent motion.

We can do this two different ways. We can either solve for things as a function of coordinate time *t*, or as a function of proper time τ. Let's first do it in terms of coordinate time. With this electric field and initial conditions, $\mathbf{u} = u\,\hat{\mathbf{x}}$, the spatial part of (16.5.1) is

$$\frac{d}{dt} m\gamma_u u = qE_o \tag{16.5.2}$$

The speed is zero at *t = 0*. The solution of this equation for *u(t)* is given below in (16.5.3). Also, we give the limits for *small* times and for *large* times.

$$\gamma_u u = \frac{u}{\sqrt{1-\dfrac{u^2}{c^2}}} = \frac{qE_o t}{m} \qquad u(t) = \frac{\dfrac{qE_o t}{m}}{\sqrt{1+\left(\dfrac{qE_o t}{mc}\right)^2}} = \begin{cases} \dfrac{qE_o t}{m} & t \ll \dfrac{mc}{qE_o} \\ c & t \gg \dfrac{mc}{qE_o} \end{cases} \tag{16.5.3}$$

If we integrate our expression in (16.5.3) once more with respect to time to obtain *x(t)*, we find the expression given in Eq., and we also give expressions for *x(t)* for small and large time limits, as above.

$$x(t) = \frac{mc^2}{qE_o}\left[\sqrt{1+\left(\frac{qE_o t}{mc}\right)^2} - 1\right] = \begin{cases} \dfrac{1}{2}\dfrac{qE_o}{m}t^2 & t \ll \dfrac{mc}{qE_o} \\ ct & t \gg \dfrac{mc}{qE_o} \end{cases} \tag{16.5.4}$$

Lets do this another way. We solve this problem from scratch as a function of proper time τ. If we look at the space and time components of (16.5.1), we can derive a second order differential equation for $\gamma_u(\tau)$. This equation, and its solutions given our initial conditions, is as follows:

$$\frac{d}{d\tau}\gamma_u = \frac{qE_o}{m}\gamma_u u \tag{16.5.5}$$

$$\frac{d}{d\tau}\gamma_u u = \frac{qE_o}{m}\gamma_u \tag{16.5.6}$$

$$\frac{d^2}{d\tau^2}\gamma_u = \left(\frac{qE_o}{mc}\right)^2 \gamma_u \tag{16.5.7}$$

$$\gamma_u = \cosh\left(\frac{qE_o}{mc}\tau\right) \tag{16.5.8}$$

We can now get *t* as a function of τ by using $dt = \gamma_u d\tau$, as follows:

$$dt = \gamma_u d\tau = \cosh\left(\frac{qE_o}{mc}\tau\right)d\tau \tag{16.5.9}$$

$$\frac{qE_o}{mc}t = \sinh\left(\frac{qE_o}{mc}\tau\right) \tag{16.5.10}$$

For short times, $t << \frac{mc}{qE_o}$, we have $u << c$, and the proper time τ and the coordinate time t are equal. For long times $t >> \frac{mc}{qE_o}$, u is about c, and the relation between proper time and coordinate time is

$$\frac{qE_o}{mc}t = \frac{1}{2}e^{\left(\frac{qE_o}{mc}\tau\right)} \qquad \text{or} \qquad \frac{qE_o}{mc}\tau = \ln\left[\frac{2qE_o}{mc}t\right] \tag{16.5.11}$$

17 Radiation by a charge in arbitrary motion

17.1 Learning Objectives

We consider time dilation and space contraction. We then return to the subject of radiation, and look at the radiation emitted by a charge in arbitrary motion, including relativistic motion.

17.2 Time dilation and space contraction

17.2.1 Time dilation

Moving clocks run slower. To see this consider a clock at rest in our barred system. Let one point in space-time be $(c\bar{t},\bar{x},\bar{y},\bar{z}) = (0,0,0,0)$ and another point in space-time be $(c\bar{t},\bar{x},\bar{y},\bar{z}) = (c\Delta\bar{t},0,0,0)$. Then using the Lorentz transformation that takes us from the barred to the unbarred frame, that is

$$x^\mu = \sum_{\nu=0}^{3}\left(\Lambda^{-1}\right)^\mu{}_\nu\,\bar{x}^\nu \qquad \left(\Lambda^{-1}\right)^\mu{}_\nu = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{17.2.1}$$

we have that the origin in the barred frame transforms into the origin in the unbarred frame, and that $(c\Delta\bar{t},0,0,0)$ transforms into $(c\gamma\,\Delta\bar{t},\gamma\,\mathrm{v}\Delta\bar{t},0,0)$. Thus the time interval of $\Delta t = \gamma\,\Delta\bar{t}$ in the unbarred frame looks like a longer time as compared to the barred frame, i.e. the clock at rest in the barred frame is running slower as observed in the unbarred frame. That is, time is dilated.

17.2.2 Space Contraction

Moving rulers are shorter. To see this consider a ruler at rest in our barred system. Let the left end of the ruler be located at $(c\bar{t},\bar{x},\bar{y},\bar{z}) = (0,0,0,0)$ and the right end at $(c\bar{t},\bar{x},\bar{y},\bar{z}) = (c\bar{t},\bar{L},0,0)$ (in a minute you will see why we leave the time $\bar{t}$ unspecified). Then using the Lorentz transformation that takes us from the barred to the unbarred frame we have that the origin in the barred frame transforms into the origin in the unbarred frame, and that $(c\bar{t},\bar{L},0,0)$ transforms into $(c\gamma\bar{t}+\beta\gamma\bar{L}+,\gamma\mathrm{v}\bar{t}+\gamma\bar{L},0,0)$. If we want to know the length of this moving ruler as seen in the unbarred frame, we must measure the position of the left end at the same time as we measure the position of the right end in the unbarred frame, which requires that $c\gamma\bar{t}+\beta\gamma\bar{L}=0$, or that we must take $c\bar{t}=-\beta\bar{L}$. When we make sure we are measuring the position of the left and right end at the same time in the unbarred frame, we thus measure a distance

$$\gamma\mathrm{v}\bar{t}+\gamma\bar{L}=\gamma\mathrm{v}\left(-\bar{L}\beta/c\right)+\gamma\bar{L}=\gamma\left(-\bar{L}\beta^2\right)+\gamma\bar{L}=\gamma\bar{L}\left(1-\beta^2\right)=\bar{L}/\gamma \tag{17.2.2}$$

Thus a ruler at rest in the barred frame with length $\bar{L}$ has a shorter length $L=\bar{L}/\gamma$ as seen in the unbarred frame. That is, length in the direction of motion is contracted.

Space contraction can help us understand in part the way that sources and fields transform. Consider for example a line charge along the *x*-axis at rest in the unbarred frame with charge per unit length λ, due to elemental charges of charge +*e* spaced a distance ΔL apart. So we have $\lambda = e/\Delta L$. If we make the assumption that the elemental charge +e is a Lorentz invariant, then in the barred frame the charge per unit length $\bar{\lambda}$ will be larger, because space contraction will lead to $\Delta\bar{L}=\Delta L/\gamma$, and therefore $\bar{\lambda}=e/\Delta\bar{L}=\gamma\,\Delta L$. This explains the way the fields transform in this case (see (15.5.8)) as well as why the sources transform the way they do in this case (see (15.2.5).

17.3 The Lienard-Wiechert potentials

I want to find the electromagnetic fields associated with a point charge in arbitrary motion (even relativistic). Let us first return to the general solution to the time-dependent equations of electromagnetism that I arrived at in (6.2.12), which I reproduce below.

$$\phi(\mathbf{r},t)=\frac{1}{4\pi\,\varepsilon_o}\int_{all\,time}dt'\int_{all\,space}\frac{\rho(\mathbf{r}',t')}{|\mathbf{r}-\mathbf{r}'|}dt'\delta(t-t'-|\mathbf{r}-\mathbf{r}'|/c)d^3x' \tag{17.3.1}$$

$$\mathbf{A}(\mathbf{r},t)=\frac{\mu_o}{4\pi}\int_{all\,time}dt'\int_{all\,space}\frac{\mathbf{J}(\mathbf{r}',t')}{|\mathbf{r}-\mathbf{r}'|}dt'\delta(t-t'-|\mathbf{r}-\mathbf{r}'|/c)d^3x' \qquad (17.3.2)$$

When I was applying equations (17.3.1) and (17.3.2) to the radiation from extended sources of charge and current, I first used the delta functions to do the dt' integrations, with the result that we ended up with expressions that looked like

$$\phi(\mathbf{r},t)=\frac{1}{4\pi\,\varepsilon_o}\int_{all\,space}\frac{\rho(\mathbf{r}',t'_{ret})}{|\mathbf{r}-\mathbf{r}'|}d^3x' \qquad (17.3.3)$$

where

$$t'_{ret}=t-|\mathbf{r}-\mathbf{r}'|/c \qquad (17.3.4)$$

Then I carried out the d^3x' integrations over the extended sources, and that occupied a large fraction of the effort leading to formulas like those for electric dipole radiation, and so on. However, when I am considering *from the outset* point sources, it is more appropriate to insert the charge density and current of a point charge into equations (17.3.1) and (17.3.2), and then do the d^3x' integrations using the delta functions associated with the point charge.

Consider a point charge whose position in space as a function of time is given by $\mathbf{X}(t')$. We can easily define its "ordinary" velocity and acceleration (see the discussion in Section 16.2 above) by

$$\mathbf{u}(t)=\frac{d\,\mathbf{X}(t)}{d\,t} \qquad \text{and} \qquad \mathbf{a}(t)=\frac{d\,\mathbf{u}(t)}{d\,t}=\frac{d^2\,\mathbf{X}(t)}{d\,t^2} \qquad (17.3.5)$$

The charge and current densities associated with this point particle are then given by

$$\rho(\mathbf{r}',t')=q\,\delta^3(\mathbf{r}'-\mathbf{X}(t')) \qquad \mathbf{J}(\mathbf{r}',t')=q\,\mathbf{u}(t')\,\delta^3(\mathbf{r}'-\mathbf{X}(t')) \qquad (17.3.6)$$

If we now insert these expressions into (17.3.1) and (17.3.2), we obtain

$$\phi(\mathbf{r},t)=\frac{1}{4\pi\,\varepsilon_o}\int dt'\int d^3x'\,q\,\delta^3(\mathbf{r}'-\mathbf{X}(t'))\,\frac{\delta\left(t-t'-|\mathbf{r}-\mathbf{r}'|/c\right)}{|\mathbf{r}-\mathbf{r}'|} \qquad (17.3.7)$$

and

$$\mathbf{A}(\mathbf{r},t)=\frac{\mu_o}{4\pi}\int dt'\int d^3x'\,q\,\mathbf{u}(t')\delta^3(\mathbf{r}'-\mathbf{X}(t'))\,\frac{\delta\left(t-t'-|\mathbf{r}-\mathbf{r}'|/c\right)}{|\mathbf{r}-\mathbf{r}'|} \qquad (17.3.8)$$

We now use the $\delta^3(\mathbf{r}'-\mathbf{X}(t'))$ delta functions to do the d^3x' integrations, giving for $\phi(\mathbf{r},t)$, for example

$$\phi(\mathbf{r},t) = \frac{1}{4\pi\,\varepsilon_o}\int dt'\, q\; \frac{\delta\left(t - t' - |\,\mathbf{r} - \mathbf{X}(t')| / c\right)}{|\mathbf{r} - \mathbf{r}'|} \tag{17.3.9}$$

Note that I have in one fell swoop gotten rid of all the pain that I went through in considering electric dipole radiation, for example, when I was dealing with d^3x' integrations over spatially extended sources. However, I have exchanged one form of pain for another, because I still have to do (and interpret) the remaining integral in (17.3.9).

This integral, because of the $|\,\mathbf{r} - \mathbf{X}(t')|$ in the argument of the delta function, is of the form $\int d\eta\; f(\eta)\, \delta(\lambda(\eta))$. If I change variables and integrate with respect to λ instead of η, I have

$$\int d\eta\; f(\eta)\, \delta(\lambda(\eta)) = \int d\lambda \left[\frac{d\eta}{d\lambda}\right] f(\eta(\lambda))\, \delta(\lambda) = f(\eta_0) \Big/ \left|\frac{d\lambda}{d\eta}\right|_{\eta=\eta_o} \tag{17.3.10}$$

where η_0 is a zero of $\lambda(\eta)$, that is $\lambda(\eta_o) = 0$. The absolute value signs appear in (17.3.10) for reasons which are explained fairly clearly in *Griffiths*..

So, I need to evaluate

$$\left|\frac{d}{dt'}\left[t - t' - |\,\mathbf{r} - \mathbf{X}(t')\,| / c\right]\right| = \left|-1 - \frac{d}{c\,dt'}\left|\,\mathbf{r} - \mathbf{X}(t')\right|\right| = 1 - \frac{\left[\mathbf{r} - \mathbf{X}(t')\right]\cdot \mathbf{u}(t')}{c\,|\,\mathbf{r} - \mathbf{X}(t')\,|} \tag{17.3.11}$$

I define the unit vector from the particle to the observer at time t' to be

$$\hat{\mathbf{n}}(t') = \frac{\mathbf{r} - \mathbf{X}(t')}{|\,\mathbf{r} - \mathbf{X}(t')\,|} \tag{17.3.12}$$

and the vector $\boldsymbol{\beta}(t')$ to be

$$\boldsymbol{\beta}(t') = \frac{\mathbf{u}(t')}{c} \tag{17.3.13}$$

then (17.3.11) becomes

$$\left|\frac{d}{dt'}\left[t - t' - |\,\mathbf{r} - \mathbf{X}(t')\,| / c\right]\right| = 1 - \hat{\mathbf{n}}(t')\cdot\boldsymbol{\beta}(t') \tag{17.3.14}$$

and equation (17.3.9) becomes, in light of equation (17.3.10)

$$\phi(\mathbf{r},t) = \frac{1}{\left[1-\hat{\mathbf{n}}(t'_{ret})\cdot\boldsymbol{\beta}(t'_{ret})\right]}\frac{q}{4\pi\,\varepsilon_o}\frac{1}{\left|\mathbf{r}-\mathbf{X}(t'_{ret})\right|} \tag{17.3.15}$$

where t'_{ret} is the zero of the argument of the delta function in equation (17.3.9), and therefore satisfies

$$c\left(t-t'_{ret}\right) = \left|\mathbf{r}-\mathbf{X}(t'_{ret})\right| \quad \text{or} \quad t'_{ret} = t-\left|\mathbf{r}-\mathbf{X}(t'_{ret})\right|/c \tag{17.3.16}$$

Similarly, the vector potential is given by

$$\mathbf{A}(\mathbf{r},t) = \frac{1}{\left[1-\hat{\mathbf{n}}(t'_{ret})\cdot\boldsymbol{\beta}(t'_{ret})\right]}\frac{\mu_o q\,\mathbf{u}(t'_{ret})}{4\pi}\frac{1}{\left|\mathbf{r}-\mathbf{X}(t'_{ret})\right|} = \frac{\mathbf{u}(t'_{ret})}{c^2}\phi(\mathbf{r},t) \tag{17.3.17}$$

These are the famous Lienard-Wiechert potentials.

17.4 The electric and magnetic fields of a point charge

We can obtain the fields from the potentials in the usual manner, that is by taking differentials in space and time *with respect to the observer's coordinates*, **r** and t. This is complicated, as we have seen before, because not only are there explicit dependences in ϕ and **A** on the observers coordinates, but there are implicit dependencies through the retarded time. It is clear from equation (17.3.16) that there is a complicated (and generally transcendental) relationship between *t*, **r**, and t'_{ret}. That is, t'_{ret} depends both on *t* and **r**, so the derivatives with respect to any function of t'_{ret} are involved. The treatment of Griffiths is the standard one, and we quote only the result here. The electric fields of a point charge in arbitrary motion are

$$\mathbf{E}(\mathbf{r},t) = \left[\frac{q}{4\pi\,\varepsilon_o}\frac{\hat{\mathbf{n}}-\boldsymbol{\beta}}{\gamma_u^2(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^3R^2}\right]_{ret} + \left[\frac{q}{4\pi\,\varepsilon_o}\frac{1}{c}\frac{\hat{\mathbf{n}}\ \mathbf{x}\left\{(\hat{\mathbf{n}}-\boldsymbol{\beta})\ \mathbf{x}\ \dot{\boldsymbol{\beta}}\right\}}{(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^3R}\right]_{ret} \tag{17.4.1}$$

and

$$\mathbf{B}(\mathbf{r},t) = \frac{1}{c}\left[\hat{\mathbf{n}}\ \mathbf{x}\ \mathbf{E}\right]_{ret} \tag{17.4.2}$$

where

$$\dot{\boldsymbol{\beta}}(t') = \frac{d}{dt'}\boldsymbol{\beta}(t') = \frac{\mathbf{a}(t')}{c},\quad R = \left|\mathbf{r}-\mathbf{X}(t')\right|,\ \text{and}\ \gamma_u^2 = 1/\left(1-\frac{u^2}{c^2}\right) \tag{17.4.3}$$

For velocities small compared to the speed of light (β small compared to 1), we recover the non-relativistic results we expect..

We emphasize that all of the derivatives in (17.4.1) are taken with respect to the particle's coordinate time , t', and not the observer's coordinate time, *t*, (both measured in the same coordinate system), and then the various terms are evaluated at the retarded time $t'_{ret} = t - \left| \mathbf{r} - \mathbf{X}(t'_{ret}) \right| / c$. . To show how different this is from taking ***the time derivatives with respect to the observer's time***, consider the expression for $\mathbf{E}(\mathbf{r},t)$ due originally to Heaviside and rediscovered and popularized by Feynman

$$\mathbf{E}(\mathbf{r},t) = \frac{q}{4\pi\,\varepsilon_o}\left[\frac{\hat{\mathbf{n}}}{R^2} + \frac{1}{Rc}\frac{d}{dt}\hat{\mathbf{n}} + \frac{1}{c^2}\frac{d^2}{dt^2}\hat{\mathbf{n}}\right]_{ret} \tag{17.4.4}$$

Amazingly enough, equation (17.4.4) is equivalent to equation (17.4.1) (which you can show after about 10 pages of equations). The difference is that the time derivatives in equation (17.4.4) are taken with respect to the observer's time *t*, and then evaluated at the retarded time t'_{ret}, whereas those in equation (17.4.1) are taken with respect to the particle time t' and then evaluated at the retarded time t'_{ret}. Clearly, there must be a complicated relationship between the observer's time and the retarded time, which we explore below.

But first, let us quote a few results for the rate at which energy is radiated using these fields. The angular distribution of the energy radiated into solid angle *dΩ*, *per unit time observer time t*, is given by (compare equation (10.3.6), and using (17.4.1))

$$\frac{dW_{rad}}{d\Omega\, dt} = \frac{r^2(\mathbf{E}_{rad}\,\mathbf{x}\,\mathbf{B}_{rad})\cdot\hat{\mathbf{n}}}{\mu_o} = \frac{r^2 E_{rad}^{\;2}}{c\mu_o} = \frac{r^2}{c\mu_o}\left[\frac{q}{4\pi\,\varepsilon_o}\frac{1}{c}\frac{\hat{\mathbf{n}}\times\left\{(\hat{\mathbf{n}}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}\right\}}{(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^3 R}\right]^2 \tag{17.4.5}$$

or

$$\frac{dW_{rad}}{d\Omega\, dt} = \frac{q^2}{\left(4\pi\right)^2 c\varepsilon_o}\frac{\left|\hat{\mathbf{n}}\times\left[(\hat{\mathbf{n}}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}\right]\right|^2}{(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^6} \tag{17.4.6}$$

Comparing this expression to (11.2.3), we see that in the non-relativistic limit we recover the dipole radiation rate per unit solid angle, as we expect.

Another expression we will need is the expression for the angular distribution of the energy radiated into solid angle *dΩ*, *per unit time along the particle trajectory time* t', which is given by

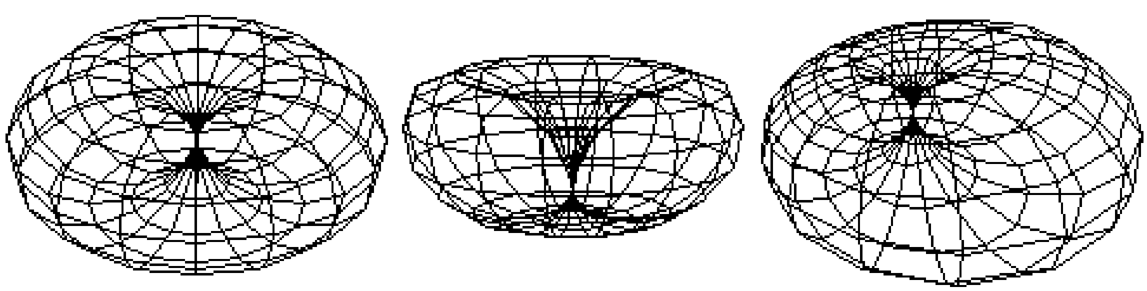
$$\frac{dW_{rad}}{d\Omega\, dt'} = \frac{dW_{rad}}{d\Omega\, dt}\frac{dt}{dt'} = \frac{q^2}{\left(4\pi\right)^2 c\varepsilon_o}\frac{\left|\hat{\mathbf{n}}\times\left[(\hat{\mathbf{n}}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}\right]\right|^2}{(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^5} \tag{17.4.7}$$

The $(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})$ term appears to the fifth power instead of the sixth power, because we need to multiply (17.4.6) by a factor of $(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})$ to convert from rate with respect to t to the rate along the particle trajectory t', as explained in more detail below. The total energy radiated into all solid angles, again per unit time along the particle's trajectory, can be found by integrating (17.4.7) over solid angle, giving

$$\frac{dW_{rad}}{dt'} = \frac{1}{4\pi\varepsilon_o}\frac{2q^2}{3c}\gamma^6\left[\left|\dot{\boldsymbol{\beta}}\right|^2 - \left|\boldsymbol{\beta}\ \mathbf{x}\ \dot{\boldsymbol{\beta}}\right|^2\right] \tag{17.4.8}$$

This quantity turns out to be a Lorentz scalar, and $c^2\gamma^6\left[\left|\dot{\boldsymbol{\beta}}\right|^2 - \left|\boldsymbol{\beta}\ \mathbf{x}\ \dot{\boldsymbol{\beta}}\right|^2\right]$ is the square of the acceleration of the charge in its instantaneous rest frame (see Section 17.7).

The $(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^5$ in the denominator of equation (17.4.7) causes a distortion in the radiation pattern we have previously seen for non-relativistic particles. Figure 17-1 gives examples of this. Figure 17-1(a) is just our familiar non-relativistic angular distribution of radiation.. The particle is at rest, and the acceleration is upward, and we get a distribution of radiation that is proportional to the square of the sine of the polar angle. In Figure 17-1(b), the acceleration is again upwards, but now there is a velocity of 0.3 the speed of light, also upwards. We see an enhancement of the radiation along the direction of the velocity. In Figure 17-1(c), the acceleration is still upwards, but the velocity is 0.1 the speed of light to the right, perpendicular to the acceleration. Again we see an enhancement along the direction of the velocity. As the velocity becomes closer and closer to the speed of light, the radiation is more and more beamed into the direction of the velocity. This happens because of the factor of $(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})$ to the fifth power in the denominator of (17.4.6). What is the physical origin of this beaming?



(a) (b) (c)

Figure 17-1: The radiation pattern for a charge in arbitrary motion

17.5 Appearance, reality, and the finite speed of light

There are two approaches to understand physically what is going on with this extreme emphasis on the forward direction (that is, the direction along the velocity vector) when the motion is relativistic. First, there is the effect of how the differential time dt is related to dt'. Remember that equation (17.3.16) holds, and if we differentiate this equation, we obtain

$$\frac{dt'_{ret}}{dt} = 1 - \frac{d}{c\,dt}\left|\mathbf{r}-\mathbf{X}(t'_{ret})\right| = 1 + \frac{1}{c}\frac{\mathbf{r}-\mathbf{X}(t'_{ret})}{\left|\mathbf{r}-\mathbf{X}(t'_{ret})\right|}\cdot\frac{d}{dt}\mathbf{X}(t'_{ret}) = 1 + \frac{1}{c}\hat{\mathbf{n}}\cdot\frac{d\,\mathbf{X}(t'_{ret})}{dt'_{ret}}\frac{dt'_{ret}}{dt} \quad (17.5.1)$$

or

$$\frac{dt'_{ret}}{dt} = 1 + \hat{\mathbf{n}}\cdot\boldsymbol{\beta}\frac{dt'_{ret}}{dt} \quad (17.5.2)$$

which can be written as

$$\frac{dt'_{ret}}{dt} = \frac{1}{(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})} \quad (17.5.3)$$

and also as

$$dt = (1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})\,dt'_{ret} \quad (17.5.4)$$

What does equation (17.5.4) mean? *Remember, this has nothing to do with time as measured in different co-moving frames. We are measuring all these times in one coordinate frame.* What (17.5.4) means is the following. The observer is sitting and watching the particle move in space, and recording what appears to happen. This is very different from using our infinite grid of observers who essentially function as all-seeing and all-knowing . When we limit ourselves to one observer and ask what that one observer "sees" as a function of time, then we get the effects of the finite propagation time for light to go from source to observer, and that is what (17.5.4) encompasses.

For example, suppose a particle is moving straight toward our observer at speed V, and emitting a "beep" of radiation every $\Delta t'$ seconds, which spreads out at the speed of light from the place where it was emitted. What will the observer say is the time interval between these beeps when they arrive at her position? Well, consider two beeps. Assume that the first beep is emitted at $t = 0$. Suppose also at time $t = 0$ that the observer and the source are separated by a distance D (see Figure 17-2)

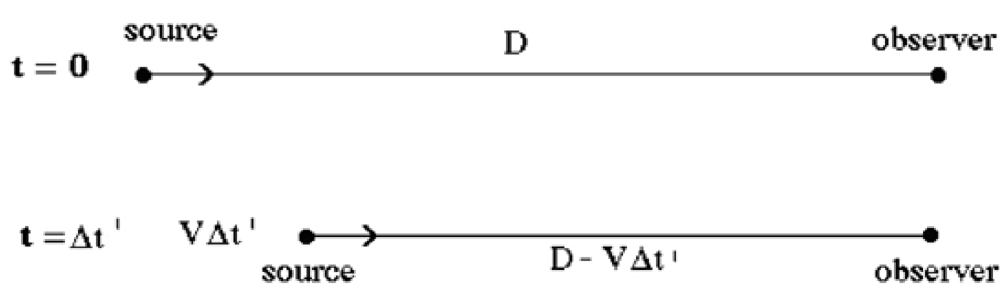


Figure 17-2: Beeps emitted by source as seen by the observer

The observer will see the first beep arrive at a time $t_1 = D/c$. How about the second beep? Well, if it is emitted a time $\Delta t'$ after the first beep, and the source is assumed to be moving directly toward the observer at speed *V*, the source will be only a distance $D - V\Delta t'$ from the observer when the second beep is emitted (see Figure 17-2). The second beep will then arrive at the observer at a time $t_2 = \Delta t' + \left(D - V\Delta t'\right)/c$ after the first beep. Since the time difference between beeps as seen by the observer is $\Delta t = t_2 - t_1$, we have

$$\Delta t = t_2 - t_1 = \left[\Delta t' + (D - V\Delta t')/c\right] - D/c = (1 - V/c)\Delta t' \tag{17.5.5}$$

This makes perfect sense. When the observer records the arrival of the beeps at her position, she finds a time interval between their arrival that is shorter than the time interval between the times that they were actually emitted along the particle's trajectory, because the second beep was emitted closer to her than the first beep, and therefore arrives sooner than one would expect based on the time $\Delta t'$ between the emitted times at the source.

Furthermore, this is exactly the effect predicted by equation (17.5.4) (remember the unit vector $\hat{\mathbf{n}}$ points from source to observer). One can generalize this argument for any angle between $\hat{\mathbf{n}}$ and the velocity of the source, and obtain the result (17.5.4) as a general result. For example, if the source is moving directly *away from* the observer, the time between beeps as recorded by the observer will be $(1+V/c)\Delta t'$, that is, the beeps will arrive further apart in time than $\Delta t'$, and that is again exactly what one expects from simple arguments like the one leading to (17.5.5). There is nothing fancy here to do with time dilation or times measured in different frames--all these times are measured in the same frame, we are just talking about the time separation between events on the particle's trajectory *as it appears to an observer relying on information propagating at finite speeds*.

Moreover, it makes sense that this effect will give a peak of emission rates in the forward direction. Suppose the beeps above were not light pulses, but pulses from a laser cannon on board the *Enterprise*, and the observer is not a Course 8 major but a Klingon warship at rest. Picard has accelerated the *Enterprise* up to $\left(1-10^{-6}\right)$ of the speed of light, and is heading directly toward the Klingon warship, firing his laser cannons at a rate one per second, as seen by our infinite grid of observers at rest with respect to the Klingon war ship (how fast Picard sees his cannons firing is a different question, which involves going from one co-moving frame to another). When these pulses reach the Klingon warship, they hit at a rate of 10^6 per second! Why? Because Picard is laying down these bursts in space one right behind the other, since the *Enterprise* is moving almost as fast as the bursts. He has an opportunity to lay down a huge number of them in the space in front of the *Enterprise*, and all of these bursts arrive at the warship in a very short time compared to the time Picard has been laying them down. This is all as seen in the same coordinate time, the time as measured a coordinate system at rest with respect to the Klingon warship. This is exactly the advantage a supersonic jet attack--if you rely only on sound to tell you that an attack is under way, then as soon as you hear that they are coming, they are already there.

17.6 How Nature counts charge

There is another way to get some physical understanding of why the factor $(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})$ pops up all over the place. We have already discussed this in part in Section 6.3. This perspective is also discussed at length in *Griffiths*, Section 10.3, and is in fact the way he comes to this factor in the Lienard-Wiechert potentials (we in contrast have given the standard mathematical derivation above).

First of all, let's introduce the notion of a space-time diagram (see Figure 17-3). In this kind of diagram, the vertical axis is the time axis, converted to distance by multiplying by the speed of light. The two horizontal axes are spatial axes. Any event in space-time can be described by the spatial coordinates at which it happened, and the time at which it happened, and is a point in this diagram. The *motion* of any physical object can be represented by a trajectory in space-time. If you are sitting at rest at the spatial origin, your space-time trajectory is simply a vertical line through the origin. If you are moving in the *x*-direction at a constant speed close to the speed of light, your space-time trajectory is a straight line in the *ct*-*x* plane at an angle just greater than 45 degrees to the x-axis. In Figure 17-3, we show two space-time trajectories. The one on the right is

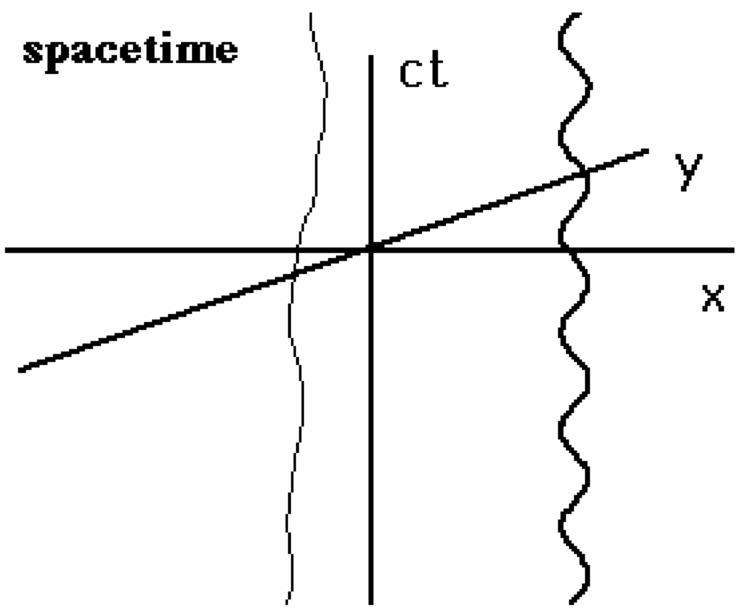


Figure 17-3: The space-time diagram

the trajectory of a particle moving in a circle in the *x*-*y* plane at a constant angular speed, with a speed around the circle of close to the speed of light. In a space-time diagram, this trajectory is just a spiral upwards. In the space-time diagram trajectory on the left, we have a particle meandering about at sub-light speeds.

Now, suppose that we are an observer sitting at the origin of space-time. What is it we see at a particular time, say $t = 0$? Well, what we see is all the things in the past which emitted radiation which has just arrived at the origin at $t = 0$. That is, we will see back in time to all events in space-time whose times $t < 0$. and positions (x,y) satisfy

$$r = \sqrt{(x^2 + y^2)} = -c\,t \tag{17.6.1}$$

This equation defines the backward light cone (see Figure 17-4 and also Figure 6-2). The backward light cone is a cone of events in this space-time diagram defined by the above equation. At any given instant in time $t < 0$, the cross-section of this cone in ordinary space (the x-y plane) is a circle of radius $-ct$. A signal radiated at $t < 0$ from any source on this circle of radius $r = -ct$ will reach the origin precisely at $t = 0$. . Conversely, any signal that is traveling at the speed of light and reaches the origin at precisely $t = 0$ must have been radiated from a source in space and time located somewhere on the backward light cone, and only on the backward light cone. Thus what the observer at the origin of space receives via radiation at $t = 0$ is information about all events in space and time located on his backward light cone, and no others. When you look around you, you are looking into the past, and what you "see" are things in the past that happened on your backward light cone.

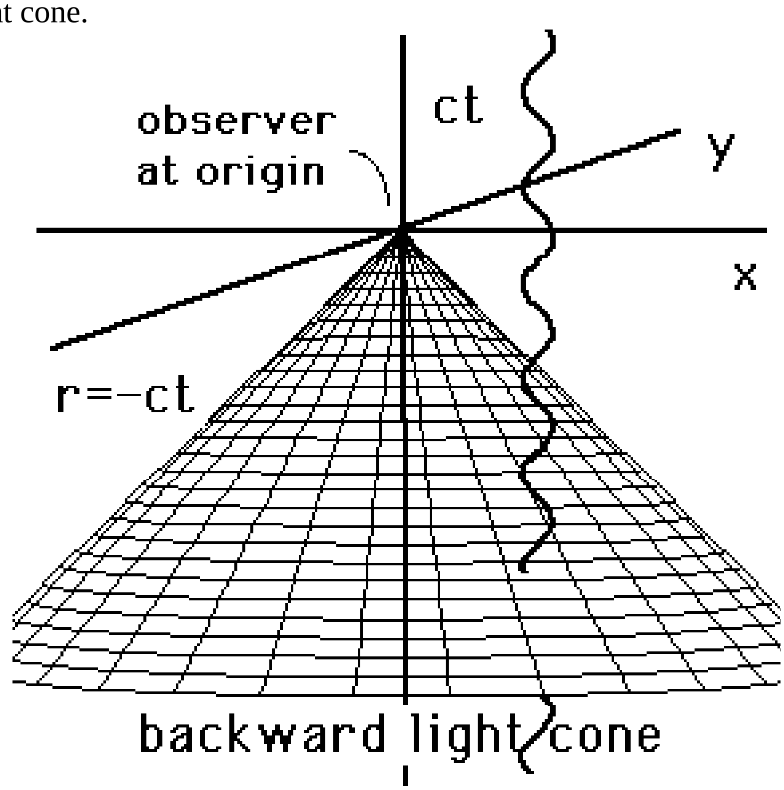


Figure 17-4: The observer's backward light cone.

Moreover, with a little thought you can see that if we plot the trajectory of a radiating charge in space-time, the condition that links the observers time t to the retarded time t'_{ret} (see equation (17.3.16)) is, for the case that the observer is at the origin at $t = 0$, $t'_{ret} = -\left|\mathbf{X}(t'_{ret})\right|/c$. But this is just saying that at the retarded time the particle's space-time trajectory lies on the observer's backward light cone. We illustrate this for a particular space-time trajectory in the above diagram. At a given observer time, the observer will "see" radiation from that point along the space-time trajectory of the charge which intersects or "cuts" his backward light cone. It is fairly easy to see geometrically that as long as the charge's speed is everywhere less than the speed of light, the space-time trajectory of the charge can intersect the observers backward light cone at one and only one point in space-time, and this happens at the time t'_{ret}..

There is another way of defining the things that one "sees" at a given instant of time, as we discussed in Section 6.3.1, due to Panofsky and Phillips, and this is the concept of the information gathering, sphere collapsing toward the observer's position at the speed of light. Everything that you "see" at some time t, say $t = 0$, is gathered by an information collecting sphere centered on your position at $t = 0$, that has been collapsing toward that you at the speed of light since the beginning of time. For example, tonight you can go out and look up at the night sky and see the Andromeda galaxy and Saturn at the same time. Both of these are on your backward light cone when you do this--with the information about Andromeda being collected about two million years before you look up, and the information about Saturn collected 90 minutes before you look up, with all of this information arriving at your eye just as you look up, dutifully carried by the information gathering sphere which has been collapsing toward you at the speed of light for long time in the past, which passed through the Andromeda galaxy three million years before you looked up. .

The interesting thing about this view of things is that we can use it to get some feel for the strange way that Nature presents us with information collected in this way, in particular about the amount of charge on a moving particle. Let us forget about point charges for the moment and return to extended distributions of charge. We have seen that the potential due to a finite distribution of charge is given by equation (6.2.15)

(17.6.2)

where again we emphasize that everything is measured in the same coordinate frame, and where (for a spatially extended distribution of charge) $t'_{ret} = t - \left|\mathbf{r} - \mathbf{r}'\right|/c$. Now, suppose our charge is distributed on a a rod lying on the x-axis, carrying a charge per unit length λ with length L, with total charge $q_o = \lambda L$.[4] The rod is moving at some speed V along the x-axis, directly toward the observer. The observer sits on the x-axis at $x = 0$, calculating the potential there at $t = 0$ using the prescription in (17.6.2). The potential that the observer calculates according to this prescription (assuming the rod is far away) is

[4]In the rest frame of the rod, the charge per unit length of the rod is λ/γ and the length is γL, where **Error! Objects cannot be created from editing field codes.** , and therefore the product $\lambda L = q_o$ is a Lorentz invariant (the same in all co-moving frames).

$$\phi(0,0) = \frac{1}{4\pi\,\varepsilon_o}\int \lambda(x',t'_{ret})\frac{dx'}{|x'|} \cong \frac{1}{4\pi\,\varepsilon_o D}\int \lambda(x',t'_{ret})\,dx' \qquad (17.6.3)$$

How do we do the integral on the right in equation (42)? We know from the above discussion that evaluating $\lambda(x',t'_{ret})$ at t'_{ret} give us zero if no part of the rod is on the backward light cone of the observer, and λ if any part of the rod is on our backward light cone. This means that we should add up the charge per unit length times dx' for all the space-time points on the rod which lie on the backward light cone of the observer sitting at the origin at $t = 0$.

That is, the instruction given in the integral on the right of the above equation is to find the charge on the rod by taking its the *apparent* length as seen by the collapsing, information gathering sphere as it moves though the distributed charge at the speed of light, and multiplying that apparent length times λ. What will that give us?

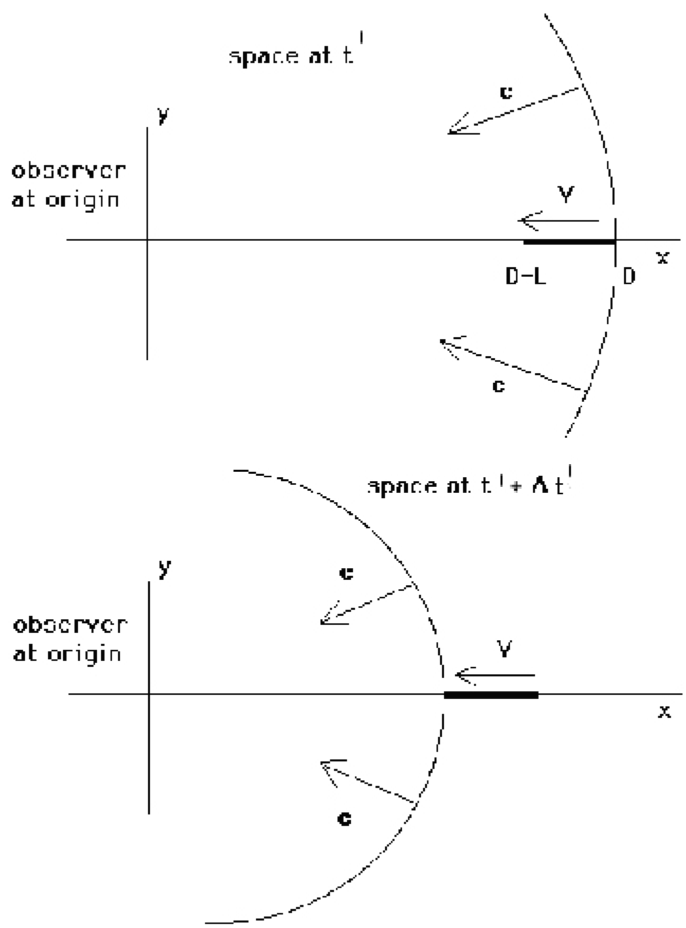


Figure 17-5: Sampling a line charge moving toward the observer.

Consider Figure 17-5. In the top diagram, we show space at the instant of time $t' < 0$. At this time, our information gathering sphere collapsing toward the origin at the speed of light has just encountered the right end of the rod, also moving toward the origin, but at speed $V < c$. In the bottom diagram, we show space at the time $t' + \Delta t'$. At this time, our information gathering sphere has just left the left end of the rod, heading toward the origin to deliver the information about how much charge it saw when it passed through the rod. The time $\Delta t'$ is thus the length of time that the rod is on the backward light cone of the observer. How long is $\Delta t'$? Well, in time $\Delta t'$, the rod moved toward the origin a distance $V\Delta t'$, so that the total distance between when the sphere first encountered the rod to when if last encountered the rod is $L + V\Delta t' = c\Delta t'$. Solving this equation for $\Delta t'$ gives $L/(c-V)$. The distance $\Delta x'$ on the backward light cone that the sphere was in contact with the rod ($\Delta x' = L + V\Delta t' = c\Delta t'$) is therefore given by $\Delta x' = c\Delta t' = L/(1-V/c)$. This is the length over which the integral in equation (17.6.3) will be non-zero, so that we have

$$\phi(0,0) \cong \frac{1}{4\pi\,\varepsilon_o\,D}\int \lambda(x', t'_{ret})dx^* = \frac{1}{4\pi\,\varepsilon_o\,D}\left[\frac{\lambda L}{(\,1-V/c\,)}\right] = \frac{1}{4\pi\,\varepsilon_o\,D}\left[\frac{q_o}{1-V/c}\right] \qquad (17.6.4)$$

Thus the potential we are instructed to calculate by equation (17.6.3) is that due to a larger charge than is actually on the rod, by a factor of $L/(\,1-V/c\,)$. This can be a huge factor if V is close enough to c. Furthermore, since the dimensions of the rod have disappeared from our final result in (17.6.4), this "enhancement" of charge will be true for a point charge moving directly toward the observer as well. In the more general case, the charge in (17.6.4) will be $q_o/(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})$, so that the way the information gathering sphere estimates charge can either be an overestimate (the case above) or an underestimate (for example, the rod moving directly away from the observer). The space-time diagram below in Figure 17-6 illustrates these two cases for a rod moving directly toward the origin along the x-axis at 0.8 c, and for another rod moving directly away from the origin along the negative y-axis at 0.8 c. Clearly the distance over which the backward light cone intersects these two moving rods varies dramatically, which leads to the dramatic variation in the Lienard-Weichert potentials with $(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})$.

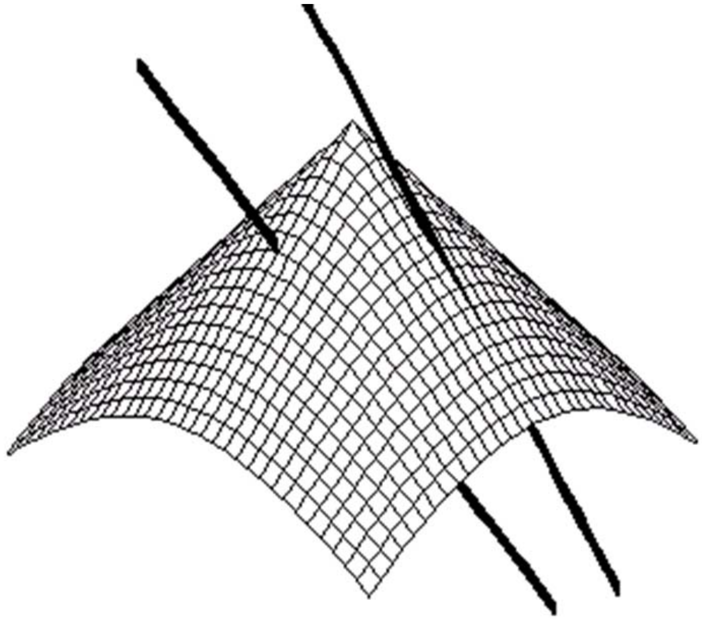


Figure 17-6: Two charges cutting the backward light cone of the observer.

You may think that this is really a strange way to decide how much charge is out there, and in fact Panofsky and Phillips offer the following analogy to show how strange it is. It is as if we were taking a census in a city to determine the total population in the following way. The census takers form a circle around the outskirts of the city, and then begin to converge toward the center of the city moving at some predetermined speed. They count the number of people in the following way. They observe the density of people around them, and then they multiply that density by the area they cover in a given time to get the number of people they have seen in that time.

Clearly this is a bad way to do a census, because if all the people are moving in toward the center of the city, the census takers will overestimate the number of people. For example, if there are only 1000 people in the city, but all 1000 gather in a circular ring that includes the census taking circle, and move along with the census takers at the same velocity, the census takers will count an enormous number of people, because the local density of people will stay the same for a really large distance. And vice versa--if people are moving out, they will undercount. This is no way to take a census, but it is precisely the way that Nature instructs us to figure out how much charge is out there. And this method really gives a lot of weight to charge moving toward us at near the speed of light. Hence the factor of $1/(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})$ in the Lienard-Wiechert potentials.

17.7 Synchrotron radiation

To show how dramatic this enhancement in the forward direction can be, let us suppose we have a charged particle moving in a circle or radius a in the *x*-*y* plane with angular frequency ω_o and speed $V_o = \omega_o a$ (see Figure 17-7). The particle's position $\mathbf{X}(t')$ is given by

$$\mathbf{X}(t') = a\,[\hat{\mathbf{x}}\cos(\omega_o\, t') + \hat{\mathbf{y}}\sin(\omega_o\, t')] \tag{17.7.1}$$

and the period $T = 2\pi / \omega_o$. The figures on the next page show the intensity of radiation as seen at a given instant of time by observers in the *x*-*y* plane at z = 0, for different values of the speed of the particle. The enormous change in the distribution as the speed approaches the speed of light is all due to the enhancement of the radiation in the forward direction. Essentially, at very relativistic speeds, the observers in the *x*-*y* plane only see radiation when the velocity vector of the particle is pointed right toward them (at the retarded time, of course). Otherwise, they see very little radiation, and if they are not in the *x*-*y* plane, they never see very much, because the velocity vector never points directly towards them.

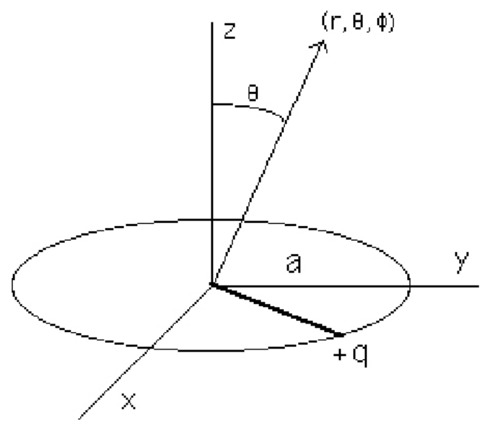


Figure 17-7: A charge moving in a circle in the *x*-*y* plane

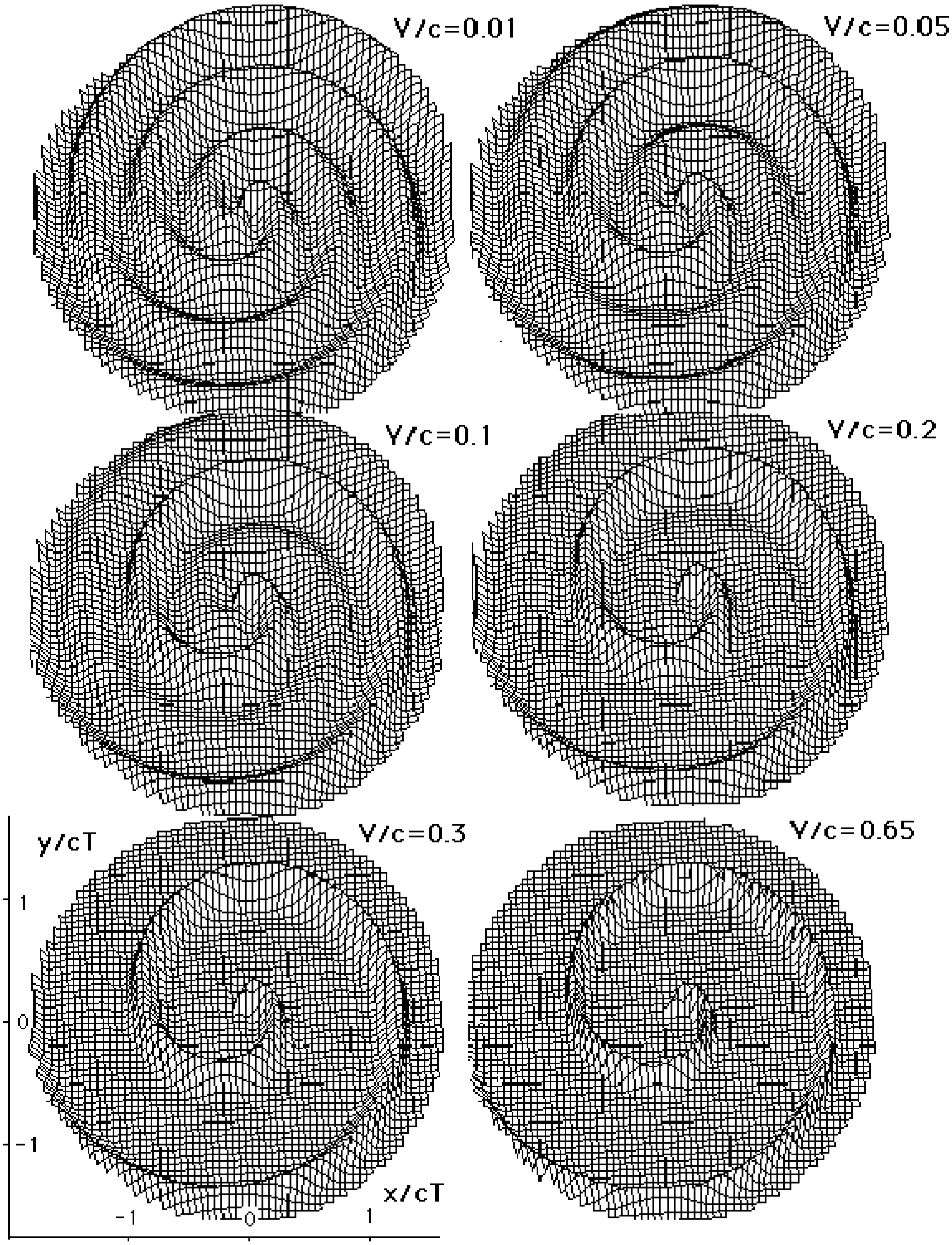


Figure 17-8: The spatial distribution of synchrotron radiation in the *x*-*y* plane.

It is interesting to look at the radiation electric fields as seen by a single observer as a function of retarded time. Suppose we have an observer far out on the positive x axis, at a distance $D >> a$. In this case, we have

$$t'_{ret} = t - \left| D\hat{\mathbf{x}} - \mathbf{X}(t'_{ret}) \right| / c = t - \frac{1}{c}\sqrt{(D - a\cos(\omega_o t'_{ret}))^2 + a^2 \sin(\omega_o t'_{ret})^2} \qquad (17.7.2)$$

and if we keep terms only to order *a/D*, this equation can be written

$$t'_{ret} = t - \frac{D}{c} + \frac{a}{c}\cos(\omega_o t'_{ret}) \qquad (17.7.3)$$

If we differentiate equation (17.7.3), we find that

$$dt = dt'_{ret}(1 + \frac{\omega_o a}{c}\sin(\omega_o t'_{ret})) \qquad (17.7.4)$$

or that

$$(1 - \hat{\mathbf{n}} \cdot \boldsymbol{\beta}) = (1 + \frac{\omega_o a}{c}\sin(\omega_o t'_{ret})) = (1 + \beta_o \sin(\omega_o t'_{ret})) \qquad (17.7.5)$$

This expression for $(1 - \hat{\mathbf{n}} \cdot \boldsymbol{\beta})$ is smallest when $\omega_o t'_{ret}$ is $3\pi/2$, that is, when the velocity vector is pointed along the positive x-axis, directly toward our observer (see Figure 17-7).

Now, what electric fields does this observer along the x-axis see? Well, we have equation (17.4.1) and we deduce that the radiation field has only a *y* component. If we plot E_y as a function of t'_{ret}, we get curves that look like those in Figure 17-9, for various values of β_o . We have adjusted the normalization at each value of β_o so that the electric field is a maximizes at unity. In fact, there is a tremendous enhancement of the electric field for the same acceleration as β_o approaches unity.

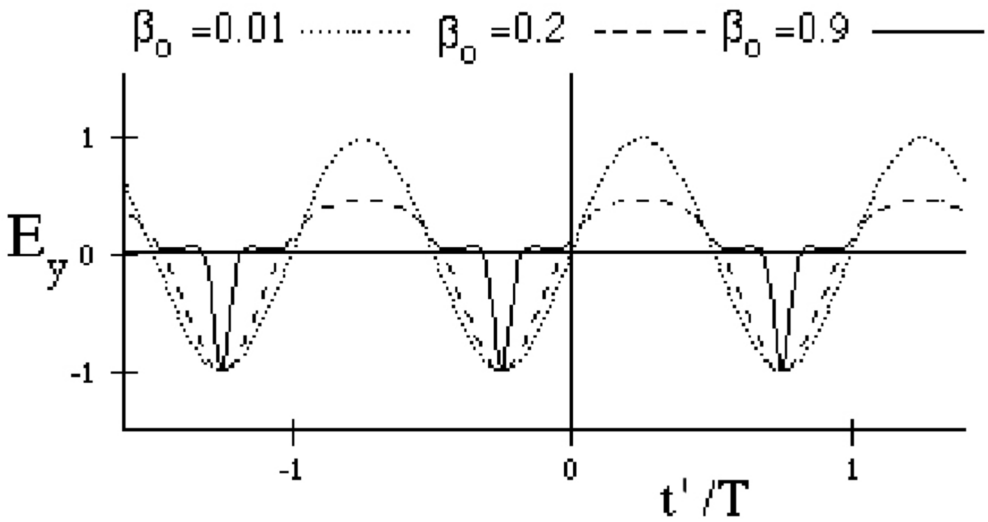


Figure 17-9: E_y as a function of time for various values of V_o/c.

One final point about the amount of radiation this accelerating charge radiates. The four acceleration defined by (16.3.4) has the following Lorentz invariant length

$$\Xi_\mu \Xi^\mu = \gamma_u{}^4 \left[-\left(\gamma_u{}^2 \left(\frac{\mathbf{u}\cdot\mathbf{a}}{c} \right) \right)^2 + \left(\mathbf{a}(t) + \gamma_u{}^2 \mathbf{u} \left(\frac{\mathbf{u}\cdot\mathbf{a}}{c^2} \right) \right)^2 \right] = \gamma_u{}^6 \left[|\mathbf{a}|^2 - \left| \frac{\mathbf{a}\,\mathbf{x}\,\mathbf{u}}{c} \right|^2 \right] \qquad (17.7.6)$$

But we know that in the instantaneous rest frame of the charge, the form for Ξ^μ is

$$\Xi^\mu \Big|_{\text{rest frame}} = \gamma_u{}^2 \left(\begin{array}{c} \gamma_u{}^2 \left(\dfrac{\mathbf{u}\cdot\mathbf{a}}{c} \right) \\ \mathbf{a}(t) + \gamma_u{}^2 \mathbf{u} \left(\dfrac{\mathbf{u}\cdot\mathbf{a}}{c^2} \right) \end{array} \right) \Bigg|_{\text{rest frame}} = \left(\begin{array}{c} 0 \\ \mathbf{a}_{\text{rest frame}} \end{array} \right) \qquad (17.7.7)$$

Equations (17.7.6) and (17.7.7) tell us that if we measure the three acceleration and three velocity in any frame, we can compute the acceleration in the instantaneous rest frame of the particle by computing the following quantity

$$\mathbf{a}^2{}_{rest\ frame} = \gamma_u{}^6 \left[|\mathbf{a}|^2 - \left| \frac{\mathbf{a}\,\mathbf{x}\,\mathbf{u}}{c} \right|^2 \right] \qquad (17.7.8)$$

This looks familiar. If you look back at equation (17.4.8), we had

$$\frac{dW_{rad}}{dt^*} = \frac{1}{4\pi\varepsilon_o} \frac{2q^2}{3c} \gamma^6 \left[|\dot{\boldsymbol{\beta}}|^2 - |\boldsymbol{\beta}\ \mathbf{x}\ \dot{\boldsymbol{\beta}}|^2 \right] \qquad (17.7.9)$$

but using (17.7.8), this is just

$$\frac{dW_{rad}}{dt'} = \frac{1}{4\pi\,\varepsilon_o}\frac{2\,q^2}{3c^3}\gamma^6\left[\left|\mathbf{v}\right|^2 - \left|\frac{\mathbf{v}\times\dot{\mathbf{v}}}{c}\right|^2\right] = \frac{1}{4\pi\,\varepsilon_o}\frac{2\,q^2}{3c^3}\mathbf{a}^2_{\;rest\;frame} \tag{17.7.10}$$

17.8 The Radiation Reaction Force

We now turn to a self force term that we have so far neglected. An accelerating charge must feel a "back reaction force" due to its own self-fields. This form of this force is generally written as

$$\mathbf{F}_{\text{radiation reaction}} = \frac{1}{4\pi\varepsilon_o}\frac{2}{3}\frac{q^2}{c^3}\frac{d^2\mathbf{V}}{dt^2} \tag{17.8.1}$$

This is the force that must be exerted on the charged object due to the fact that it is radiating and losing energy irreversibly to infinity due to that radiation. Something must be providing energy to power the radiation, and it is either coming from the kinetic energy of the charge, or if the charge is not losing kinetic energy, it must be coming from an external agent which is doing work at a rate sufficient to account for the radiation. In any case there must be an electric field at the charge due to the fact that it is radiating, and the force associated with that electric field has to have the form given in (17.8.1).

The standard way to get a form for this radiation reaction force for a point charge is to say that whatever the form of this force $\mathbf{F}_{\text{radiation reaction}}\,\mathbf{F}_{rr}$, if we are going to maintain constant kinetic energy of the charge, the rate at which *we* do work against it to offset its effects is given by $-\mathbf{F}_{\text{radiation reaction}}\cdot\mathbf{v}$, where **v** is the velocity of the charge. This energy must eventually end up as energy radiated away, so that if we compute the total work we do over all time, conservation of energy demands that we must have

$$W_{us} = -\int \mathbf{F}_{rr}\cdot\mathbf{v}\;\,dt = \int\frac{dW_{rad}}{dt}dt = \int\frac{1}{4\pi\varepsilon_o}\frac{2}{3}\frac{q^2a^2}{c^3}dt \tag{17.8.2}$$

In arriving at the last term in (17.8.2), we have used the standard form for the rate at which energy is radiated into electric dipole radiation by a non-relativistic charge, and we are integrating over all time in equation (17.8.2). If we assume that the particle is at rest at the endpoints of our integration, we can integrate the right side of equation (17.8.2) by parts to obtain the form we want for $\mathbf{F}_{\text{radiation reaction}}$

$$\int\frac{1}{4\pi\varepsilon_o}\frac{2}{3}\frac{q^2a^2}{c^3}dt = \int\frac{1}{4\pi\varepsilon_o}\frac{2}{3}\frac{q^2}{c^3}\dot{\mathbf{v}}\cdot\dot{\mathbf{v}}\,dt = -\int\frac{1}{4\pi\varepsilon_o}\frac{2}{3}\frac{q^2}{c^3}\ddot{\mathbf{v}}\cdot\mathbf{v}\,dt \tag{17.8.3}$$

If we compare equation (17.8.2) and (17.8.3), it is natural to conclude that

$$\mathbf{F}_{\text{radiation reaction}} = \frac{1}{4\pi\varepsilon_o}\frac{2}{3}\frac{q^2}{c^3}\ddot{\mathbf{v}} = \frac{2}{3}m_e\frac{r_e}{c}\ddot{\mathbf{v}} = \frac{2}{3}m_e\tau_e\ \ddot{\mathbf{v}} \tag{17.8.4}$$

where

$$\tau_e = \frac{r_e}{c} \tag{17.8.5}$$

is the speed of light transit time across the classical electron raduis. This is the form we have in (17.8.1).

It is important to emphasize that the radiation reaction force represents an irreversible loss of energy to infinity. We never get this energy back, it disappears forever from the system.

18 Basic Electrostatics

18.1 Learning Objectives

We first motivate what we are going to do in the next four or five sections. We then go back to the origins of electromagnetism—electrostatics, and spend some time going through the classic aspects of this subject, including the energy we put into assembling a configuration of charges.

18.2 Where are we going?

In Section 1.3.1, I enumerated what I consider to be the profound part of classical electromagnetism, which I repeat here.

1) The existence of fields which carry energy and momentum, and the ways in which they mediate the interactions of material objects.

2) The nature of light and the radiation process.

3) The explicit prescription for the way that space and time transform which is contained in Maxwell's equations.

We have finished with (2) and (3) above, and we have touched on various aspects of (1). We now focus our attention on (1). In particular we will be looking at the electromagnetic interactions of particles and fields in the "near zone", where we neglect the effects of radiative losses, and look at the reversible exchange of energy, momentum, and angular momentum between charged particles and fields, and how that proceeds.

As a preview of the sorts of things we want to explore, consider the application showing the interaction of charged particles, as shown in Figure 18-1 and Figure 18-2.