8.07 Class Notes Fall 2010

Most of the animations and visualizations referred to herein are at
://web.mit.edu/viz/EM/

Created on a PC running XP with Word 2003 and MathType 6.
Comments and questions to @mit.edu

Table of Contents
(*LIVE LINKS*)

Figure Captions

1 Fields and their Depiction

1.1 Learning objectives

The learning objectives for this section are first to get an overview of electromagnetism and how it changed the way we view the world. I then discuss the way I present the material in this course, and how it differs from the traditional treatments. I will tell you why I choose to teach topics in this order, and why I think this organization will help you get a deeper feel for electromagnetic theory, rather than you becoming lost in the mathematics.

1.2 Maxwell's electromagnetism as a fundamental advance

Classical electromagnetic field theory emerged in more or less complete form in 1873 in James Clerk Maxwell's *A Treatise on Electricity and Magnetism*. Maxwell's treatise had an impact for electromagnetism similar to that of Newton's *Principia* for classical mechanics. It not only provided the mathematical tools for the investigation of electromagnetic theory, but it also altered the basic intellectual framework of physics, in that it finally displaced *action at a distance* and substituted for it the concept of *fields.*

What is action at a distance? It is a world view in which the interaction between two material objects requires no mechanism other than the objects themselves and the empty space between them. That is, two objects exert a force on each other simply because they are present. Any mutual force between them (for example, gravitational attraction or electric repulsion) is instantaneously transmitted from one object to the other through empty space. There is no need to take into account any method or agent of transmission of that force, or any finite speed for the propagation of that agent of transmission. This is known as *action at a distance* because objects exert forces on one another (*action*) with nothing but empty space (*distance*) between them. No other agent or mechanism is needed.

Many natural philosophers, including Newton (1693)[1], criticized the action at a distance model because in our everyday experience, forces are exerted by one object on another only when the objects are in direct contact. In the field theory view, this is always true in some sense. That is, objects that are not in direct contact (objects separated by apparently empty space) exert a force on one another through the presence of an intervening medium or mechanism existing in the space between the objects. The force between the two objects is transmitted by direct *contact* from the first object to an intervening mechanism immediately surrounding that object, and then from one element of space to a neighboring element, in a continuous manner, until it is transmitted to the

[1] "That Gravity should be innate, inherent and essential to Matter, so that one body may act upon another at a Distance thro' a *Vacuum*, without the Mediation of anything else, by and through which their Action and Force may be conveyed from one to another, is to me so great an Absurdity, that I believe no Man who has in philosophical Matters a competent Faculty of thinking, can ever fall into it."

region of space contiguous to the second object, and thus ultimately to the second object itself.

Thus, although the two objects are not in direct contact with one another, they *are* in direct contact with a medium or mechanism that exists between them. The force between the objects is transmitted (at a finite speed) by stresses induced in the intervening space by the presence of the objects. The field theory view in classical electromagnetism thus avoids the concept of action at a distance and replaces it by the concept of *action by continuous contact*. The contact is provided by a stress, or field, induced in the space between the objects by their presence.

This is the essence of field theory, and is the foundation of all modern approaches to understanding the world around us. Field theory has of course evolved far beyond these beginnings. In the modern view, every aspect of reality is due to *quantized* fields:

> *In its mature form, the idea of quantum field theory is that quantum fields are the basic ingredients of the universe, and the particles are just bundles of energy and momentum of the fields.*
>
> Weinberg, 1999

Our task here is to understand electromagnetism, before quantization, with emphasis on the energy and momentum carried by fields. This is of interest in itself, and will also give us insight, by analogy, into aspects of "matter fields", and how they can carry energy and momentum in their quantized, particle-like realizations.

1.3 Why this course is different

1.3.1 The profound parts of E&M first

The standard way to approach this subject is to present the various topics in electromagnetism in the *historical* order in which they were developed—e.g. electrostatics first, then magnetostatics, then Faraday's Law, and finally the displacement current and radiation, followed by special relativity and the manifestly covariant form of Maxwell's equations. Although there is much to recommend this approach, and perhaps it is the best one to follow in a course that spans two or more semesters, I do not follow it here.

The reasons are as follows. At MIT, Electromagnetism II, 8.07, is a one term course on a semester system, and thus the course is approximately 12 weeks or 37 one hour classes long. I have taught this course many times at MIT, and invariably with the traditional organization I get to the most interesting and profound material near the end of that 12 weeks , when the students (and myself) are exhausted. In contrast, I traditionally have spent a lot of time at the beginning on material which is mathematically difficult but not profound, that is electrostatics and magnetostatics. I count the most profound aspects of classical electromagnetism to be as follows:

The existence of fields which carry energy and momentum

How these fields mediate the interactions of material objects, especially the fact that the shape *of the fields is predictive of the stresses they transmit.*

The nature of light and of the radiation processes by which it is created.

Maxwell's equations contain the way that space and time transform.

In this course I propose to address the above, more profound, aspects of electromagnetism first. I begin with general solution for the electromagnetic fields given known sources of charge and current density. I will apply these solutions to many different cases, the first of which will be to consider in a relatively brief treatment the static solutions far from sources which do not vary in time. Then I will immediately move on to find the fields far from a spatially localized set of sources which slowly vary in time. Then I discuss the conservation of energy and momentum at length. I then address special relativity. At the end of the course I return to statics and also consider the effects of the presence of material media.

1.3.2 The easy E&M and the hard E&M

Another reason that I depart from the traditional sequence is that there is an "easy" electromagnetism and a "hard" electromagnetism. The first occurs when the behavior of the sources of electromagnetic fields, that is, charges and currents, is given, and that behavior cannot be influenced by the fields that they produce. The second occurs when the behavior of the sources of the fields can be affected by the fields that they produce.

It is in this second situation that electromagnetism becomes difficult, and in many situations intractable--when the fields that are produced by sources can affect the sources that produce them. When we are dealing with linear dielectric or magnetic material media, this situation obtains, but because of the linearity there are straightforward ways to deal with it. But in other contexts there is no good analytic approach. For example, in the traditional approach to the subject the really hard E&M appears almost immediately, in boundary value problems in electrostatics, and much effort is expended in investigating the details of solutions to these difficult kinds of problems, which is in parallel with the historical development of the field.

However, it turns out that looking at the easy part of E&M first is more than enough to show you the nature of fields, the energy and momentum that they carry, the nature of radiation, and the way in which space and time transform. For that reason, I prefer to separate the treatment of electromagnetism into the easy part and the hard part, doing the easy part first. This allows me to spend more time addressing the profound issues, leaving the less profound (and frequently more mathematically difficult) issues until later. The question of the effects of material media can also be grouped with the "easy" electromagnetism, but there are a number of (not so profound) complexities in this, and I therefore leave this subject to the end of my treatment.

1.3.3 Energy and momentum in fields

One of the many amazing things about classical electromagnetic fields is that they carry energy, momentum, and angular momentum just as "ordinary matter" does, and there is a constant interchange of these quantities between their mechanical forms and their electromagnetic field forms. Although all texts in electromagnetism make this point, and derive the appropriate conservation laws, actual examples showing the interchange are rare. In this text I put a lot of emphasis on the processes by which energy and momentum are created in fields and the manner in which that energy and momentum flows around the system, to and from the fields and particles, thereby mediating the interaction of the particles.

For example, consider the electromagnetism of the head-on collision of two charged particles of equal mass and charge. In this process, energy is stored in the field and then retrieved, electromagnetic momentum flux is created at the location of one charge and momentum flows via the electromagnetic field to the other charge, in a dazzling array of interaction between matter and fields. But this problem is almost never discussed in these terms. I will try in all circumstances in this text to describe how the field mediates the interaction of material objects by taking up energy and momentum from them and by transferring this energy and momentum from one particle to the other. This is a different emphasis than the traditional approach, and one which illustrates more clearly the essence and the importance of fields.

1.3.4 Animations and visualizations

> *In order therefore to appreciate the requirements of the science [of electromagnetism], the student must make himself familiar with a considerable body of most intricate mathematics, the mere retention of which in the memory materially interferes with further progress.*
>
> James Clerk Maxwell (1855)

I will spend some time on vector fields and examples of vector fields, and the methods we will use to visualize these fields in the course. The mathematics I delve into is fierce, and as is well known, the level of the mathematics obscures the physical reality that the equations represent. The quote from Maxwell above is from one of his first papers on the subject. I try to offset the tyranny of the mathematics by using many visual depictions of the electromagnetic field, both stills and movies, and also interactive visualizations, as appropriate to the topic at hand.

1.4 Reference texts

There are many excellent electromagnetism texts available, the most popular being the text by Griffiths (1999), which is the course textbook. I will also refer to a number of other classic texts in this area, most notably the various editions of Jackson (1975), but also to texts by Panofsky and Phillips (1962), Hertz (1893), Jefimenko (1989) and others. If you are interested in the history of electromagnetism, the book I would recommend if you read only one is *The Maxwellians* by Hunt (2005).

1.5 Representations of vector fields

1.5.1 The vector field representation of a vector field

A field is a function that has a value at every point in space. A scalar field is a field for which there is a single number associated with every point in space. A vector field is a field in which there is a vector associated with every point in space—that is, three numbers instead of only the single number for the scalar field. An example of a vector field is the velocity of the Earth's atmosphere, that is, the wind velocity.

Figure 1-1-1 is an example of a "vector field" representation of a field. We show the charges that would produce this field if it were an electric field due to two point charges, one positive (the orange charge) and one negative (the blue charge). We will always use this color scheme to represent positive and negative charges.

In the vector field representation, we put arrows representing the field direction on a rectangular grid. The direction of the arrow at a given location represents the direction of the vector field at that point. In many cases, we also make the length of the vector proportional to the magnitude of the vector field at that point. But we also may show only the direction with the vectors (that is make all vectors the same length), and color-code the arrows according to the magnitude of the vector. Or we may not give any information about the magnitude of the field at all, but just use the arrows on the grid to indicate the direction of the field at that point.

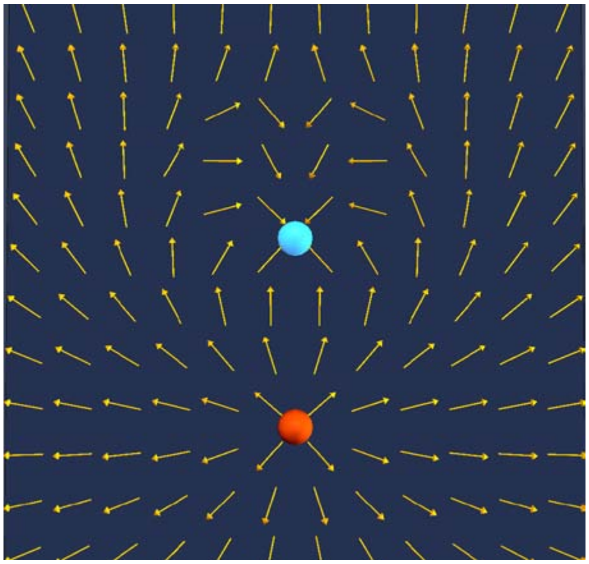


Figure 1-1-1: A vector representation of the field of two point charges.

Figure 1-1-1 is an example of the latter situation. That is, we use the arrows on the vector field grid to simply indicate the direction of the field, with no indication of the magnitude of the field, either by the length of the arrows or their color. Note that the arrows point away from the positive charge (the positive charge is a "source" for electric field) and towards the negative charge (the negative charge is a "sink" for electric field). In this case the magnitude of the positive charge is five times the magnitude of the negative charge.

1.5.2 The field line representation of a vector field

There are other ways to represent a vector field. One of the most common is to draw field lines. To draw a field line, start out at any point in space and move a very short distance in the direction of the local vector field, drawing a line as you do so. After that short distance, stop, find the new direction of the local vector field at the point where you stopped, and begin moving again in that new direction. Continue this process indefinitely. Thereby you construct a line in space that is everywhere tangent to the local vector field. If you do this for different starting points, you can draw a set of field lines that give a good representation of the properties of the vector field. Figure 1-1-2 below is an example of a field line representation for the same two charges we used in Figure 1-1-1.

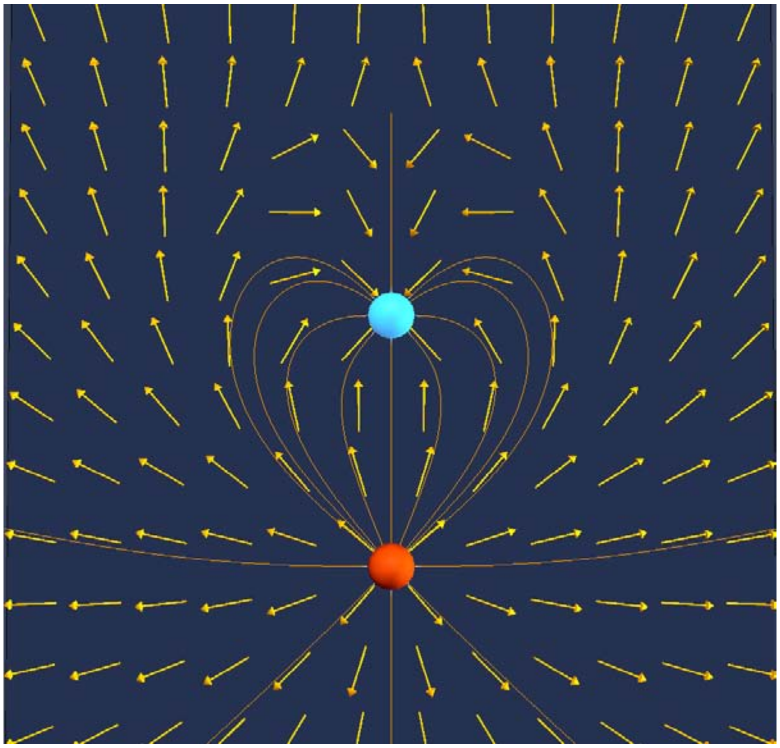


Figure 1-1-2: A vector field and field line representation of the same field.

The mathematics of constructing field lines follows directly from the description above. Let us parameterize a line in space by the arc length along the line, *s*, where the line goes through a point in space $\mathbf{r}_o = (x_o, y_o, z_o)$ and we measure arc length along the field line from the point $\mathbf{r}_0$. A field line going through $\mathbf{r}_0$ of the vector field **F** is a line in space $\mathbf{r}(s) = (x(s), y(s), z(s))$ that satisfies at every point along the line the equations in Cartesian coordinates

$$\frac{dx(s)}{ds} = \frac{F_x}{F} \qquad \frac{dy(s)}{ds} = \frac{F_y}{F} \qquad \frac{dz(s)}{ds} = \frac{F_z}{F} \tag{1.5.1}$$

In some situations we can solve (1.5.1) for the equation describing the field lines analytically, but we can always generate them numerically using this definition.

1.5.3 Line integral convolution representations

The final representation of vector fields is the line integral convolution representation (LIC). The advantage of this representation lies in its spatial resolution. The use of field lines has the disadvantage that small scale structure in the field can be missed depending on the choice for the spatial distribution of the field lines. The vector field grid representation has a similar disadvantage in that the associated icons limit the spatial resolution because of the size of the icons and because of the spacing between icons needed for clarity. These two factors limit the usefulness of the vector field representation in showing small scale structure in the field.

In contrast, the LIC method of Cabral and Leedom (1993) avoids both of these problems by the use of a texture pattern to indicate the spatial structure of the field at a resolution close to that of the display. Figure is a LIC representation of the electric field for the same charges as in the earlier Figures. The local field direction is in the direction in which the texture pattern in this figure is correlated. The variation in color over the figure has no physical meaning; the color is used simply to give visual information about the local field direction. This LIC representation gives by far the most information about the spatial structure of the field. By its nature it cannot indicate the direction of the field: the texture pattern indicates either the field direction or the direction 180 degrees from the field direction.

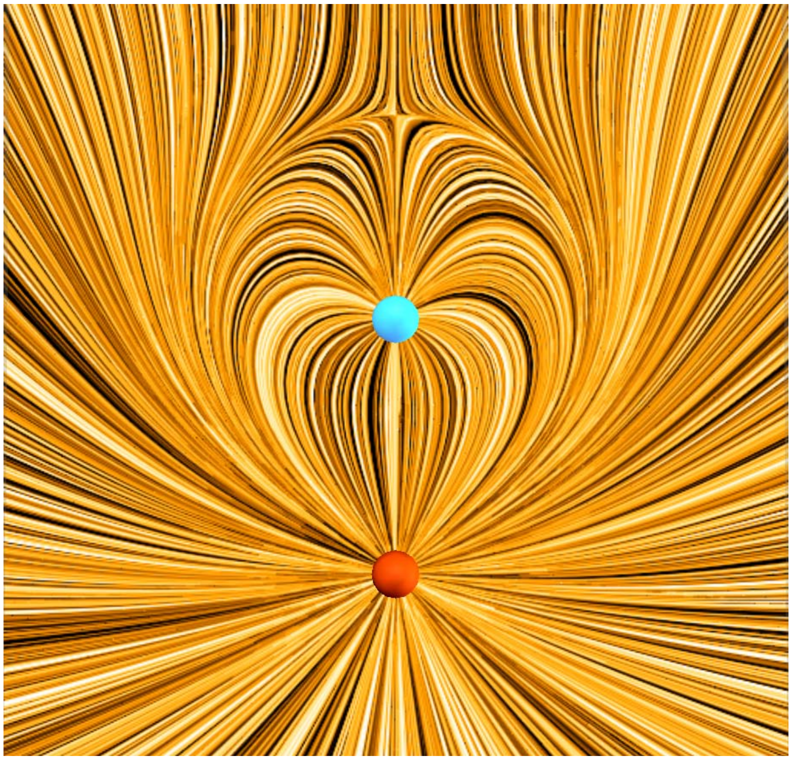


Figure 1-3: A LIC representation of the same field as in earlier figures

The complexity of vector field topologies can be amazing. For example, in Figure 1-4 we show the LIC representation of the vector field

$$\mathbf{F}(\mathbf{r}) = \sin\left(y^2\right)\hat{\mathbf{x}} + \cos(x^2)\hat{\mathbf{y}} \tag{1.5.2}$$

Although this is a simple analytic form for a vector field, the visual representation of the field is complex. This vector field has zero curl, as is apparent from looking at the representation below.

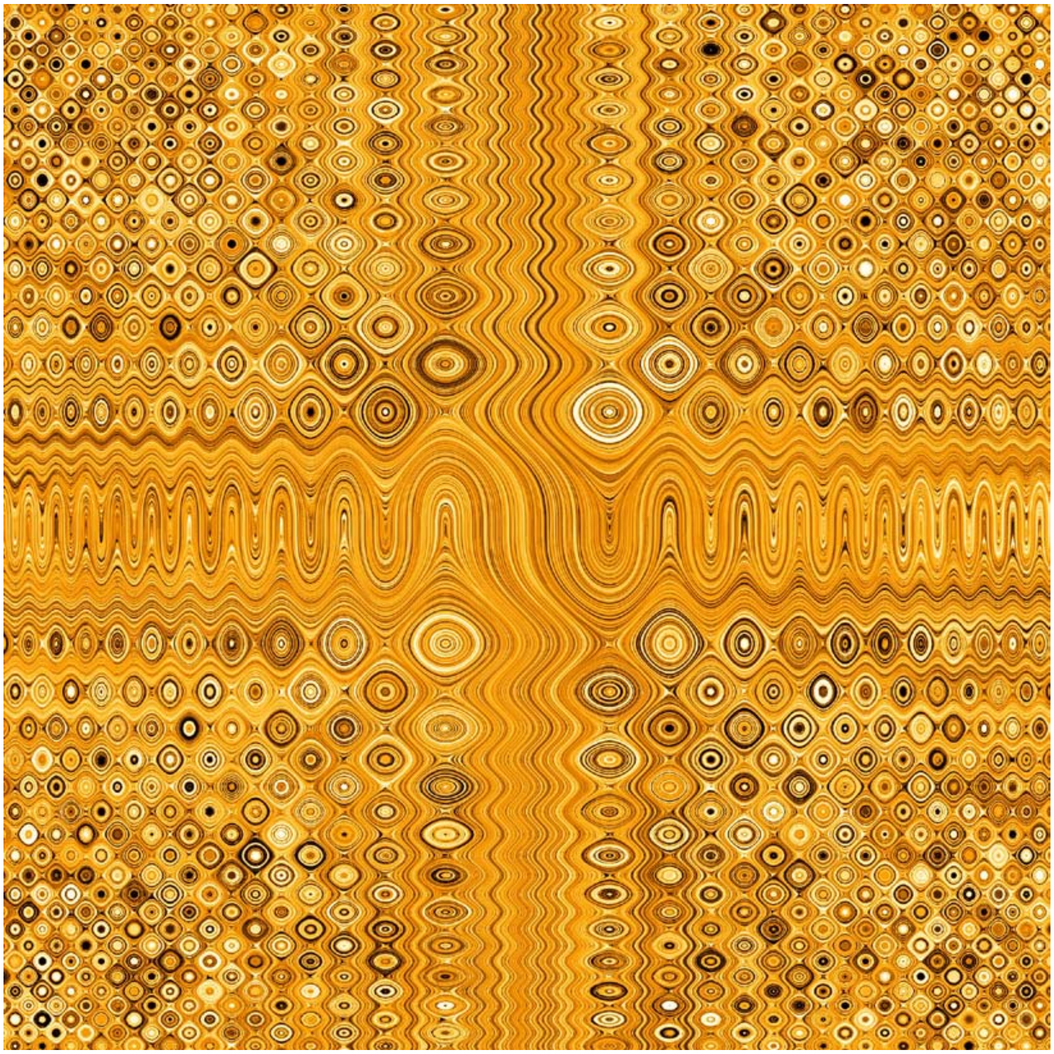


Figure 1-4: A LIC representation of a vector field with no sources

2 Conservations Laws for Scalar and Vector Fields

2.1 Learning objectives

The learning objectives for this handout are to get a feel for the mathematics of vector fields, using fluid flow as an example. We also introduce various nomenclatures that were first introduced in the study of fluids (e.g., flux, sources, sinks) and subsequently taken over to electromagnetism. We learn here what mathematical form we expect to see for conservation of both scalar (e.g. mass) and vector (e.g. momentum) quantities. We then introduce the properties of an important "improper" function, the Heaviside-Dirac delta function. Finally, we discuss briefly the topic of complete sets of functions.

2.2 Conservation laws for scalar quantites in integral and differential form

2.2.1 Density and creation rate of a conserved scalar quantity

Much of the terminology we use in electromagnetism comes from the theory of fluids, and fluid flow represents a tangible example of a vector field which is easily visualized. We thus briefly review the properties of fluid flow. We particularly focus on conservations laws, since we will see many such laws in electromagnetism, and understanding the meaning of these laws is a central part of understanding the physics.

Consider conservations laws for scalar quantities first. Suppose the mass per unit volume of the fluid at $\left(\mathbf{r},t\right)$ is given by $\rho_{mass}(\mathbf{r},t)$ and the velocity of a fluid element at $\left(\mathbf{r},t\right)$ is given by $\mathbf{v}(\mathbf{r},t)$. Suppose also that mass is being created at $\left(\mathbf{r},t\right)$ at a rate given by $s_{mass}(\mathbf{r},t)$ (units of *s* are mass per unit volume per unit time). Consider an arbitrary closed surface *S* containing a volume *V*, as shown in **Error! Reference source not found.**. We assume that the surface and the enclosed volume are fixed in space and time (e.g. the surface of the volume is not moving). At time *t*, the amount of mass inside of the closed surface is given by the volume integral of the mass density, and the rate at which matter is being created inside the volume is given by the volume integral of the mass creation rate per unit volume. On physical grounds, it is obvious that the time rate of change of the amount of matter inside of the volume is given by the rate at which it is being created inside the volume minus the rate at which matter is flowing out through the stationary surface of the volume. That is,

$$\frac{d}{dt}\int_V \rho_{mass}\left(\mathbf{r},t\right)d^3x = \int_V s_{mass}\left(\mathbf{r},t\right)d^3x - \text{rate of mass flow out of volume} \quad (2.2.1)$$

2.2.2 The flux and flux density of a conserved scalar quantity

What is the rate at which mass is flowing out of the volume? Consider a infinitesimal surface element on the surface $\hat{\mathbf{n}}\,da$, where $\hat{\mathbf{n}}\,da$ is the local normal to the surface, defined so that it points away from the volume of interest (in this case outward). The rate at which mass is flowing through this surface element at time t is $\rho_{mass}\mathbf{v}\cdot\hat{\mathbf{n}}\,da$. To see this, imagine that you are an observer sitting on da and you measure the total amount of matter which flows across da in a time dt. This amount is the volume of the matter that will cross da in time dt, $\left[\mathbf{v}\cdot\hat{\mathbf{n}}\,da\,dt\right]$, times the mass density, that is, $\rho_{mass}\left[\mathbf{v}\cdot\hat{\mathbf{n}}\,da\,dt\right]$. Note that the dot product of $\mathbf{v}$ and $\hat{\mathbf{n}}$ in this expression is very important. If there is no component of the flow velocity along $\hat{\mathbf{n}}$, there is no matter flow across da.

To find the *rate* of matter flow we simply divide this amount of matter in time dt by the infinitesimal time dt, giving us $\rho_{mass}\mathbf{v}\cdot\hat{\mathbf{n}}\,da$. If we want the total rate at which mass if flowing through the entire surface S we integrate this quantity over the entire surface. Then (2.2.1) becomes

$$\frac{d}{dt}\int_V \rho_{mass}\left(\mathbf{r},t\right)d^3x = \int_V s_{mass}\left(\mathbf{r},t\right)d^3x - \int_S \left[\rho_{mass}\mathbf{v}\right]\cdot\hat{\mathbf{n}}\,da \qquad (2.2.2)$$

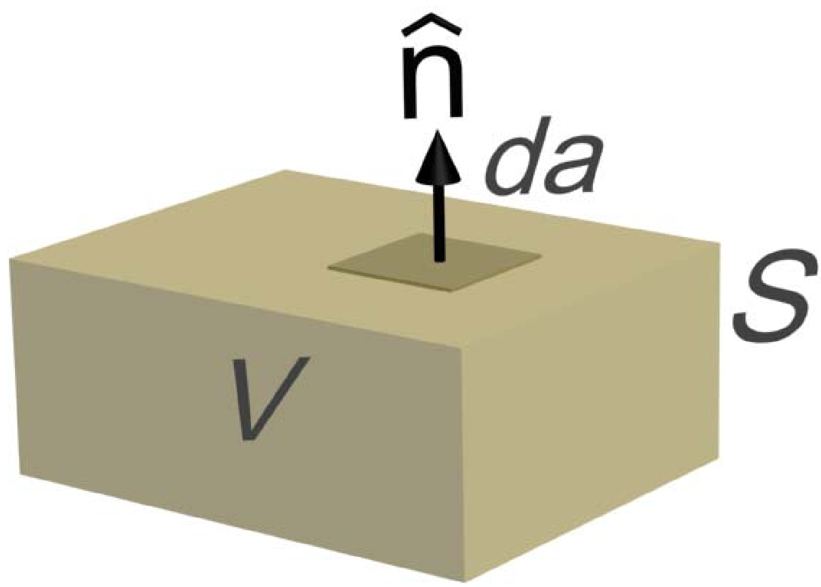


Figure 2-1: A closed surface *S* with enclosed volume *V*

We call $\int_S \left[\rho_{mass}\mathbf{v}\right]\cdot\hat{\mathbf{n}}\,da$ the flux (flow) of matter through the surface S. This quantity is a scalar. A point of frequent confusion is that the vector quantity $\rho_{mass}\mathbf{v}$ is sometimes also called the flux. In these notes we will always call quantities like $\rho_{mass}\mathbf{v}$ the flux density (a vector), and reserve the use of flux to mean the scalar we get from integrating the flux density over a specific surface.

The more standard way of writing (2.2.2) is to bring the flux to the left and side, that is,

$$\frac{d}{dt}\int_V \rho_{mass}(\mathbf{r},t)\,d^3x + \int_S [\rho_{mass}\mathbf{v}]\cdot\hat{\mathbf{n}}\,da = \int_V s_{mass}(\mathbf{r},t)\,d^3x \tag{2.2.3}$$

2.2.3 Gauss's Theorem and the differential form

We now want to write (2.2.3) in differential form. To do this we invoke one of the fundamental theorems of vector calculus, the divergence theorem. This theorem states that if **F** is a reasonably behaved vector function, then the surface integral of **F** over a closed surface is related to the volume integral over the volume that surface encloses of the divergence of **F**

$$\int_S \mathbf{F}\cdot\hat{\mathbf{n}}\,da = \int_V \nabla\cdot\mathbf{F}\,d^3x \tag{2.2.4}$$

where in Cartesian coordinates

$$\nabla\cdot\mathbf{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \tag{2.2.5}$$

If we use (2.2.5) in (2.2.3), and use the fact that our surface *S* is arbitrary, we see that we must have for the conservation of mass in differential form the equation

$$\frac{\partial \rho_{mass}}{\partial t} + \nabla\cdot[\rho_{mass}\mathbf{v}] = s \tag{2.2.6}$$

2.3 Conservation laws for vector quantities in integral and differential form

Above we were talking about the conservation of a scalar quantity, e.g. the scalar field $\rho_{mass}(\mathbf{r},t)$. However many of our conserved quantities in electromagnetism are vector quantities, for example electromagnetic momentum. To see how we handle conservation laws involving vector quantities, let us first consider the fluid context, specifically the conservation of the momentum density of the fluid. The momentum density of the fluid is a vector, and is given by the mass density of the fluid times the vector velocity of the fluid element, that is $\rho_{mass}\mathbf{v}$. If the momentum of fluid is a conserved quantity, we should have a statement similar to (2.2.1), that is

$$\frac{d}{dt}\int_V \rho_{mass}\mathbf{v}\,d^3x = \int_V [\text{creation rate of momentum}]\,d^3x - \text{rate of momentum flow out of volume} \tag{2.3.1}$$

Note now that our creation rate of momentum is a vector quantity, as is the rate at which momentum flows out of the volume through *da*. In analogy with the mass flowing out of the volume through $\hat{\mathbf{n}}\,da$, the momentum flowing out of the volume as carried by the flow is $\left[\rho_{mass}\mathbf{v}\right]\mathbf{v}\cdot\hat{\mathbf{n}}\,da$. That is, we take the density of momentum, $\left[\rho_{mass}\mathbf{v}\right]$, multiply it by the volume of the fluid which flows across *da* in time *dt*, $\mathbf{v}\cdot\hat{\mathbf{n}}\,da\,dt$, and divide by *dt* to get the rate. Thus (2.3.1) becomes (bringing the flux of momentum to the left hand side),

$$\frac{d}{dt}\int_V \rho_{mass}\mathbf{v}\,d^3x + \int_S \left[\rho_{mass}\mathbf{v}\right]\mathbf{v}\cdot\hat{\mathbf{n}}\,da = \int_V \left[\text{creation rate of momentum}\right]\,d^3x \qquad (2.3.2)$$

So the quantity $\int_S \left[\rho_{mass}\mathbf{v}\right]\mathbf{v}\cdot\hat{\mathbf{n}}\,da$ is the flux of momentum through the surface S, and the flux density of momentum is a second rank tensor, $\rho_{mass}\mathbf{v}\,\mathbf{v}$. This makes sense because the flux of a scalar quantity like mass is a vector, e.g. $\rho_{mass}\mathbf{v}$, and therefore the flux of a vector quantity like momentum must be more complicated than a vector, just as a vector is a more complicated object than a scalar. We only mention second rank tensors in passing here. Later we will consider them in detail.

In differential form, the conservation for momentum is written as

$$\frac{\partial}{\partial t}\left[\rho_{mass}\mathbf{v}\right] + \nabla\cdot\left[\rho_{mass}\mathbf{v}\,\mathbf{v}\right] = \left[\text{volume creation rate of momentum}\right] \qquad (2.3.3)$$

3 The Dirac delta function and complete othorgonal sets of fucntions

3.1 Basic definition

One of the functions we will find indispensable in proving various vector theorems in electromagnetism is the Dirac[2] delta function. In one dimension, the delta function is defined by

$$\delta(x) = 0 \qquad x \neq 0 \qquad (3.1.1)$$

$$\int_{-\infty}^{\infty} \delta(x)dx = 1 \qquad (3.1.2)$$

From (3.1.2) we see that the one-dimensional delta function must have the dimensions of one over the dimensions of its argument. To get an intuitive feel for this “improper”

---

[2] Although this function is usually attributed to Dirac, it was first introduced by Heaviside, and is a good example of the zeroth theorem of the history of science: a discovery named after someone often did not originate with that person. See Jackson (2008).

function, it is probably best to think of it as the limit of a sequence of proper functions. There are many sequences which in a limit become delta functions, for example the function

$$\mathrm{H}_a(x) = \frac{1}{a\sqrt{\pi}} e^{-x^2/a^2} \tag{3.1.3}$$

It is easy to show that $\mathrm{H}_a(x)$ satisfies (3.1.1) and (3.1.2) in the limit that *a* goes to zero. Our intuitive picture of a delta function is thus a function which is highly peaked around *x = 0*, with unit area under the function. It is then plausible that if we take a well-behaved continuous function $f(x)$, we have

$$\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0) \tag{3.1.4}$$

In higher dimensions we define the delta function as a product of one-dimensional delta functions, e.g.

$$\delta^3(\mathbf{r}) = \delta(x)\,\delta(y)\,\delta(z) \tag{3.1.5}$$

with

$$\delta^3(\mathbf{r}) = 0 \qquad |\mathbf{r}| \neq 0 \tag{3.1.6}$$

and

$$\int_{all\,space} \delta^3(\mathbf{r})\, d^3x = 1 \tag{3.1.7}$$

One of the relations that we will find most useful in this course is the following:

$$\delta^3(\mathbf{r}) = -\frac{1}{4\pi} \nabla^2 \frac{1}{r} \tag{3.1.8}$$

To prove this relation, we first note that it is easy to show that $\nabla^2 \frac{1}{r} = 0$ for $r \neq 0$. To see that (3.1.7) holds, we use the divergence theorem (2.2.4)

$$\begin{aligned} \int_{all\,space} \nabla^2 \frac{1}{r} d^3x &= \int_{all\,space} \nabla \cdot \left[\nabla \frac{1}{r} \right] d^3x = \int_{surface\,sphere} \left[\nabla \frac{1}{r} \right] \cdot \hat{\mathbf{n}}\, da = \\ &= \int_{surface\,sphere} \left[-\frac{\hat{\mathbf{n}}}{r^2} \right] \cdot \hat{\mathbf{n}}\, da = \int_{surface\,sphere} \left[-\frac{r^2 d\Omega}{r^2} \right] = -4\pi \end{aligned} \tag{3.1.9}$$

3.2 Useful relations

From the definitions above, there are a number of useful relations for the delta function which we list here

$$\int_{-\infty}^{\infty} \delta(x-a)\, f(x)dx = f(a) \tag{3.2.1}$$

$$\int_{-\infty}^{\infty} \delta'(x-a)\, f(x)dx = -f'(a) \tag{3.2.2}$$

$$\int_{-\infty}^{\infty} f(x)\,\delta\left[g(x)\right]dx = \frac{f(a)}{\left|g'(a)\right|} \quad where \quad g(a)=0 \tag{3.2.3}$$

$$\nabla^2 \frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|} = -4\pi\,\delta^3(\mathbf{r}-\mathbf{r}') \tag{3.2.4}$$

3.3 Complete sets of orthorgonal functions on a finite interval

A denumerably infinite set of functions $\left\{ f_n(x)\right\}\Big|_{n=0}^{\infty}$ of the interval $[-1,1]$ is *orthorgonal* if for any *n* and *m*,

$$\int_{-1}^{1} f_n(x) f_m(x)\, dx = C_n \delta_{nm} \tag{3.3.1}$$

Here δ_{jn} is the Kronecker delta (δ_{jn} is 1 if $j = n$ and 0 otherwise). The set is complete is for any "nice" function $g(x)$ defined on $[-1,1]$, we can expand $g(x)$ as

$$g(x) = \sum_{n=0}^{\infty} a_n f_n(x) \tag{3.3.2}$$

where $\left\{a_n\right\}\Big|_{n=0}^{\infty}$ is a set of constants. We can use the orthogonality property of our complete set of functions by first multiplying (3.3.2) by $f_m(x)$

$$g(x)\, f_m(x) = \sum_{n=0}^{\infty} a_n f_n(x)\, f_m(x) \tag{3.3.3}$$

and then integrating (3.3.3) from -1 to 1, yielding

$$\int_{-1}^{1} g(x) f_m(x) dx = \sum_{n=0}^{\infty} a_n \int_{-1}^{1} f_n(x) f_m(x) dx = \sum_{n=0}^{\infty} a_n C_n \delta_{nm} \tag{3.3.4}$$

I thus find that my constants are given by

$$a_m = \frac{1}{C_m} \int_{-1}^{1} g(x) f_m(x) dx \tag{3.3.5}$$

If for every *m*, $C_m = 1$, then my set of functions is said to be *orthonormal.*

3.4 Representation of a delta function in terms of a complete set of functions

Suppose my function $g(x)$ is a delta function; that is, suppose

$$g(x) = \delta(x - x_o) \tag{3.4.1}$$

Then from (3.3.5) I have

$$a_m = \frac{1}{C_m} \int_{-1}^{1} \delta(x - x_o) f_m(x) dx = \frac{1}{C_m} f_m(x_o) \tag{3.4.2}$$

and if I insert this into (3.3.2) I find that

$$\delta(x - x_o) = \sum_{n=0}^{\infty} \frac{f_n(x) f_n(x_o)}{C_n} \tag{3.4.3}$$

Thus I have a representation of a one dimensional delta function for every complete set of functions, and there are literally an infinite number of complete sets of function on any finite interval.

4 Conservation of energy and momentum in electromagnetism

4.1 Learning objectives

In Section 2 above I talked about the general form of conservation laws for scalar and vector quantities. I now turn to the question of the energy flow and momentum flow in electromagnetism. I introduce the Poynting flux vector and the Maxwell Stress Tensor.

4.2 Maxwell's Equations

Electric and magnetic fields are produced by charges and currents, and Maxwell's equations tell us how the fields are produced by the charge density ρ and current density $\mathbf{J}$. Maxwell's equations relating the fields to their sources are

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_o} \tag{4.2.1}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{4.2.2}$$

$$\nabla \times \mathbf{B} = \mu_o \mathbf{J} + \mu_o \varepsilon_o \frac{\partial \mathbf{E}}{\partial t} \tag{4.2.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{4.2.4}$$

4.3 Conservation of charge

If I take the divergence of (4.2.3) and use (4.2.1), I obtain the differential equation for the conservation of charge (that is, Maxwell's equations contain charge conservation)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \tag{4.3.1}$$

If I consider (4.3.1) in light of our treatment of conservation laws for a scalar quantity in Section 2.2 above, I see that **J** is the flux density of charge and that the volume creation rate for charge is zero, that is, electric charge is neither created nor destroyed.

4.4 Conservation of energy

If I use a vector identity in the first step below, and then use Maxwell's equations (4.2.2) and (4.2.3) in the second step, I easily have

$$\begin{aligned} \nabla \cdot \left(\frac{\mathbf{E} \times \mathbf{B}}{\mu_o} \right) &= \frac{1}{\mu_o} \left(\mathbf{B} \cdot \nabla \times \mathbf{E} - \mathbf{E} \cdot \nabla \times \mathbf{B} \right) = \frac{1}{\mu_o} \left(\mathbf{B} \cdot \left(-\frac{\partial \mathbf{B}}{\partial t} \right) - \mathbf{E} \cdot \left(\mu_o \mathbf{J} + \mu_o \varepsilon_o \frac{\partial \mathbf{E}}{\partial t} \right) \right) \\ &= \frac{1}{2\mu_o} \frac{\partial \mathbf{B} \cdot \mathbf{B}}{\partial t} - \frac{\varepsilon_o}{2} \frac{\partial \mathbf{E} \cdot \mathbf{E}}{\partial t} - \mathbf{J} \cdot \mathbf{E} \end{aligned} \tag{4.4.1}$$

This can be re-written as

$$\frac{\partial}{\partial t} \left[\frac{1}{2} \varepsilon_o E^2 + \frac{B^2}{2\mu_o} \right] + \nabla \cdot \left(\frac{\mathbf{E} \times \mathbf{B}}{\mu_o} \right) = -\mathbf{E} \cdot \mathbf{J} \tag{4.4.2}$$

From the general form of the conservations laws I considered above, I see that I can interpret $\frac{1}{2}\varepsilon_o E^2 + \frac{B^2}{2\mu_o}$ as the energy density of the electromagnetic field (joules per cubic meter), $\frac{\mathbf{E}\times\mathbf{B}}{\mu_o}$ as the flux density of electromagnetic energy (joules per square meter per second), and $-\mathbf{E}\cdot\mathbf{J}$ as the volume creation rate of electromagnetic energy (joules per cubic meter per second).

4.5 Conservation of momentum and angular momentum

I define the Maxwell Stress Tensor $\overleftrightarrow{\mathbf{T}}$ as

$$\overleftrightarrow{\mathbf{T}} = \varepsilon_o\left[\mathbf{EE} - \frac{1}{2}\overleftrightarrow{\mathbf{I}}\,E^2\right] + \frac{1}{\mu_o}\left[\mathbf{BB} - \frac{1}{2}\overleftrightarrow{\mathbf{I}}\,B^2\right] \tag{4.5.1}$$

In Problem Set 1, you proved the vector identity

$$\nabla\cdot\left[\mathbf{AA} - \frac{1}{2}\overleftrightarrow{\mathbf{I}}\,A^2\right] = \mathbf{A}\left(\nabla\cdot\mathbf{A}\right) + \left(\nabla\times\mathbf{A}\right)\times\mathbf{A} \tag{4.5.2}$$

Using this identity and Maxwell's equations, I have

$$\begin{aligned}\nabla\cdot\overleftrightarrow{\mathbf{T}} &= \varepsilon_o\nabla\cdot\left[\mathbf{EE} - \frac{1}{2}\overleftrightarrow{\mathbf{I}}\,E^2\right] + \frac{1}{\mu_o}\nabla\cdot\left[\mathbf{BB} - \frac{1}{2}\overleftrightarrow{\mathbf{I}}\,B^2\right] \\ &= \varepsilon_o\mathbf{E}\nabla\cdot\mathbf{E} + \varepsilon_o\left(\nabla\times\mathbf{E}\right)\times\mathbf{E} + \frac{1}{\mu_o}\mathbf{B}\nabla\cdot\mathbf{B} + \frac{1}{\mu_o}\left(\nabla\times\mathbf{B}\right)\times\mathbf{B} \\ &= \mathbf{E}\rho + \varepsilon_o\left(-\frac{\partial\mathbf{B}}{\partial t}\right)\times\mathbf{E} + \mathbf{J}\times\mathbf{B} + \varepsilon_o\frac{\partial\mathbf{E}}{\partial t}\times\mathbf{B} = \rho\mathbf{E} + \mathbf{J}\times\mathbf{B} + \varepsilon_o\frac{\partial\left(\mathbf{E}\times\mathbf{B}\right)}{\partial t}\end{aligned} \tag{4.5.3}$$

Rearranging terms gives me the following equation

$$\frac{\partial}{\partial t}\left[\varepsilon_o\mathbf{E}\times\mathbf{B}\right] + \nabla\cdot\left(-\overleftrightarrow{\mathbf{T}}\right) = -\left[\rho\mathbf{E} + \mathbf{J}\times\mathbf{B}\right] \tag{4.5.4}$$

Given my general form of a conservation law for a vector quantity, equation (2.3.3), I identify $\varepsilon_o\mathbf{E}\times\mathbf{B}$ as electromagnetic momentum density, $-\overleftrightarrow{\mathbf{T}}$ as the flux density of electromagnetic momentum, and $-\left[\rho\mathbf{E} + \mathbf{J}\times\mathbf{B}\right]$ as the volume creation rate of electromagnetic momentum. Let me show that these three quantities have the appropriate units, e.g. momentum per cubic meter, momentum per square meter per sec, and momentum per cubic meter per sec, respectively.

The relation between the units of E and the units of B is $E = B\,L/T$ (I know this because the force on a charge is $q\left(\mathbf{E}+\mathbf{V}\times\mathbf{B}\right)$). Therefore the units of $\varepsilon_o EB$ are the units of $\varepsilon_o E^2 T/L$, and since $\varepsilon_o E^2$ is an energy density, it has units of $energy/L^3$. But *energy* has units of *force* times *distance*, and *force* has units of *momentum/T*, so *energy density* has units of *momentum /TL²*. So the units of $\varepsilon_o EB$ are *momentum/L³*, as desired. The tensor $\overleftrightarrow{\mathbf{T}}$ has units of energy density, and energy density has units of *momentum /TL²*, or momentum flux density. The units of $-\left[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}\right]$ are force per unit volume, and since *force* is *momentum over time*, this has units of *momentum per cubic meter per sec*, as I expect for a creation rate for electromagnetic momentum density.

If I look at (4.5.4) in integral form, I see that for any volume V contained by a closed surface S, I have

$$\frac{d}{dt}\int_V\left[\varepsilon_o\mathbf{E}\times\mathbf{B}\right]d^3x+\int_S\left(-\overleftrightarrow{\mathbf{T}}\right)\cdot\hat{\mathbf{n}}\,da=-\int_V\left[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}\right]d^3x \tag{4.5.5}$$

The corresponding equations for angular momentum are

$$\frac{\partial}{\partial t}\mathbf{r}\times\left[\varepsilon_o\mathbf{E}\times\mathbf{B}\right]+\nabla\cdot\left(-\mathbf{r}\times\overleftrightarrow{\mathbf{T}}\right)=-\mathbf{r}\times\left[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}\right] \tag{4.5.6}$$

and

$$\frac{d}{dt}\int_V\mathbf{r}\times\left[\varepsilon_o\mathbf{E}\times\mathbf{B}\right]d^3x+\int_S\left(-\mathbf{r}\times\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}\right)da=-\int_V\mathbf{r}\times\left[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}\right]d^3x \tag{4.5.7}$$

4.5.1 The Maxwell stress tensor in statics

To get some idea of the properties of the Maxwell stress tensor, I first look at it in cases where this is no time dependence, that is in electrostatics and magnetostatics. In this case, (4.5.5) can be written as

$$\int_V\left[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}\right]d^3x=\int_S\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}\,da \tag{4.5.8}$$

The term on the left above is the volume force density in electromagnetism, integrated over the volume, so it is the total electromagnetic force on all the charges and currents inside the volume V. What (4.5.8) tells me is that I can compute this force in two different ways. First I can do it the obvious way, by sampling the volume, looking at the charge on each little volume element, and adding that up to get the total force. The right hand side says that I can do this calculation a totally different way. I do not ever have to go inside the volume and look at the individual charges and currents and the fields at the location of those charges and currents. Rather I can simply move around on the surface of the sphere containing the charges and currents, and simply look at the fields on the surface of that sphere, calculating $\ddot{\mathbf{T}} \cdot \hat{\mathbf{n}}\, da$ at each little area element, and noting that this depends only on the fields at that area element. Isn't that amazing?

Well actually it is not so amazing, because it is exactly what I should expect for any decent field theory. Remember the fields are the agents which transmit forces between material objects and I should be able to look at the fields themselves and figure out what kind of stresses they are transmitting. As an analogy to illustrate this point, consider the theory of pegboards interacting via connecting strings, as illustrated in Figure 4-1. The right pegboard exerts a net force on the left pegboard because the strings connecting the pegs carry tension. I can calculate the force on the left pegboard in two ways. I can move around in the interior of the pegboard, find each peg, and the strings attached to it, and add up the total force for that peg, and then move on to the other pegs and thus compute the total force on the left pegboard. This process is analogous to doing the volume integral of the electromagnetic force density in (4.5.8).

Or, I surround the left pegboard by an imaginary sphere, as shown in Figure 4-2, and simply walk around on the surface of that sphere, never looking inside the volume it contains. Whenever I see a string piercing the surface of the sphere, I know that that sting is transmitting a force across the imaginary surface and I can measure the direction and magnitude of that force. I explore the entire surface, add up the tension due to all the strings, and then have the total force on the left pegboard. This is analogous to doing the surface integral on the right side of (4.5.8)

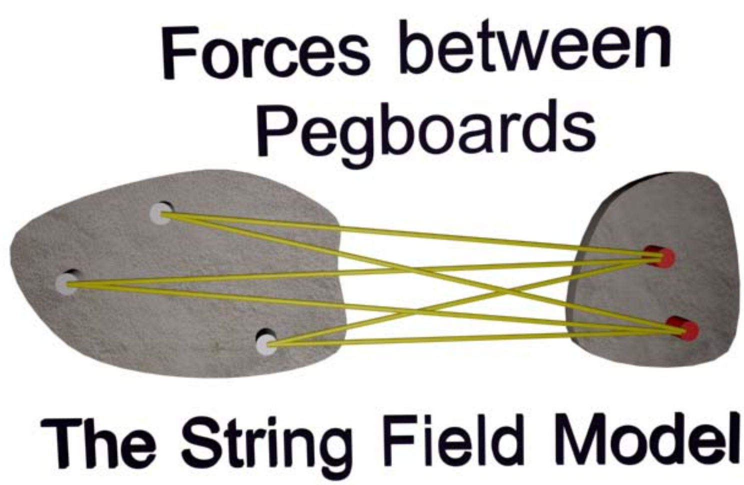


Figure 4-1: Two pegboards interacting through strings attached to the pegs

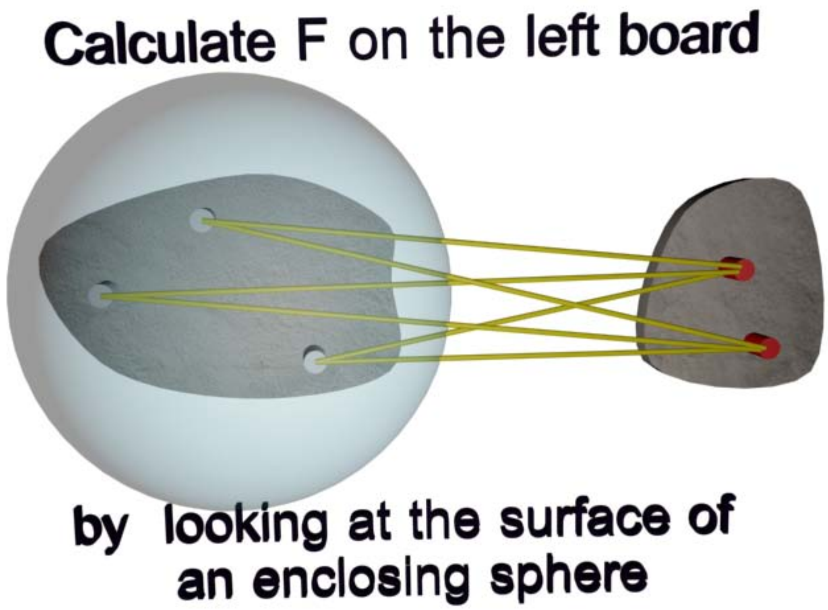


Figure 4-2: Enclosing the left pegboard by a sphere and exploring its surface.

4.5.2 Calculating $\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}\,da$

So this sounds neat, let me see what is actually involved in calculating $\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}\,da$ in electrostatics, for example. Figure 4-3 shows a surface element and the local electric field. Unless **E** and $\hat{\mathbf{n}}$ are co-linear, they determine a plane. Let the *x* axis in that plane be along the $\hat{\mathbf{n}}$ direction, and the *y* axis be perpendicular to the $\hat{\mathbf{n}}$ direction and in the **E**-$\hat{\mathbf{n}}$ plane. Let **E** make an angle θ with $\hat{\mathbf{n}}$.

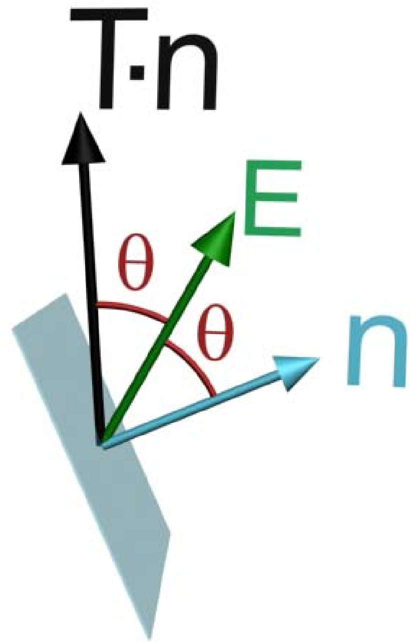


Figure 4-3: The relation between the directions of E, n, and $\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}$

In this coordinate system, I thus have the components of **E** as

$$\mathbf{E} = E\cos\theta\,\hat{\mathbf{x}} + E\sin\theta\,\hat{\mathbf{y}} \tag{4.5.9}$$

and $\hat{\mathbf{n}} = \hat{\mathbf{x}}$. If I look at the definition of $\overleftrightarrow{\mathbf{T}}$, and the definition of $\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}$, I see that $\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}$ is a vector with components

$$\left(\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}\right)_i = T_{ij}n_j = T_{ix} \tag{4.5.10}$$

and therefore

$$\left(\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}\right) = T_{xx}\hat{\mathbf{x}} + T_{yx}\hat{\mathbf{y}} + T_{yx}\hat{\mathbf{z}} = \varepsilon_o\left(E_xE_x - \frac{1}{2}E^2\right)\hat{\mathbf{x}} + \varepsilon_o E_x E_y\hat{\mathbf{y}} \tag{4.5.11}$$

If I insert the values of the components of **E** into (4.5.11), and use some trig, I have

$$\begin{aligned}\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}} &= \varepsilon_o E^2\left(\cos^2\theta - \frac{1}{2}\right)\hat{\mathbf{x}} + \varepsilon_o E^2\sin\theta\cos\theta\,\hat{\mathbf{y}} \\ &= \frac{1}{2}\varepsilon_o E^2\left(\cos 2\theta\,\hat{\mathbf{x}} + \sin 2\theta\,\hat{\mathbf{y}}\right)\end{aligned} \tag{4.5.12}$$

So that I can conclude the following about $\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}$:

1. $\overset{\leftrightarrow}{\mathbf{T}}\cdot\hat{\mathbf{n}}$ lies in the plane defined by **E** and $\hat{\mathbf{n}}$.
2. If I go an angle θ to get to **E** from $\hat{\mathbf{n}}$, I have to go an angle 2θ to get to $\overset{\leftrightarrow}{\mathbf{T}}\cdot\hat{\mathbf{n}}$, in the same sense.
3. The magnitude of $\overset{\leftrightarrow}{\mathbf{T}}\cdot\hat{\mathbf{n}}$ is always $\frac{1}{2}\varepsilon_o E^2$

Since $\overset{\leftrightarrow}{\mathbf{T}}$ is has the same form for **E** and **B**, the only change in these rules for **B** is that the magnitude of $\overset{\leftrightarrow}{\mathbf{T}}\cdot\hat{\mathbf{n}}$ is $B^2/2\mu_o$

Let me take some particular configurations of **E** and $\hat{\mathbf{n}}$ and see what these rules tell me. Figure 4-4 shows various configurations. From studying this figure, I conclude that if **E** and $\hat{\mathbf{n}}$ are parallel or anti-parallel, the **E** field transmits a *pull* across the surface ($\overset{\leftrightarrow}{\mathbf{T}}\cdot\hat{\mathbf{n}}$ is along $\hat{\mathbf{n}}$ and thus out of the volume of interest). If **E** and $\hat{\mathbf{n}}$ are perpendicular, the **E** field transmits a *push* across the surface ($\overset{\leftrightarrow}{\mathbf{T}}\cdot\hat{\mathbf{n}}$ is opposite $\hat{\mathbf{n}}$ and thus into the volume of interest). For other orientations $\overset{\leftrightarrow}{\mathbf{T}}\cdot\hat{\mathbf{n}}$ is a combination of a push or a pull and a shear force.

Note that there is no force exerted on *da*. I am just evaluating the force transmitted by the field across *da*, just as I was looking at the force transmitted across *da* in our pegboard string field model above.

I conclude this discussion with an actual calculation for two opposite charges. Two charges with equal magnitude *q* but with opposite signs are located a distance *2d* apart, one along the positive *x*-axis a distance *d* from the origin, and the other along the negative *x*-axis also a distance *d* from the origin. The charges are glued in place (that is there is a mechanical force on each that keeps them from moving under the Coulomb repulsion). I enclose the charge on the negative *x*-axis inside a cube of side *H*, with *H* going to infinity. One of the faces of the cube lies in the *z*-*y* plane at *x = 0*, as shown in Figure 4-5.

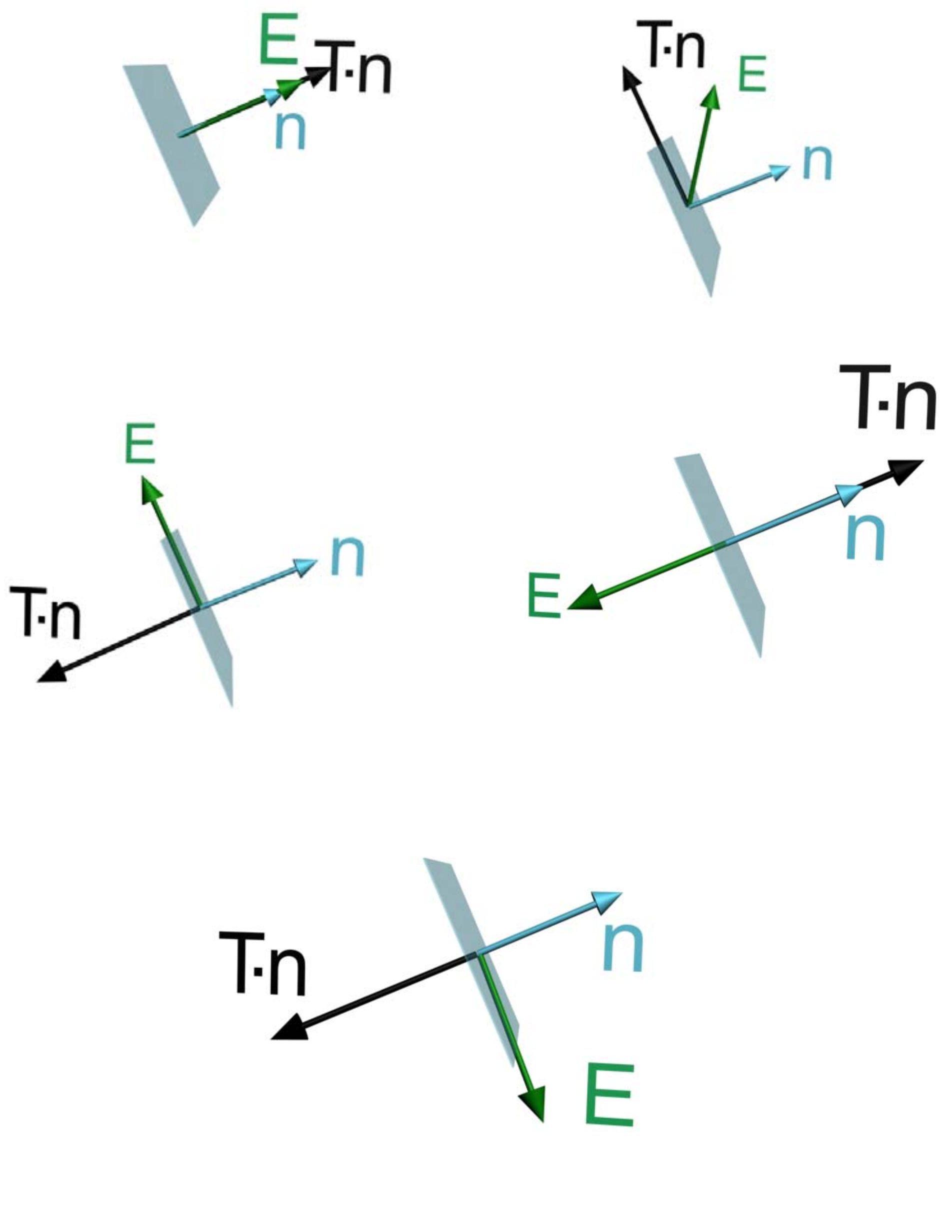


Figure 4-4: Orientation of $\overleftrightarrow{T} \cdot \hat{n}$ for various orientations of $\hat{n}$ and E.

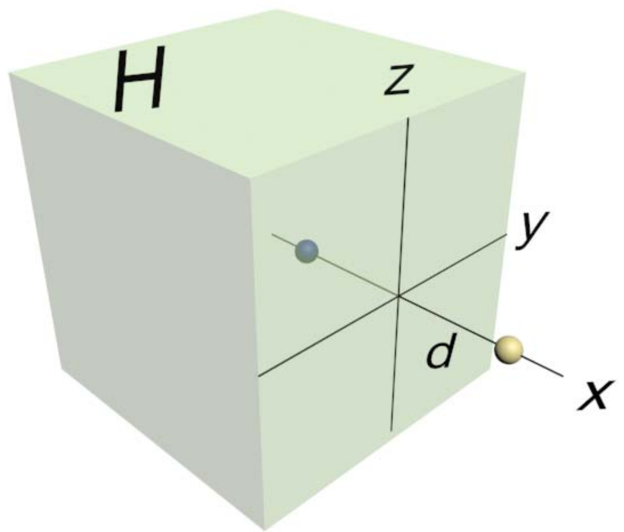
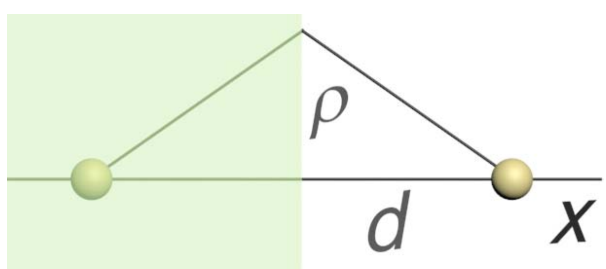


Figure 4-5: An electrostatics problem

If I look at the field configuration on the right *zy* face of the cube, it is parallel or anti-parallel to the normal to that face, and therefore the electrostatic force is exerting a *pull* across the cube's surface (that is, a force in the positive x direction). I can calculate that pull by integrating $\frac{1}{2}\varepsilon_o E^2$ over that surface, using the solution to this simple electrostatics problem, and I show that explicit calculation below. I find that the force is a force of attraction is $\frac{q^2}{4\pi\varepsilon_o(2d)^2}$, as I expect. The integral of the stress tensor over the other faces goes to zero as *H* goes to infinity.



Here is the explicit calculation of the force. If I look at any face of the cube except the *yz* face at $x = 0$, the stress tensor on that face (which goes as the square of the electric field) will fall off as one over distance to the fourth power and the area will only grow as distance squared, so I will get no contribution from these faces as H goes to infinity. I need only need to evaluate the stress tensor at $x = 0$, so I only need the electric field in the *yz* plane at $x = 0$. This electric field is always perpendicular to the x-axis in this plane, and has magnitude

$$2\frac{q}{4\pi\varepsilon_o}\frac{1}{d^2+\rho^2}\frac{\rho}{\sqrt{d^2+\rho^2}} \tag{4.5.13}$$

where ρ is the perpendicular distance from the *x*-axis to a given point in the *yz* plane. In our stress tensor calculations I need to calculate the following integral:

$$\int_{yz\ plane}\frac{1}{2}\varepsilon_o E^2 da = \frac{\varepsilon_o}{2}\int_0^\infty \rho d\rho \int_0^{2\pi} d\phi E^2 = \pi\varepsilon_o\int_0^\infty \rho d\rho E^2$$

$$= \pi\varepsilon_o\int_0^\infty \rho d\rho\left[\frac{q}{2\pi\varepsilon_o}\frac{\rho}{\left(d^2+\rho^2\right)^{3/2}}\right]^2 = \frac{q^2}{4\pi\varepsilon_o}\int_0^\infty \frac{\rho^3 d\rho}{\left(d^2+\rho^2\right)^3} =$$

$$\int_0^\infty \frac{\rho^3 d\rho}{\left(d^2+\rho^2\right)^3} = \int_0^\infty \frac{\rho d\rho}{\left(d^2+\rho^2\right)^3}\rho^2 = \text{by parts} = \frac{1}{2}\int_0^\infty \frac{\rho d\rho}{\left(d^2+\rho^2\right)^2} = -\left.\frac{1}{4\left(d^2+\rho^2\right)}\right|_0^\infty = \frac{1}{4d^2} \tag{4.5.14}$$

So

$$\int_{yz\ plane}\frac{1}{2}\varepsilon_o E^2 da = \frac{q^2}{4\pi\varepsilon_o}\frac{1}{\left(2d\right)^2}$$

From the properties of the stress tensor, I see that on the *yz* face at $x = 0$, since the electric field is perpendicular to the local normal, $\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}$ is a push on the cube, and a push is the negative *x* direction. Therefore

$$\int_{cube\ surface}\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}\,da = -\hat{\mathbf{x}}\int_{yz\ plane}\frac{1}{2}\varepsilon_o E^2 da = -\hat{\mathbf{x}}\frac{q^2}{4\pi\varepsilon_o}\frac{1}{\left(2d\right)^2}$$

Therefore the total electromagnetic momentum flux out of the cube surrounding the left charge is

$$\int_{cube\ surface}\left(-\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}\right)da = +\hat{\mathbf{x}}\frac{q^2}{4\pi\varepsilon_o}\frac{1}{\left(2d\right)^2}$$

You may well ask how there can be a flux of electromagnetic momentum when the density of electromagnetic momentum ($\varepsilon_o\mathbf{E}\times\mathbf{B}$) is zero. Consider the example of a current carrying wire. The positive ions are at rest and the electrons are moving. So there is no net electric charge, but there ***is*** a flux of electric charge (a current).

If we ask about the total rate at which electromagnetic momentum is being created in the cube, it is

$$\int\limits_{cube\ volume} \left(-\rho\mathbf{E}\right) d^3x = +\hat{\mathbf{x}}\frac{q^2}{4\pi\varepsilon_o}\frac{1}{\left(2d\right)^2} \tag{4.5.15}$$

and thus I see from the last two equations that electromagnetic momentum is being created at exactly the rate at which it is leaving the cube, as I expect, since this is a static situation.

Physically what is happening is the following. Some external agent (the glue holding the left charge down) is applying a force that exactly balances the electrostatic force of repulsion on the positive charge on the negative x-axis, that is there is a mechanical force in the $+\hat{\mathbf{x}}\frac{q^2}{4\pi\varepsilon_o}\frac{1}{\left(2d\right)^2}$. A force is a momentum per time, and thus the external agent is creating momentum at this rate. But the charge is not moving because the electrostatic force just balances this mechanical force. Since I cannot locally put the momentum into the charge, what happens is that the mechanical force is creating electromagnetic momentum, which then flows away from the charge on the left out of the volume containing it, and ultimately to the charge on the right. There it is absorbed by the mechanical force on the charge on the right, which is a sink of momentum in the +x direction.

5 The Helmholtz Theorem

5.1 Learning Objectives

We continue on with the theory of vector fields, particularly as applied to fluids. In this handout, the most important thing that we learn is that under certain assumptions about behavior at infinity, a vector field is specified uniquely by its divergence and curl. We then consider various flow fields derived from specific sources, to get some familiarity with vector fields given their divergence and curl.

5.2 The Helmholtz Theorem

If we specify both the curl $\mathbf{c}(\mathbf{r},t)$ and the divergence $s(\mathbf{r},t)$ of a vector function $\mathbf{F}(\mathbf{r},t)$, and postulate that these functions fall off at least as fast as $1/r^2$ at infinity, and that $\mathbf{F}(\mathbf{r},t)$ itself goes to zero at infinity, then that function $\mathbf{F}(\mathbf{r})$ is uniquely determined. This is known as the Helmholtz Theorem. To prove this theorem, we first construct a function that has the given divergence and curl, and then we show that this function is unique.

5.2.1 Construction of the vector function F

We construct the following function using our given curl and divergence functions.

$$\mathbf{F}(\mathbf{r},t) = -\nabla \frac{1}{4\pi} \int_{all\,space} \frac{s(\mathbf{r}',t)\, d^3x'}{|\mathbf{r}-\mathbf{r}'|} + \nabla \times \frac{1}{4\pi} \int_{all\,space} \frac{\mathbf{c}(\mathbf{r}',t)\, d^3x'}{|\mathbf{r}-\mathbf{r}'|} \qquad (5.2.1)$$

The divergence of this function is given by

$$\nabla \cdot \mathbf{F}(\mathbf{r},t) = -\nabla^2 \frac{1}{4\pi} \int_{all\,space} \frac{s(\mathbf{r}',t)\, d^3x'}{|\mathbf{r}-\mathbf{r}'|} + \frac{1}{4\pi} \nabla \cdot \left[\nabla \times \int_{all\,space} \frac{\mathbf{c}(\mathbf{r}',t)\, d^3x'}{|\mathbf{r}-\mathbf{r}'|} \right] \qquad (5.2.2)$$

The divergence of the curl of any function is zero, and we can switch the order of integration and differentiation in (5.2.2) and use (3.2.4), giving

$$\nabla \cdot \mathbf{F}(\mathbf{r},t) = -\frac{1}{4\pi} \int_{all\,space} s(\mathbf{r}',t) \nabla^2 \frac{1}{|\mathbf{r}-\mathbf{r}'|} d^3x' = \int_{all\,space} s(\mathbf{r}',t) \delta^3(\mathbf{r}-\mathbf{r}') d^3x' = s(\mathbf{r},t) \qquad (5.2.3)$$

Thus the vector we have constructed with (5.2.1) has the divergence that we desire. In a similar fashion we can show that the curl of this vector field (5.2.1) is $\mathbf{c}(\mathbf{r},t)$. It can be shown that our construction in (5.2.1) is unique as long as the divergence and curl fall off at least as fast as $1/r^2$ at infinity.

5.2.2 The inverse of the Helmholtz Theorem

The theorem above has the two important corollaries. If we have a function $\mathbf{F}(\mathbf{r},t)$ which falls off faster than $1/r$ at infinity, then

$$\nabla \times \mathbf{F} = 0 \quad \Rightarrow \quad \mathbf{F} = -\nabla \phi(\mathbf{r},t) \qquad (5.2.4)$$

or

$$\nabla \cdot \mathbf{F} = 0 \quad \Rightarrow \quad \mathbf{F} = \nabla \times \mathbf{A}(\mathbf{r},t) \qquad (5.2.5)$$

5.2.3 The Helmholtz Theorem in two dimensions

For future reference, we note that if the vector function **F** is entirely within the *x-y* plane and does not depend on *z*, that is

$$\mathbf{F}(\mathbf{r},t) = F_x(x,y,t)\,\hat{\mathbf{x}} + F_y(x,y,t)\,\hat{\mathbf{y}} \qquad (5.2.6)$$

then (5.2.1) becomes

$$\mathbf{F}(x,y,t) = -\nabla \frac{1}{2\pi} \int_{xy\ plane} s(x',y',t) \ln\left[\sqrt{(x-x')^2 + (y-y')^2}\right] dx'dy' \\ + \nabla \times \frac{1}{2\pi} \int_{xy\ plane} \mathbf{c}(x',y',t) \ln\left[\sqrt{(x-x')^2 + (y-y')^2}\right] dx'dy' \tag{5.2.7}$$

This can be shown to be the correct construction of F (that is, that it has the proper divergence and curl) by using the two-dimensional version of (3.1.8),

$$\delta(x)\delta(y) = -\frac{1}{2\pi}\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right] \ln\sqrt{x^2+y^2} \tag{5.2.8}$$

5.3 Examples of incompressible fluid flows

We consider here some incompressible fluid flow examples where the divergence and curl of the flow is given, and we proceed to construct the flow itself using the Helmholtz construction in (5.2.1). Incompressible means that the density $\rho_{mass}(\mathbf{r},t)$ is constant in time and space at ρ^o_{mass}, so that for incompressible flows the flow velocity **v** satisfies (see equation (2.2.6))

$$\nabla \cdot \mathbf{v} = \frac{s}{\rho^o_{mass}} \tag{5.3.1}$$

5.3.1 Irrotational flows

Let us look at some examples of irrotational flows. An irrotational flow is a flow which satisfies

$$\nabla \times \mathbf{v} = 0 \tag{5.3.2}$$

As a first example, consider an irrotational flow whose source function *s* is given by

$$s(\mathbf{r},t) = s_o \delta^3(\mathbf{r}) = s_o \delta(x)\delta(y)\delta(z) \tag{5.3.3}$$

Since the overall dimensions of the creation rate $s(\mathbf{r},t)$ must be mass per unit time per unit volume, and $\delta^3(\mathbf{r})$ has the dimensions of inverse volume, the dimensions of s_o must be mass per time. Equation (5.2.1) and (5.3.1) tell us that the vector field **v** is given by

$$\mathbf{v}(\mathbf{r}) = -\nabla \frac{1}{4\pi} \int_{all\,space} \frac{s_0 \delta^3(\mathbf{r}')\, d^3x'}{\rho^o_{mass} \left|\mathbf{r}-\mathbf{r}'\right|} = -\nabla \frac{s_o}{4\pi\rho^o_{mass} r} = \frac{s_o \hat{\mathbf{r}}}{4\pi\rho^o_{mass} r^2} \tag{5.3.4}$$

This is just a radial outflow from the source at the origin, with the flow velocity decreasing in magnitude as inverse distance squared. Note that the total rate at which mass is flowing out through a sphere of radius *R* (the flux of mass through the sphere) is given by

$$\int_S \left[\rho_{mass}\mathbf{v}\right]\cdot\hat{\mathbf{n}}\,da = \int_S \rho^o_{mass}\left(\frac{s_o\hat{\mathbf{r}}}{4\pi\rho^o_{mass}r^2}\right)\cdot\hat{\mathbf{n}}\,da = s_o \tag{5.3.5}$$

This is what we expect to see because the rate at which matter is being created within the sphere is given by

$$\int_V s\,d^3x = \int_V s_o\delta^3(\mathbf{r})\,d^3x = s_o \tag{5.3.6}$$

and so s_o should be the rate at which we see matter flowing out through the surface of the sphere.

As a second example consider the source function for an irrotational flow given by

$$s(\mathbf{r},t) = s_1\delta^3(\mathbf{r}) - s_2\delta^3(\mathbf{r}-L\,\hat{\mathbf{z}}) \tag{5.3.7}$$

This represents a source of fluid at the origin of strength s_1 along with a sink of fluid of strength s_2 at $L\,\hat{\mathbf{z}}$. Using (5.2.1), we can easily find that in this case

$$\mathbf{v}(\mathbf{r}) = \left[-\nabla\frac{s_1}{4\pi\rho^o_{mass}r} + \nabla\frac{s_2}{4\pi\rho^o_{mass}\left|\mathbf{r}-L\,\hat{\mathbf{z}}\right|}\right] = \left[\frac{s_1\hat{\mathbf{r}}}{4\pi\rho^o_{mass}r^2} - \frac{s_2\left(\mathbf{r}-L\,\hat{\mathbf{z}}\right)}{4\pi\rho^o_{mass}\left|\mathbf{r}-L\,\hat{\mathbf{z}}\right|^3}\right] \tag{5.3.8}$$

Note that the total rate at which mass is flowing out through a sphere (the flux of mass through the sphere) of very large radius $R >> d$ is given by

$$\int_S \left[\rho_{mass}\mathbf{v}\right]\cdot\hat{\mathbf{n}}\,da = s_1 - s_2 \tag{5.3.9}$$

This is what we expect to see because the rate at which matter is being created within this sphere is given by

$$\int_V s\,d^3x = s_1 - s_2 \tag{5.3.10}$$

If $s_1 = 5s_2$, then the velocity field topologically has the same form as the field plotted in Figure 1-1-1 and Figure 1-1-2 above.

5.3.2 Flows with rotation

Finally let us consider two flows where the curl of the flow is not zero. The first is a "sourceless" flow, that is $\nabla \cdot \mathbf{v} = 0$, but one for which the curl is given by

$$\nabla \times \mathbf{v} = \left[\delta(x)\delta(y-100) - 2\delta(x)\delta(y+140)\right]\hat{\mathbf{z}} \tag{5.3.11}$$

Figure 5-1: A flow with no source but with rotation

To see a movie of this kind of sourceless flow, follow the link below

://web.mit.edu/viz/EM/visualizations/vectorfields/FluidFlows/FluidFlowCurlCurl02/ffcurlcurl02.htm

The second example of a flow with rotation is a flow with the same curl as (5.3.11) but now instead of a zero divergence, a divergence given by

$$\nabla \cdot \mathbf{v} = \delta(x-250)\delta(y) \tag{5.3.12}$$

Figure 5-2: A flow with a source and rotation

There are a number of movies of different kinds of flows at the link below

://web.mit.edu/viz/EM/visualizations/vectorfields/FluidFlows/

6 The Solution to the Easy E&M

6.1 Learning Objectives

In this handout, I dive into electromagnetic theory and write out explicitly the solutions to the easy electromagnetism. By easy electromagnetism I refer to the situation where the sources of electromagnetic fields are known for all space and time. The sources of electromagnetic field are charges and currents (*"currents"* are moving charges). *"Known for all space and time"* means that someone gives us those functions, and your task is to deduce the electric and magnetic fields that these sources produce. This task is "easy" in that I can immediately write down a solution for **E** and **B** which can be solved with a straightforward and perfectly well defined algorithmic procedure.

6.2 The Easy Electromagnetism

6.2.1 The Solution to Maxwell's Equations

To solve the equations given in Section 4.2 for **E** and **B** given ρ and $\mathbf{J}$, I first introduce the vector potential **A**. Because the divergence of **B** is zero, we know from the Helmhotz theorem that **B** can be written as the curl of a vector, the vector potential **A**.

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{6.1.1}$$

If we insert (6.1.1) into (4.2.2), we have that

$$\nabla \times \left[\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right] = 0 \tag{6.1.2}$$

Since the curl of the vector $\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}$ is zero, we can write it as the gradient of a scalar function, which we will denote by ϕ, so that

$$\mathbf{E} = -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} \tag{6.1.3}$$

If we insert (6.1.1) and (6.1.2) into (4.2.3), we have

$$\nabla \times \nabla \times \mathbf{A} = \mu_o \mathbf{J} + \mu_o \varepsilon_o \frac{\partial}{\partial t} \left[-\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} \right] \tag{6.1.4}$$

or

$$\nabla \left(\nabla \cdot \mathbf{A} \right) - \nabla^2 \mathbf{A} = \mu_o \mathbf{J} - \mu_o \varepsilon_o \nabla \frac{\partial \phi}{\partial t} - \mu_o \varepsilon_o \frac{\partial^2}{\partial t^2} \mathbf{A} \tag{6.1.5}$$

If we let

$$c^2 = \frac{1}{\mu_o \varepsilon_o} \tag{6.1.6}$$

then

$$\left[\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \mathbf{A} + \nabla \left(\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} \right) = -\mu_o \mathbf{J} \tag{6.1.7}$$

We still have the freedom to specify the divergence of **A**, since up to this point we have only specified its curl, and we choose the divergence of **A** so that it satisfies

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} = 0 \tag{6.1.8}$$

This is knownas the Lorentz gauge condition, although it should be more properly called the Lorenz gauge condition (see Jackson (2008)).

With (6.1.8), (6.1.7) becomes

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\mathbf{A} = -\mu_o \mathbf{J} \tag{6.1.9}$$

And we can easily show by inserting (6.1.3) into (4.2.1) and using (6.1.8) that

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\phi = -\frac{\rho}{\varepsilon_o} \tag{6.1.10}$$

6.2.2 The free space-time dependent Green's function

To solve (6.1.9) and (6.1.10), we first need to solve the equation

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]G(\mathbf{r},t) = \delta^3\left(\mathbf{r}\right)\delta(t) \tag{6.2.1}$$

The function $G(\mathbf{r},t)$ is the response of the system to a point disturbance in space and time. Once we know this we can write down a general solution for sources distributed in space and time. The solution to (6.2.1) is

$$G(\mathbf{r},t) = -\frac{1}{4\pi}\frac{\delta(t - r/c)}{r} \tag{6.2.2}$$

which can be verified by direct substitution, as follows. First, we have

$$\nabla\left[\frac{\delta(t-\frac{r}{c})}{r}\right] = \delta(t-\frac{r}{c})\nabla\frac{1}{r} + \frac{1}{r}\nabla\delta(t-\frac{r}{c}) = -\delta(t-\frac{r}{c})\frac{\hat{\mathbf{r}}}{r^2} - \frac{\hat{\mathbf{r}}}{cr}\delta'(t-\frac{r}{c}) \tag{6.2.3}$$

Dotting ∇ into (6.2.3) gives

$$\nabla^2\left[\frac{\delta(t-\frac{r}{c})}{r}\right]=\nabla\cdot\nabla\left[\frac{\delta(t-\frac{r}{c})}{r}\right]$$
$$=-\nabla\cdot\left[\delta(t-\frac{r}{c})\frac{\hat{\mathbf{r}}}{r^2}+\frac{\hat{\mathbf{r}}}{cr}\delta'(t-\frac{r}{c})\right]$$
$$=-\frac{1}{r^2}\frac{\partial}{\partial r}\left[\delta(t-\frac{r}{c})+\frac{r}{c}\delta'(t-\frac{r}{c})\right]$$
$$=-4\pi\delta(t-\frac{r}{c})\delta^3(\mathbf{r})+\frac{1}{cr^2}\delta'(t-\frac{r}{c})-\frac{1}{cr^2}\delta'(t-\frac{r}{c})+\frac{1}{c^2r}\delta''(t-\frac{r}{c})$$
$$=-4\pi\delta(t-\frac{r}{c})\delta^3(\mathbf{r})+\frac{1}{c^2r}\delta''(t-\frac{r}{c}) \qquad (6.2.4)$$

But

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left[\frac{\delta(t-\frac{r}{c})}{r}\right]=\left[\frac{\delta''(t-\frac{r}{c})}{c^2r}\right] \qquad (6.2.5)$$

Subtracting the two expressions (6.2.4) and (6.2.5), we have

$$\left[\nabla^2-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\left[\frac{\delta(t-\frac{r}{c})}{r}\right]=-4\pi\delta^3(\mathbf{r})\,\delta(t-\frac{r}{c})=-4\pi\delta^3(\mathbf{r})\,\delta(t) \qquad (6.2.6)$$

where the last form on the right hand side is true because the delta function in **r** means that we only have a contribution when $r = 0$, so we can take r to be zero in the argument of the delta function in time.

Then, we see (by shifting the origin of space and time by $\mathbf{r}'$ and t') that

$$\left[\nabla^2-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]G(\mathbf{r},\mathbf{r}',t,t')=\delta^3\left(\mathbf{r}-\mathbf{r}'\right)\delta(t-t') \qquad (6.2.7)$$

where

$$G(\mathbf{r},\mathbf{r}',t,t')=-\frac{1}{4\pi}\frac{\delta(t-t'-\left|\mathbf{r}-\mathbf{r}'\right|/c)}{\left|\mathbf{r}-\mathbf{r}'\right|} \qquad (6.2.8)$$

6.2.3 The solution for $(\phi, \mathbf{A})$ given $(\rho, \mathbf{J})$ for all space and time

We now assert that the solution to (6.1.9) is

$$\mathbf{A}(\mathbf{r},t) = -\mu_o \int_{all\,time} dt' \int_{all\,space} G(\mathbf{r},\mathbf{r}',t,t')\mathbf{J}(\mathbf{r}',t')\,d^3x' \qquad (6.2.9)$$

To see that this is indeed the solution to (6.1.9), we apply the operator

$$\Box^2 = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \qquad (6.2.10)$$

to equation (6.2.9), yielding (using (6.2.7))

$$\begin{aligned}\Box^2\mathbf{A}(\mathbf{r},t) &= -\mu_o \int_{all\,time} dt' \int_{all\,space} \Box^2 G(\mathbf{r},\mathbf{r}',t,t')\mathbf{J}(\mathbf{r}',t')\,d^3x' \\ &= -\mu_o \int_{all\,time} \delta(t-t')dt' \int_{all\,space} \delta^3(\mathbf{r}-\mathbf{r}')\,\mathbf{J}(\mathbf{r}',t')\,d^3x' = -\mu_o\mathbf{J}(\mathbf{r},t)\end{aligned} \qquad (6.2.11)$$

where we have used the delta functions to carry out the time and space integrations. Our solution (6.2.9) can thus be written as (using (6.2.8))

$$\mathbf{A}(\mathbf{r},t) = \frac{\mu_o}{4\pi}\int_{all\,time} dt' \int_{all\,space} \frac{\mathbf{J}(\mathbf{r}',t')}{|\mathbf{r}-\mathbf{r}'|} dt'\delta(t-t'-|\mathbf{r}-\mathbf{r}'|/c)\,d^3x' \qquad (6.2.12)$$

If we use the delta function in time to do the t' integration in (6.2.12) we have finally

$$\mathbf{A}(\mathbf{r},t) = \frac{\mu_o}{4\pi}\int \frac{\mathbf{J}(\mathbf{r}',t'_{ret})}{|\mathbf{r}-\mathbf{r}'|}\,d^3x' \qquad (6.2.13)$$

where

$$t'_{ret} = t - \frac{|\mathbf{r}-\mathbf{r}'|}{c} \qquad (6.2.14)$$

and similarly we have

$$\phi(\mathbf{r},t) = \frac{1}{4\pi\,\varepsilon_o}\int \frac{\rho(\mathbf{r}',t'_{ret})}{|\mathbf{r}-\mathbf{r}'|}\,d^3x' \qquad (6.2.15)$$

6.3 What does the observer see at time *t*?

The prescription as to how to do the spatial integrals in (6.2.13) and (6.2.15) using the definition of retarded time (6.2.14) is unusual. Because of the finite propagation time from source to observer at the observer's time *t*, the spatial integrals are sampling what happened in sources more distant from the observer at an earlier time than the sources closer to the observer. That is, the value of t'_{ret} depends on the distance from the a source to the observer. Although we are integrating over all space in the equations above, we are adding up contributions from different volumes of space at different times in the observer's past.

6.3.1 The collapsing information gathering sphere

One way to understand this sampling (Panofsky and Phillips (1962)) is to consider what information is seen by an observer located at the origin at $t = 0$. The information that arrives at the observer at the origin at $t = 0$ has been collected by a sphere of radius $r' = -ct'$ that has been collapsing toward the observer at the speed of light since time began, as shown in Figure 6-1. The observer at time t will see all the light collected by this information collecting sphere. The center of the sphere is at the location of the observer at time t, and the sphere has been contracting since the beginning of time with a radial velocity c such that it has just converged on the observer at time t. The time t'_{ret} at which this information-collecting sphere passes a source at r' at any point is space is then the time at which that source produced the effect which is seen by the observer at time t.

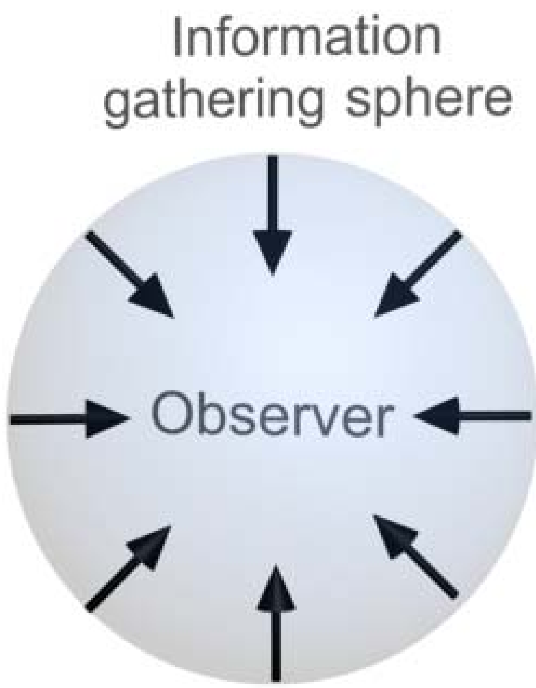


Figure 6-1: The information gathering sphere collapsing toward the origin

6.3.2 The backward light cone

Another way to envisage this process is to look at a space-time diagram. In such a diagram, we plot *ct* along the vertical axes and two spatial coordinates (say *x* and *y*) perpendicular to this axis. An 'event" in space-time, say a firecracker exploding, is located by its time of occurrence and the place at which it occurred. A burst of light emitted by the observer at the origin at $t = 0$ in this diagram propagates outward at the speed of light in all directions, and the locus of space-time points on that outwardly propagating sphere is represented by the forward light cone shown in Figure 6-2. The forward light cone is a cone whose apex is at the origin and whose opening angle is 45 degrees. Only observers at points in space-time lying on the forward light cone will see the burst of light emitted by our observer at $t = 0$. Similarly, the light seen by our observer at the origin at $t = 0$ must have originated at some point in space-time on the observer's backward light cone, since for only those points will the radiation just be arriving at the observer's position at $t = 0$.

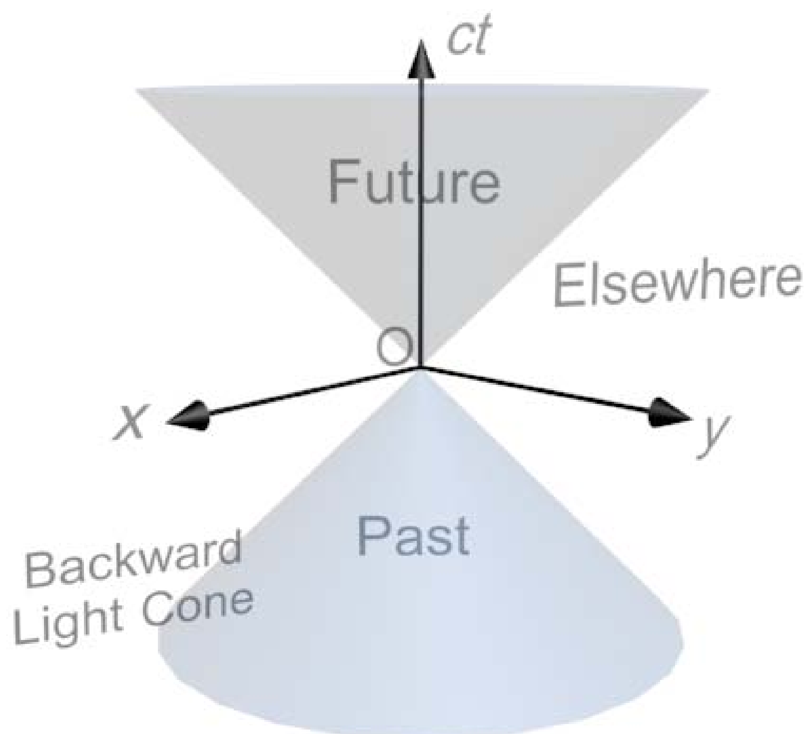


Figure 6-2: The observer's forward and backward light cone

7 E and B far from localized static sources

7.1 Learning objectives

We first investigate the form of the electric and magnetic fields far from localized sources, assuming that these sources do not vary in time. "Localized" means that our sources vanish outside of a sphere of radius d. "Far from" mean we are at radii such thar $r >> d$. Our major goal here is to show that everything "looks" the same if you get far enough away, and to introduce the idea of moments, that is, dipole moments, quadrupole moments, and so on.

7.2 A systematic expansion in powers of *d/r*

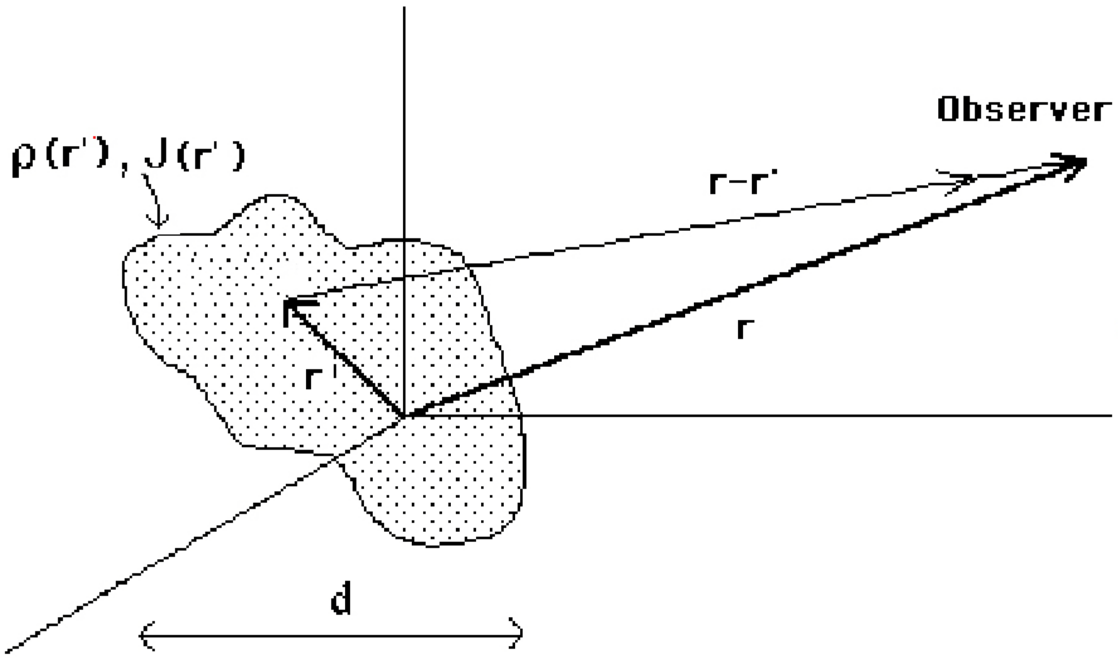


Figure 7-1: An observer far from a localized distribution of static sources.

In Figure 7-1, we show a distribution of time independent charges and currents which vanish outside a distance d from the orgin. I want to look at (6.2.13) and (6.2.15) when there is no time dependence and under the assumption that I am far away from the sources compared to d. For no time independence, I have

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_o}{4\pi}\int \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d^3x' \tag{7.2.1}$$

and

$$\phi(\mathbf{r}) = \frac{1}{4\pi\,\varepsilon_o}\int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d^3x' \tag{7.2.2}$$

Let's look at the $|\mathbf{r}-\mathbf{r}'|$ term in (7.2.2), assuming that the angle between $\mathbf{r}$ and $\mathbf{r}'$ is θ'. I have

$$|\mathbf{r}-\mathbf{r}'|^2 = (\mathbf{r}-\mathbf{r}')\cdot(\mathbf{r}-\mathbf{r}') = r^2 + (r')^2 - 2rr'\cos\theta' \tag{7.2.3}$$

I define

$$\eta = \left(\frac{r'}{r}\right)^2 - 2\frac{r'}{r}\cos\theta' = \left(\frac{r'}{r}\right)\left(\frac{r'}{r} - 2\cos\theta'\right) \tag{7.2.4}$$

and I assume that $r >> r'$, since I am taking the distance r of the observer large compared to d, and the r' value in the integral in (7.2.3) can always be assumed to be less that d, since the sources vanish outside of d. Then it is clear that η is a small quantitiy, and

$$\left|\mathbf{r} - \mathbf{r}'\right| = r\left(1+\eta\right)^{1/2} \qquad \frac{1}{\left|\mathbf{r} - \mathbf{r}'\right|} = \frac{1}{r}\left(1+\eta\right)^{-1/2} \tag{7.2.5}$$

I can expand $\left(1+\eta\right)^{-1/2}$ in a Taylor series, since I always have $\eta << 1$, giving

$$\frac{1}{\left|\mathbf{r} - \mathbf{r}'\right|} = \frac{1}{r}\left(1+\eta\right)^{-1/2} = \frac{1}{r}\left(1 - \frac{1}{2}\eta + \frac{3}{8}\eta^2 - \frac{5}{16}\eta^3 + \ldots\right) \tag{7.2.6}$$

If I use the definition of η in (7.2.4) in equation (7.2.6), and gather terms in powers of $\left(\frac{r'}{r}\right)$, I find that

$$\frac{1}{\left|\mathbf{r} - \mathbf{r}'\right|} = \frac{1}{r}\begin{pmatrix} 1 + \left(\frac{r'}{r}\right)\cos\theta' + \left(\frac{r'}{r}\right)^2 \frac{\left(3\cos^2\theta' - 1\right)}{2} \\ + \left(\frac{r'}{r}\right)^2 \frac{\left(5\cos^3\theta' - 3\cos\theta'\right)}{2} + \ldots \end{pmatrix} \tag{7.2.7}$$

If I look back at the Legendre polynomials I found in Problem Set 1, I have

$$\frac{1}{\left|\mathbf{r} - \mathbf{r}'\right|} = \frac{1}{r}\sum_{l=0}^{\infty}\left(\frac{r'}{r}\right)^l P_l\left(\cos\theta'\right) \tag{7.2.8}$$

Putting this into (7.2.1) and (7.2.2), I have

$$\begin{aligned} \mathbf{A}(\mathbf{r}) &= \frac{\mu_o}{4\pi}\int \mathbf{J}(\mathbf{r}')\left[\frac{1}{r}\sum_{l=0}^{\infty}\left(\frac{r'}{r}\right)^l P_l\left(\cos\theta'\right)\right] d^3x' \\ &= \frac{\mu_o}{4\pi}\sum_{l=0}^{\infty}\frac{1}{r^{l+1}}\int \mathbf{J}(\mathbf{r}')\left(r'\right)^l P_l\left(\cos\theta'\right) d^3x' \end{aligned} \tag{7.2.9}$$

and

$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_o}\sum_{l=0}^{\infty}\frac{1}{r^{l+1}}\int \rho(\mathbf{r}')(r')^l\, P_l(\cos\theta')\, d^3x' \tag{7.2.10}$$

We have in (7.2.9) and (7.2.10) what we wanted to achieve, an expansion in powers of $\left(\frac{d}{r}\right)$, so that each successive term is smaller than the preceeding by a factor of $\frac{d}{r} << 1$ for a distant observer.

7.3 The magnetic dipole and electric dipole terms

The first term in the sum in (7.2.9) is

$$\frac{\mu_o}{4\pi}\frac{1}{r}\int \mathbf{J}(\mathbf{r}')\, P_0(\cos\theta')\, d^3x' = \frac{\mu_o}{4\pi}\frac{1}{r}\int \mathbf{J}(\mathbf{r}')\, d^3x' \tag{7.3.1}$$

In Problem Set 3 you will show that this term vanishes in the time independent case. The second term in (7.2.9) is

$$\frac{\mu_o}{4\pi}\frac{1}{r^2}\int \mathbf{J}(\mathbf{r}')(r')\, P_1(\cos\theta')\, d^3x' = \frac{\mu_o}{4\pi}\frac{1}{r^2}\int \mathbf{J}(\mathbf{r}')(r'\cos\theta')\, d^3x' \tag{7.3.2}$$

If $\hat{\mathbf{n}} = \mathbf{r}/r$ is a unit vector which points from the origin of our coordinate system to the observer at $\mathbf{r}$, then $r'\cos\theta' = \hat{\mathbf{n}}\cdot\mathbf{r}'$, and

$$\frac{\mu_o}{4\pi}\frac{1}{r^2}\int \mathbf{J}(\mathbf{r}')(r'\cos\theta')\, d^3x' = \frac{\mu_o}{4\pi}\frac{1}{r^2}\int \mathbf{J}(\mathbf{r}')(\mathbf{r}'\cdot\hat{\mathbf{n}})\, d^3x' \tag{7.3.3}$$

In Problem Set 3, you will show that this can be written as

$$\frac{\mu_o}{4\pi}\frac{1}{r^2}\int \mathbf{J}(\mathbf{r}')(\mathbf{r}'\cdot\hat{\mathbf{n}})\, d^3x' = \frac{\mu_o}{4\pi}\frac{\mathbf{m}\times\hat{\mathbf{n}}}{r^2} \tag{7.3.4}$$

where I have defined the "magnetic dipole moment" **m** by

$$\mathbf{m} = \frac{1}{2}\int \mathbf{r}'\times \mathbf{J}(\mathbf{r}')d^3x' \tag{7.3.5}$$

Note that once you are given $\mathbf{J}(\mathbf{r}')$, this moment is a *fixed constant vector*! If I take the curl of (7.3.4), I find that the magnetic field associated with the first non-vanishing term in (7.2.9) is

$$\mathbf{B}_{dipole}(\mathbf{r}) = \frac{\mu_o}{4\pi}\frac{\left[3\hat{\mathbf{n}}(\mathbf{m}\cdot\hat{\mathbf{n}}) - \mathbf{m}\right]}{r^3} \tag{7.3.6}$$

Now let us turn to the expansion of the electric potential in (7.2.10). The first non-vanishing term in the sum is

$$\frac{1}{4\pi\varepsilon_o}\frac{1}{r}\int \rho(\mathbf{r}')P_0(\cos\theta')\,d^3x' = \frac{1}{4\pi\varepsilon_o}\frac{1}{r}\int\rho(\mathbf{r}')\,d^3x' = \frac{1}{4\pi\varepsilon_o}\frac{Q_o}{r} \tag{7.3.7}$$

where Q_o is the total charge of the static distribution. This is just the potential of a point charge, and the associated electric field is that of a point charge. The next term is

$$\begin{aligned} \frac{1}{4\pi\varepsilon_o}\frac{1}{r^2}\int\rho(\mathbf{r}')(r')P_1(\cos\theta')\,d^3x' &= \frac{1}{4\pi\varepsilon_o}\frac{1}{r^2}\int\rho(\mathbf{r}')(r'\cos\theta')\,d^3x' \\ = \frac{1}{4\pi\varepsilon_o}\frac{1}{r^2}\int\rho(\mathbf{r}')\,\mathbf{r}'\cdot\hat{\mathbf{n}}\,d^3x' &= \frac{1}{4\pi\varepsilon_o}\frac{\hat{\mathbf{n}}\cdot\mathbf{p}}{r^2} \end{aligned} \tag{7.3.8}$$

where we have defined the electric dipole moment **p** as

$$\mathbf{p} = \int \mathbf{r}'\rho(\mathbf{r}')d^3x' \tag{7.3.9}$$

If we take the gradient of (7.3.8) to find the electric field corresponding to this term, we have

$$\mathbf{E}_{dipole}(\mathbf{r}) = \frac{1}{4\pi\varepsilon_o}\frac{\left[3\hat{\mathbf{n}}(\mathbf{p}\cdot\hat{\mathbf{n}}) - \mathbf{p}\right]}{r^3} \tag{7.3.10}$$

7.4 Properties of a static dipole

Both the magnetic and electric dipole fields have the same form. I discuss the relevant properties of an electric dipole oriented along the z-axis. If **p** is along the *z*-axis, then

$$\mathbf{E}(\mathbf{r}) = \frac{2p\cos\theta}{4\pi\,\varepsilon_o r^3}\hat{\mathbf{r}} + \frac{p\sin\theta}{4\pi\,\varepsilon_o r^3}\hat{\boldsymbol{\theta}} \tag{7.4.1}$$

In spherical polar coordinates, our differential equations (1.5.1) for the field lines becomes

$$\frac{dr(s)}{ds} = \frac{E_r}{E} \qquad \frac{rd\theta(s)}{ds} = \frac{E_\theta}{E} \tag{7.4.2}$$

or we can write the differential equation $r(\theta)$ for a given field line by dividing the first equation in (7.4.2) by the second to obtain

$$\frac{1}{r}\frac{dr(\theta)}{d\theta} = \frac{E_r}{E_\theta} = \frac{2\cos\theta}{\sin\theta} \tag{7.4.3}$$

If we gather the terms involving *r* and the terms involving θ in (7.4.3), we have

$$\frac{dr}{r} = \frac{2\cos\theta}{\sin\theta}d\theta = 2\frac{d}{d\theta}\ln\left(\sin\theta\right) \tag{7.4.4}$$

which can be integrated to give the equation for a dipole field line.

$$r = L\sin^2\theta \tag{7.4.5}$$

The parameter L characterizing a given field line is the equatorial crossing distance of that field line. Figure 7-2 shows a family of such field lines with equatorial crossing distances equally space.

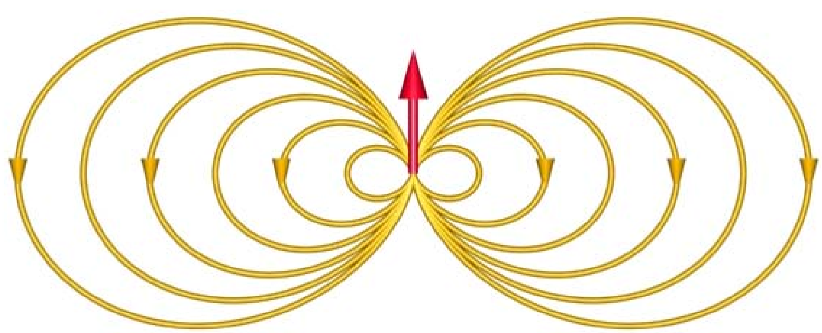


Figure 7-2: The field lines of a static dipole

7.5 The electric quadrupole term

If we go to the third term in the sum given in (7.2.10) for the electric potential, I can show that this term has the form

$$\frac{1}{4\pi\varepsilon_o}\frac{1}{2r^3}\sum_{j=1}^{3}\sum_{i=1}^{3}\hat{n}_i\hat{n}_jQ_{ij} = \frac{1}{4\pi\varepsilon_o}\frac{1}{2r^3}\hat{\mathbf{n}}\cdot\overleftrightarrow{\mathbf{Q}}\cdot\hat{\mathbf{n}} \tag{7.5.1}$$

where the electric quadrupole tensor $\overset{\leftrightarrow}{\mathbf{Q}}$ is given by

$$\left[\overset{\leftrightarrow}{\mathbf{Q}}\right]_{ij} = \int\left[3x_i'x_j' - \left(r'\right)^2 \delta_{ij}\right]\rho(\mathbf{r})d^3x' \qquad (7.5.2)$$

8 Sources varying slowly in time

8.1 Learning objectives

I now turn to the time case where our sources are localized and now are not static, but vary in time, but *slowly*, in the sense that any time dependence is slow compared to the speed of light travel time across the source, *d/c*. When I do this I will uncover the details what is arguably the most fundamental of all electromagnetic processes, the generation of electromagnetic waves. My major goal in this section is to show you how to systematically expand the solutions (6.2.13) and (6.2.15) in small parameters to get **E** and **B** for sources which vary slowly in time in the above sense. As I expect, I recover the static fields I have already seen above, but I also find much much more.

8.2 E and B fields far from localized sources varying slowly in time

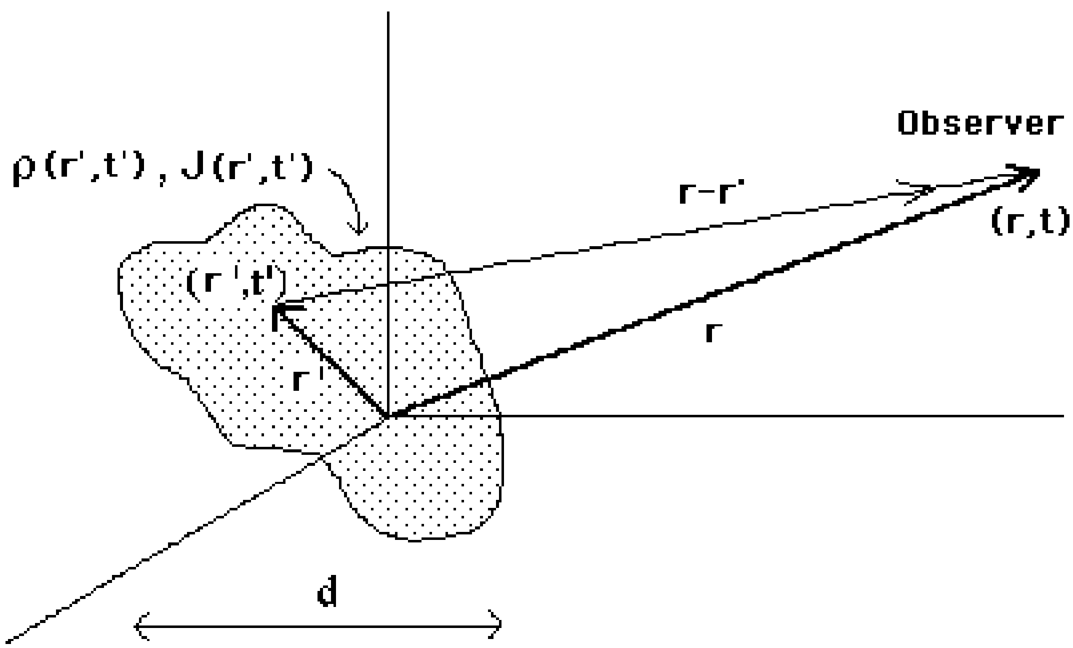


Figure 8-1: An observer far from a localized distribution of currents and charges

Suppose we have a localized distribution of charge and current near the origin, described by sources $\rho(\mathbf{r}',t')$ and $\mathbf{J}(\mathbf{r}',t')$ (Figure 8-1). The sources have a characteristic linear dimension *d*, such that the charge and current densities are zero for $r' > d$. Let the length of time for significant variation in the charge and current densities be *T*. We want to investigate the electromagnetic fields produced by these currents and charges as measured by an observer at $(\mathbf{r},t)$ located far from the sources, assuming that the sources vary slowly in time.

Again, as above, "far from" means that the observer's distance from the origin, *r*, is much greater than the maximum extent of the sources, *d*. "Slowly in time" means that the characteristic time *T* for significant variation is long compared to the speed of transit time across the sources, that is, *T >> d/c*. Under these two assumptions (*far from* in distance and *slowly* in time), we can develop straightforward expansions for $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$ to various orders in the small quantities *d/r* and *d/cT*. Note that we have made no assumption as of yet about how the distance *r* compares to *cT*, and in fact we will be interested below in **three very different cases**: $r << cT$, $r \approx cT$, and $r >> cT$, the so-called near, intermediate, and far zones (also called the quasi-static, intermediate, and radiation zones). The solutions to our equations look quite different in these three regimes.

To obtain the electromagnetic fields at $(\mathbf{r},t)$, we first calculate the electromagnetic potentials. We have the exact solutions for $\phi(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$ at any observation point $(\mathbf{r},t)$

$$\phi(\mathbf{r},t)=\frac{1}{4\pi\,\varepsilon_o}\int\rho(\mathbf{r}',t'_{ret})\frac{d^3x'}{\left|\mathbf{r}-\mathbf{r}'\right|};\qquad \mathbf{A}(\mathbf{r},t)=\frac{\mu_o}{4\pi}\int\mathbf{J}(\mathbf{r}',t'_{ret})\frac{d^3x'}{\left|\mathbf{r}-\mathbf{r}'\right|} \tag{8.2.1}$$

$$t'_{ret}=t-\left|\mathbf{r}-\mathbf{r}'\right|/c \tag{8.2.2}$$

For an observer far away ($r >> d$), I make the approximation that (see (7.2.3))

$$\begin{aligned}\left|\mathbf{r}-\mathbf{r}'\right|&=\sqrt{r^2-2\mathbf{r}\cdot\mathbf{r}'+r'^2}\approx r\sqrt{1-2\frac{\mathbf{r}\cdot\mathbf{r}'}{r^2}}\\ &\approx r\left(1-\frac{\mathbf{r}\cdot\mathbf{r}'}{r^2}\right)=r-\hat{\mathbf{n}}\cdot\mathbf{r}'+....,\qquad \hat{\mathbf{n}}=\mathbf{r}/\left|\mathbf{r}\right|\end{aligned} \tag{8.2.3}$$

where I now go only to terms of first order in $\left(\frac{r'}{r}\right)$, as opposed to keeping all orders, as I did above in (7.2.7). I also have

$$\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|}=\frac{1}{r}+\frac{\hat{\mathbf{n}}\cdot\mathbf{r}'}{r^2}+.... \tag{8.2.4}$$

Using (8.2.3) in (8.2.2) gives

$$t'_{ret}\cong t-r/c+\hat{\mathbf{n}}\cdot\mathbf{r}'/c \tag{8.2.5}$$

Expanding the exact solutions given in (8.2.1) is complicated for time varying sources because of the finite propagation time from field point to observation point. As we saw before, events which are recorded at the observation point at $(\mathbf{r},t)$ are due to time variations in the source at **r**' at a time $t'_{ret} \cong t - r/c + \hat{\mathbf{n}}\cdot\mathbf{r}'/c$, where t'_{ret} depends on **r**'. It is worth emphasizing this point.

Because of the finite propagation time from source to observer, at time *t* we are sampling what happened in parts of the source more distant from the observer at an earlier source time than parts of the source closer to the observer!

Thus to find $\phi(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$, we have to add up source variations which occur at different points in the source at *different* retarded times. Our basic assumption here is that the sources vary slowly enough in time that we can expand $\rho(\mathbf{r}',t'_{ret})$ as follows

$$\begin{aligned}\rho(\mathbf{r}',t'_{ret}) &= \rho(\mathbf{r}',t - r/c + \hat{\mathbf{n}}\cdot\mathbf{r}'/c + ...)\\ \rho(\mathbf{r}',t'_{ret}) &\cong \rho(\mathbf{r}',t - r/c) + \frac{\hat{\mathbf{n}}\cdot\mathbf{r}'}{c}\frac{\partial}{\partial t'}\rho(\mathbf{r}',t - r/c) + ...\end{aligned} \tag{8.2.6}$$

This is just a Taylor series expansion about $t - r/c$, where r/c is the propagation time from the center of the source region to the observer. Such an expansion will be good as long as the first term on the right hand side of equation (8.2.6) is much larger than the second, i.e.,

$$\frac{\hat{\mathbf{n}}\cdot\mathbf{r}'}{c}\frac{1}{\rho}\frac{\partial\rho}{\partial t'} << 1 \tag{8.2.7}$$

But $\frac{1}{\rho}\frac{\partial\rho}{\partial t'}$ is *1/T*, where *T* is a characteristic time for significant variation in ρ. Since |**r**'| is less than *d*, the maximum extent of our localized source region, for our approximation in (8.2.6) to be valid, we need

$$\frac{d}{c} << T \qquad \textit{Electric Dipole Approximation} \tag{8.2.8}$$

This approximation is known as the *electric dipole approximation*. For expansion (8.2.6) to hold, we must require (8.2.8) to hold, which says that the time required for light to propagate across our source must be small compared to characteristic times for significant variation in the source. Thus if we assume that

$$r >> d \quad \text{and} \quad \frac{d}{c} << T \tag{8.2.9}$$

then using (8.2.4) and (8.2.6), we can expand our exact solutions (8.2.1) to first order in the small quantities *d/r* and *d/cT* as

$$\phi(\mathbf{r},t)=\frac{1}{4\pi\ \varepsilon_o}\int d^3x'\left[\rho(\mathbf{r}',t')+\frac{\hat{\mathbf{n}}\cdot\mathbf{r}'}{c}\frac{\partial}{\partial t'}\rho(\mathbf{r}',t')+....\right]\left[\frac{1}{r}\left(1+\frac{\hat{\mathbf{n}}\cdot\mathbf{r}'}{r}+....\right)\right] \quad (8.2.10)$$

$$\mathbf{A}(\mathbf{r},t)=\frac{\mu_o}{4\pi\ r}\int d^3x'\mathbf{J}(\mathbf{r}',t')+\frac{\mu_o}{4\pi\ r}\int\left(\hat{\mathbf{n}}\cdot\mathbf{r}'\right)\left[\frac{\mathbf{J}(\mathbf{r}',t')}{r}+\frac{1}{c}\frac{\partial}{\partial t'}\mathbf{J}(\mathbf{r}',t')\right]d^3x'+.... \quad (8.2.11)$$

where

$$t'=t-r/c \quad (8.2.12)$$

We could keep higher order terms in (8.2.10) and (8.2.11), but we will find that what information we need is contained in the orders we've kept. These terms will allow us to look at the properties of electric dipole radiation, magnetic dipole radiation, and electric quadrupole radiation. We define the electric monopole, electric dipole moment, magnetic dipole moment, and electric quadrupole moment of our sources by the equations

$$q(t')=\int\rho(\mathbf{r}',t')d^3x'=\mathrm{Q_o} \quad (8.2.13)$$

$$\mathbf{p}(t')=\int\mathbf{r}'\rho(\mathbf{r}',t')d^3x' \qquad \mathbf{m}(t')=\frac{1}{2}\int\mathbf{r}'\times\mathbf{J}(\mathbf{r}',t')d^3x' \quad (8.2.14)$$

$$\left[\ddot{\mathbf{Q}}(t')\right]_{ij}=\int\left[3x_i'x_j'-\left(r'\right)^2\delta_{ij}\right]\rho(\mathbf{r},t')d^3x' \quad (8.2.15)$$

These are exactly analogous to our definitions in the static case, except now these vectors quantities vary in time. Note that since charge is conserved and our sources are localized (in particular there is no current flow to infinity), the total charge $q(t')$ is constant in time and equal to Q_O.

8.3 Electric dipole radiation

We gather all terms in **A** and ϕ which are proportional to Q_{O}, **p** or $d\mathbf{p}/dt$. From (8.2.10) we have

$$\phi(\mathbf{r},t)\cong\frac{1}{4\pi\ \varepsilon_o}\frac{Q_o}{r}+\frac{1}{4\pi\ \varepsilon_o}\frac{1}{r^2}\hat{\mathbf{n}}\cdot\int\mathbf{r}'\rho(\mathbf{r}',t')\ d^3x'+\frac{1}{4\pi\ \varepsilon_o}\frac{1}{c\ r}\hat{\mathbf{n}}\cdot\int\mathbf{r}'\frac{\partial}{\partial t'}\rho(\mathbf{r}',t')\ d^3x' \quad (8.3.1)$$

Using the definition of $\mathbf{p}(t)$ in (8.2.14), we have

$$\phi(\mathbf{r},t) \cong \frac{1}{4\pi\,\varepsilon_o}\frac{Q_o}{r} + \frac{1}{4\pi\,\varepsilon_o}\frac{1}{r^2}\hat{\mathbf{n}}\cdot\mathbf{p}(t') + \frac{1}{4\pi\,\varepsilon_o}\frac{1}{c\,r}\hat{\mathbf{n}}\cdot\frac{d\mathbf{p}(t')}{dt} \qquad \text{where } t' = t - r/c \tag{8.3.2}$$

What about **A**? In Problem Set 3 you showed that

$$\int \mathbf{J}(\mathbf{r}',t')\,d^3x' = -\int \mathbf{r}'\left[\nabla'\cdot\mathbf{J}(\mathbf{r}',t')\right]d^3x' \tag{8.3.3}$$

and thus the first term in (8.2.11) becomes

$$\mathbf{A}(\mathbf{r},t) \cong -\frac{\mu_o}{4\pi}\frac{1}{r}\int d\tau'\;\mathbf{r}'(\nabla'\cdot\mathbf{J}(\mathbf{r}',t')) \tag{8.3.4}$$

Remember that charge conservation in differential form is

$$\nabla'\cdot\mathbf{J}(\mathbf{r}',t') + \frac{\partial}{\partial t'}\rho(\mathbf{r}',t') = 0 \tag{8.3.5}$$

so that the term in **A** proportional to $d\mathbf{p}/dt'$ is

$$\mathbf{A}(\mathbf{r},t) \cong +\frac{\mu_o}{4\pi}\frac{1}{r}\frac{d\mathbf{p}(t')}{dt'} \qquad t' = t - r/c \tag{8.3.6}$$

The other terms in (8.2.11) are proportional to electric quadrupole or magnetic dipole moments, as discussed later, so we for the moment ignore those terms and concentrate only on the terms that involve $\mathbf{p}(t')$ and its time derivatives. We define $\dot{\mathbf{p}}(t')$ to be $d\mathbf{p}(t')/dt'$, $\ddot{\mathbf{p}}(t')$ to be $d^2\mathbf{p}(t')/dt'^2$, and so on. Since $\mathbf{B} = \nabla\mathbf{x}\mathbf{A}$ we have using (6.1.1) that

$$\mathbf{B}(\mathbf{r},t) = \nabla\times\left[\frac{\mu_o}{4\pi}\frac{1}{r}\dot{\mathbf{p}}(t')\right] = \frac{\mu_o}{4\pi}\frac{1}{r}\nabla\times\dot{\mathbf{p}}(t') - \frac{\mu_o}{4\pi}\dot{\mathbf{p}}(t')\times\nabla\left[\frac{1}{r}\right] \tag{8.3.7}$$

Obviously, $\nabla\frac{1}{r} = -\frac{1}{r^2}\hat{\mathbf{n}}$, and for functions of *t-r/c*,

$$\nabla g(t - r/c) = -\frac{\hat{\mathbf{n}}}{c}\;\frac{dg(t')}{dt'} \tag{8.3.8}$$

so that

$$\nabla\times\dot{\mathbf{p}}(t') = -\frac{\hat{\mathbf{n}}}{c}\times\ddot{\mathbf{p}}(t') \tag{8.3.9}$$

and thus

$$\mathbf{B}(\mathbf{r},t) = \frac{\mu_o}{4\pi}\left[\frac{\dot{\mathbf{p}}}{r^2} + \frac{\ddot{\mathbf{p}}}{c\,r}\right]\times\hat{\mathbf{n}} \quad \text{evaluated at } t' = t - r/c \tag{8.3.10}$$

What about **E**? Well $\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$ so from (8.3.2) and (8.3.6)

$$\mathbf{E}(\mathbf{r},t) = -\frac{1}{4\pi\,\varepsilon_o}\nabla\left[\frac{Q_o}{r} + \frac{1}{r^2}\hat{\mathbf{n}}\cdot\mathbf{p}(t') + \frac{1}{c\,r}\hat{\mathbf{n}}\cdot\dot{\mathbf{p}}(t')\right] - \frac{\mu_o}{4\pi}\frac{1}{r}\ddot{\mathbf{p}}(t') \tag{8.3.11}$$

We can derive expressions for $\nabla\left[\frac{\hat{\mathbf{n}}\cdot\mathbf{p}(t')}{r^2}\right]$ and $\nabla\left[\frac{\hat{\mathbf{n}}\cdot\dot{\mathbf{p}}(t')}{r}\right]$ as follows

$$\nabla\frac{\hat{\mathbf{n}}\cdot\mathbf{p}(t')}{r^2} = \frac{1}{r^3}\left[\mathbf{p} - 3(\mathbf{p}\cdot\hat{\mathbf{n}})\hat{\mathbf{n}}\right] - \frac{(\hat{\mathbf{n}}\cdot\dot{\mathbf{p}})}{cr^2}\hat{\mathbf{n}} \tag{8.3.12}$$

$$\nabla\frac{\hat{\mathbf{n}}\cdot\dot{\mathbf{p}}(t')}{r} = \frac{1}{r^2}\left[\dot{\mathbf{p}} - 2(\dot{\mathbf{p}}\cdot\hat{\mathbf{n}})\hat{\mathbf{n}}\right] - \frac{(\hat{\mathbf{n}}\cdot\ddot{\mathbf{p}})}{cr}\hat{\mathbf{n}} \tag{8.3.13}$$

Inserting these into equation (8.3.11) gives us

$$\mathbf{E}(\mathbf{r},t) = \underbrace{\frac{1}{4\pi\,\varepsilon_o}\frac{Q_o}{r^2}\hat{\mathbf{n}} + \frac{1}{r^3}\frac{\left[3\hat{\mathbf{n}}(\mathbf{p}\cdot\hat{\mathbf{n}}) - \mathbf{p}\right]}{4\pi\,\varepsilon_o}}_{\text{quasi-static}} + \underbrace{\frac{1}{c\,r^2}\frac{\left[3\hat{\mathbf{n}}(\dot{\mathbf{p}}\cdot\hat{\mathbf{n}}) - \dot{\mathbf{p}}\right]}{4\pi\,\varepsilon_o}}_{\text{induction}} + \underbrace{\frac{1}{rc^2}\frac{(\ddot{\mathbf{p}}\times\hat{\mathbf{n}})\times\hat{\mathbf{n}}}{4\pi\,\varepsilon_o}}_{\text{radiation}} \tag{8.3.14}$$

evaluated at $t' = t - r/c$. In the limit of no time variation, **E** is just a static monopole field and dipole field, and **B** is zero. If **p** is varying slowly in time, the static dipole field becomes quasi-static (slowly varying in time, but with the same mathematical form). In addition, there are terms proportional to $\dot{\mathbf{p}}/cr^2$ (the *induction* terms) and terms proportional to $\ddot{\mathbf{p}}/c^2r$ (the *radiation* terms).

9 Examples of electric dipole radiation

9.1 Learning objectives

We look at the solutions for electric dipole radiation for specific cases. We consider two different kinds of time behavior for the electric dipole moment, and define the near, intermediate, and far zones.

9.2 Dipole moment vector p varying in magnitude but not direction

We want to look at some specific cases so that we can understand what equations (8.3.10) and (8.3.14) mean. To do this we first assume that we have a dipole moment vector $\mathbf{p}$ that is *always in the same direction* but with a time varying magnitude $p(t)$, that is

$$\mathbf{p}(t) = \hat{\mathbf{z}}\, p(t) \tag{9.2.1}$$

If we insert (9.2.1) into the expressions for **E** and **B** above, we find that

$$\mathbf{E}(\mathbf{r},t) = \frac{p}{r^3}\frac{\left[3\hat{\mathbf{n}}(\hat{\mathbf{z}}\cdot\hat{\mathbf{n}}) - \hat{\mathbf{z}}\right]}{4\pi\,\varepsilon_o} + \frac{\dot{p}}{c\,r^2}\frac{\left[3\hat{\mathbf{n}}(\hat{\mathbf{z}}\cdot\hat{\mathbf{n}}) - \hat{\mathbf{z}}\right]}{4\pi\,\varepsilon_o} + \frac{\ddot{p}}{rc^2}\frac{(\hat{\mathbf{z}}\,\mathbf{x}\,\hat{\mathbf{n}})\,\mathbf{x}\,\hat{\mathbf{n}}}{4\pi\,\varepsilon_o} \tag{9.2.2}$$

$$\mathbf{B}(\mathbf{r},t) = \frac{\mu_o}{4\pi}\left[\frac{\dot{p}}{r^2} + \frac{\ddot{p}}{c\,r}\right]\hat{\mathbf{z}}\,\mathbf{x}\,\hat{\mathbf{n}} \tag{9.2.3}$$

Remember that $\hat{\mathbf{n}}$ is just $\hat{\mathbf{r}}$, pointing from the source located at the near the origin to the observer far from the origin. Let us specify the direction of $\hat{\mathbf{n}} = \mathbf{r}/r$ in spherical polar coordinates by the polar angle θ and the azimuth angle φ, as shown in Figure 9-1.

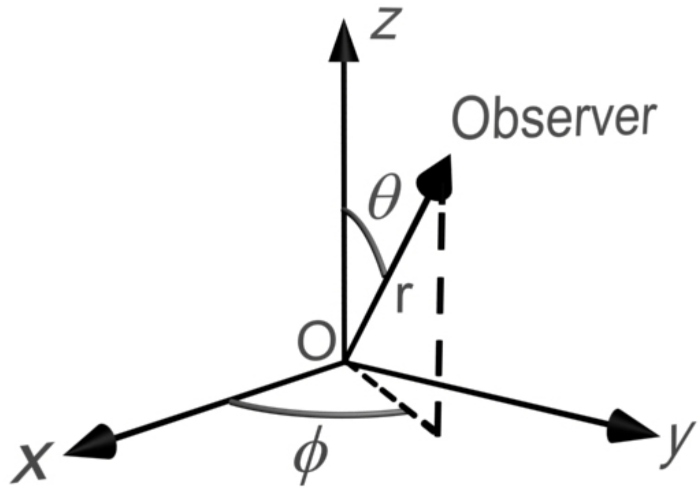


Figure 9-1: The vector to the observer

Thus we have

$$\hat{\mathbf{z}}\cdot\hat{\mathbf{n}} = \cos\theta \qquad \hat{\mathbf{z}}\,\mathbf{x}\,\hat{\mathbf{n}} = \hat{\boldsymbol{\varphi}}\sin\theta \qquad (\hat{\mathbf{z}}\,\mathbf{x}\,\hat{\mathbf{n}})\,\mathbf{x}\,\hat{\mathbf{n}} = \hat{\boldsymbol{\theta}}\sin\theta \tag{9.2.4}$$

So (9.2.2) and (9.2.3) become

$$\mathbf{E}(\mathbf{r},t) = \frac{2\cos\theta}{4\pi\,\varepsilon_o}\hat{\mathbf{r}}\left(\frac{p}{r^3} + \frac{\dot{p}}{cr^2}\right) + \frac{\sin\theta}{4\pi\,\varepsilon_o}\hat{\boldsymbol{\theta}}\left(\frac{p}{r^3} + \frac{\dot{p}}{cr^2} + \frac{\ddot{p}}{c^2 r}\right) \tag{9.2.5}$$

$$\mathbf{B}(\mathbf{r},t) = \frac{\mu_o \sin\theta}{4\pi}\left(\frac{\dot{p}}{r^2} + \frac{\ddot{p}}{c\,r}\right)\hat{} \qquad (9.2.6)$$

9.3 The near, intermediate, and far zones

If I ask which terms are the dominate terms in equation (9.2.5), I must ask about the magnitude of r compared to cT, where T is a characteristic time scale for variations in the sources. To see this, we re-write (9.2.5) as follows in order of magnitude

$$E \approx \frac{1}{4\pi\,\varepsilon_o}\left(\frac{p}{r^3} + \frac{\dot{p}}{cr^2} + \frac{\ddot{p}}{c^2 r}\right) \approx \frac{1}{4\pi\,\varepsilon_o}\frac{p}{r^3}\left(1 + \frac{r\,\dot{p}}{cp} + \frac{r^2\ddot{p}}{c^2 p}\right) \qquad (9.3.1)$$

I now set $\dot{p} \approx p/T$ and $\ddot{p} \approx p/T^2$, so that (9.3.1) becomes

$$E \approx \frac{1}{4\pi\,\varepsilon_o}\frac{p}{r^3}\left(1 + \frac{r}{cT} + \left[\frac{r}{cT}\right]^2\right) \qquad (9.3.2)$$

There are three different possibilities:

(I) $r << cT$ **STATIC OR NEAR ZONE.** The dominant term in (9.3.2) is the first one, which represents the quasi-static dipole electric fields, varying in time, but in essence just a dipole field with E falling off as $1/r^3$.

(II) $r >> cT$ **RADIATION OR FAR ZONE**. The dominant term in (9.3.2) the last one, and it falls off as *1/r*, and these are the radiation fields. As we will see below, these terms carry energy to infinity, that is energy that is lost irreversibly and cannot be recovered.

(III) $r \approx cT$ **INDUCTION OR INTERMEDIATE ZONE**, all of the terms in (9.3.2) are of comparable importance.

9.4 Examples of electric dipole radiation in the near, intermediate, and far zones

In the sections below, I look at a examples of the time behavior of *p(t)*, each chosen to illustrate various features of (9.2.5) and (9.2.6). The examples can all be grouped into time behaviors of two general types. The first type of behavior is a sinusoidal time dependence, that is

$$p(t) = p_0 + p_1 \cos\omega t \qquad (9.4.1)$$

The second type of time behavior is an electric dipole which has been constant at one value of the dipole moment, p_0, up to time $t = 0$, and then smoothly transitions to another value of the dipole moment, p_1, over a time T. In this type of behavior, I take the time dependence of $p(t)$ to be

$$p(t) = \begin{cases} p_o & \text{for } t < 0 \\ p_0 + p_1\left[6\left(\frac{t}{T}\right)^5 - 15\left(\frac{t}{T}\right)^4 + 10\left(\frac{t}{T}\right)^3\right] & \text{for } 0 < t < T \\ p_0 + p_1 & \text{for } t > T \end{cases} \tag{9.4.2}$$

The time dependence of this function and its first and second time derivatives is shown in Figure 9-2. In the case plotted the dipole moment increases by 25% from its initial value. For clarity, so that they are more easily seen, I have multiplied the first and second derivatives of the dipole moment as a function of time by a factor of 10 in Figure 9-2.

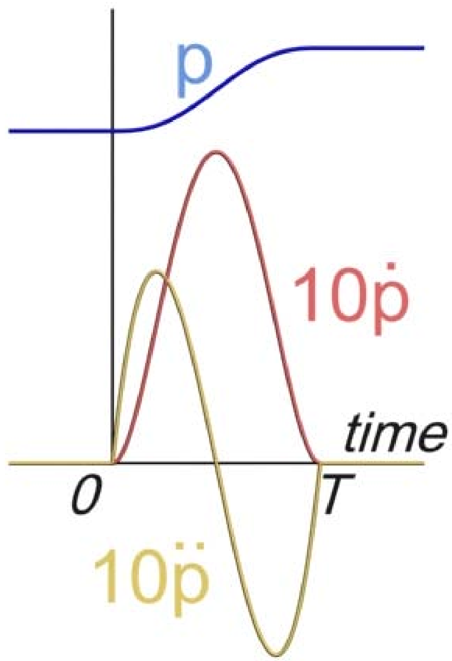


Figure 9-2: The time dependence of a dipole changing in time *T*

In each of the examples I show below, I will also show movies in class that animate the radiation time sequence at the three different scales for that example. The field lines and texture patterns move in these movies, with a velocity at each point in time given by

$$\mathbf{V}_{\text{field line}}(\mathbf{r},t) = c^2 \frac{\mathbf{E}(\mathbf{r},t)\times\mathbf{B}(\mathbf{r},t)}{E^2(\mathbf{r},t)} \tag{9.4.3}$$

I will justify the use of this velocity for a moving electric field line later on in the course. For the moment, the only thing you need to know is that this velocity is in the direction of the local value of $\mathbf{E}(\mathbf{r},t)\times\mathbf{B}(\mathbf{r},t)$ at every point in space and time, and that this vector represents the direction of electromagnetic energy flow, as we discuss at length soon. So motion in these movies represents direction of electromagnetic energy flow, and you should view these movies in this light.

9.4.1 Dipole moment varying sinusoidally with total reversal, in the near zone

In this case we set p_0 to zero in equation (9.4.1). Figure 9-3 shows the field line configuration at a time near the maximum value of the dipole moment, in the near zone. Successive figures show this pattern at different phases in the dipole cycle.

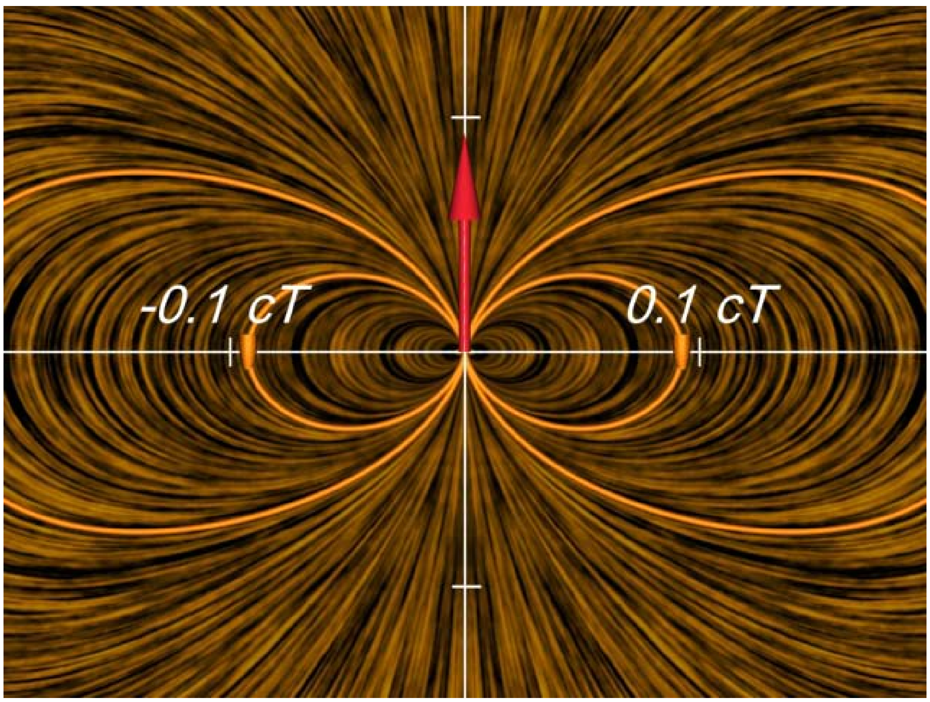


Figure 9-3: An oscillating dipole at the maximum of the dipole moment, in the near zone.

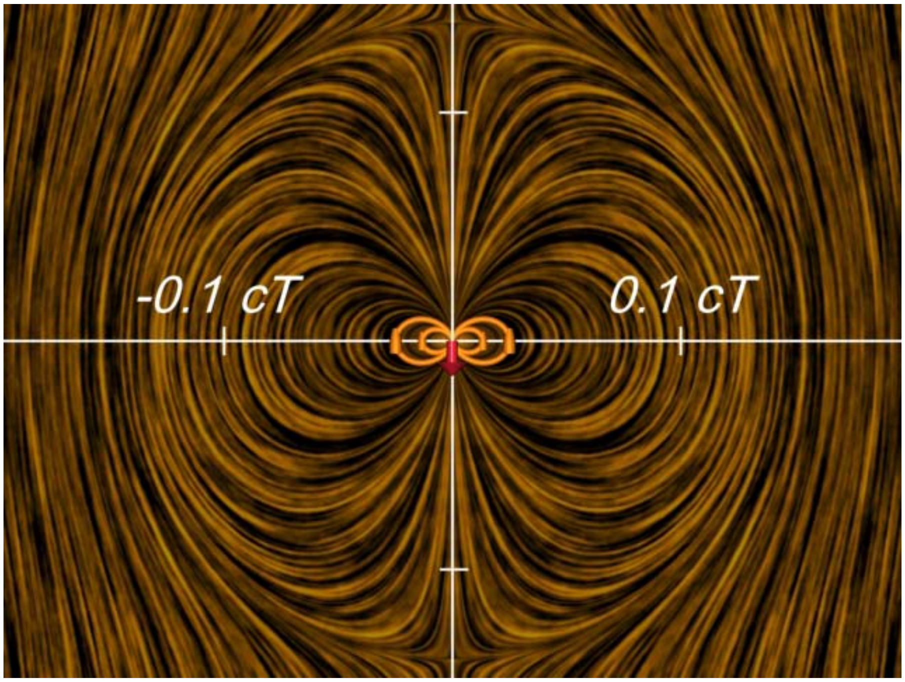


Figure 9-4: An oscillating dipole just after of the dipole moment has reversed, in the near zone.

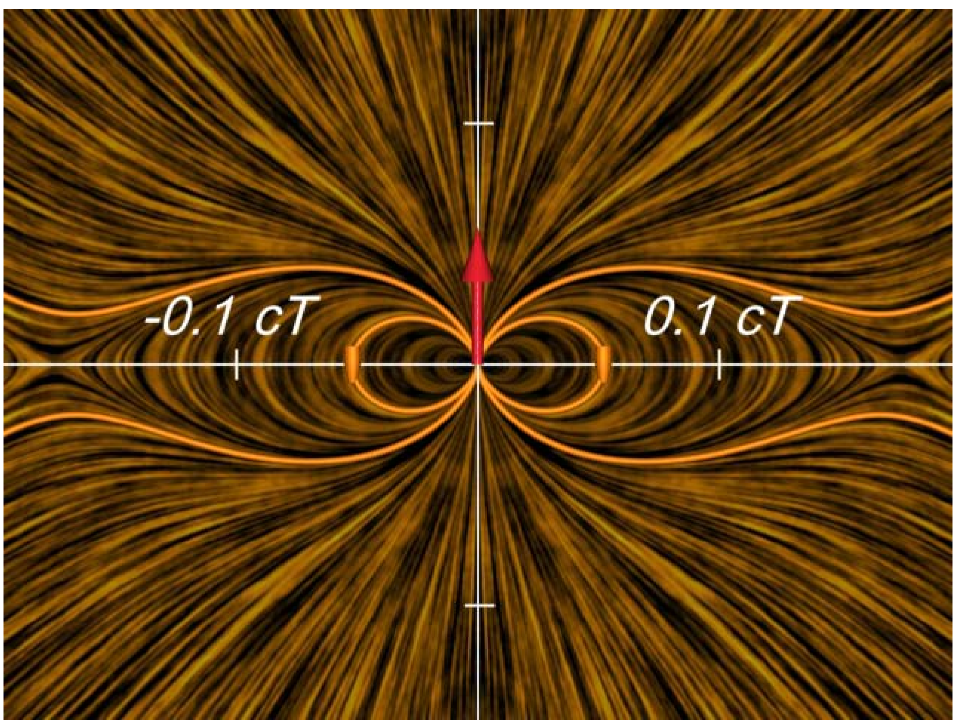


Figure 9-5: An oscillating dipole as the dipole moment approaches its maximum, in the near zone.

9.4.2 Dipole moment with total reversal, in the intermediate zone

Now we look at exactly the same thing as above, but we include distance further from the origin. The pattern changes qualitatively. The figures below show various phases.

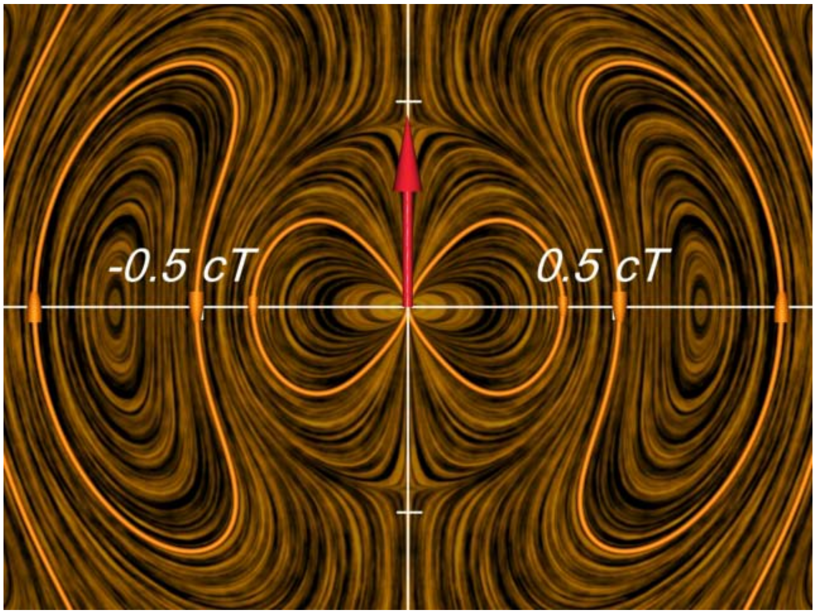


Figure 9-6: An oscillating dipole at the maximum of the dipole moment, in the intermediate zone

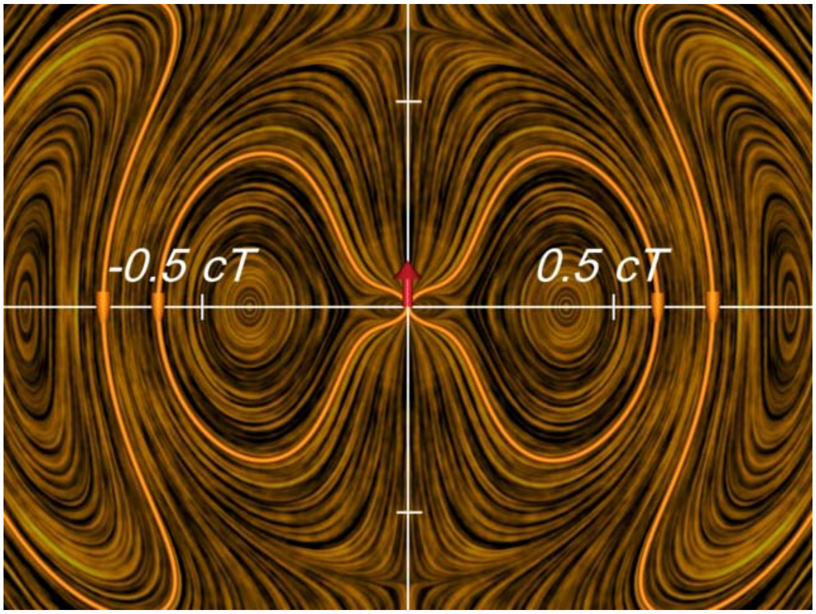


Figure 9-7: An oscillating dipole just before the dipole moment reverses, in the intermediate zone

9.4.3 Dipole moment varying sinusoidally with total reversal, in the far zone

Now we show the pattern even further out, and we see the characteristic electric dipole radiation pattern.

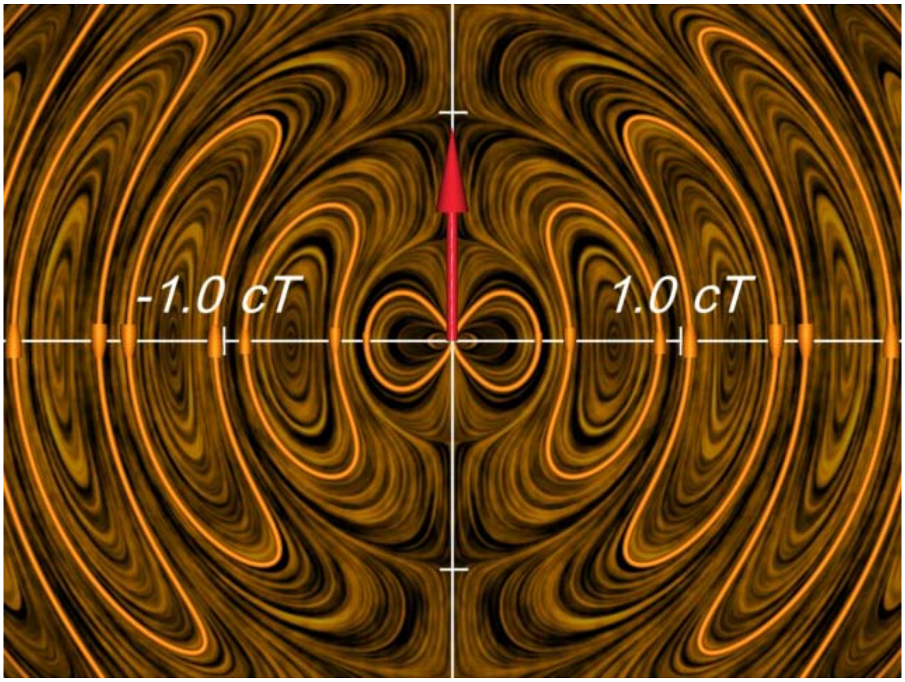


Figure 9-8: An oscillating dipole at the maximum of the dipole moment, in the far zone

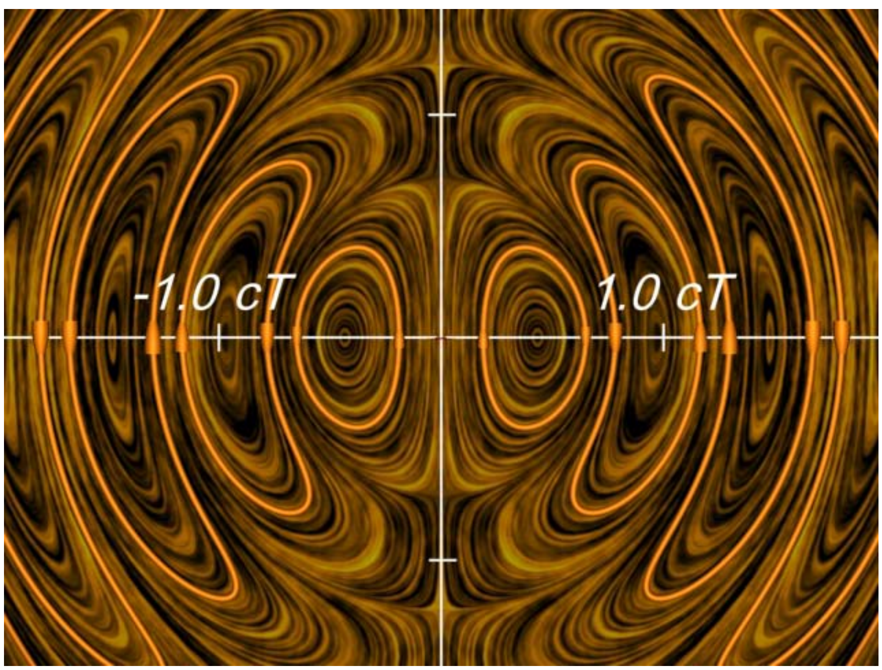


Figure 9-9: An oscillating dipole at the zero of the dipole moment, in the far zone

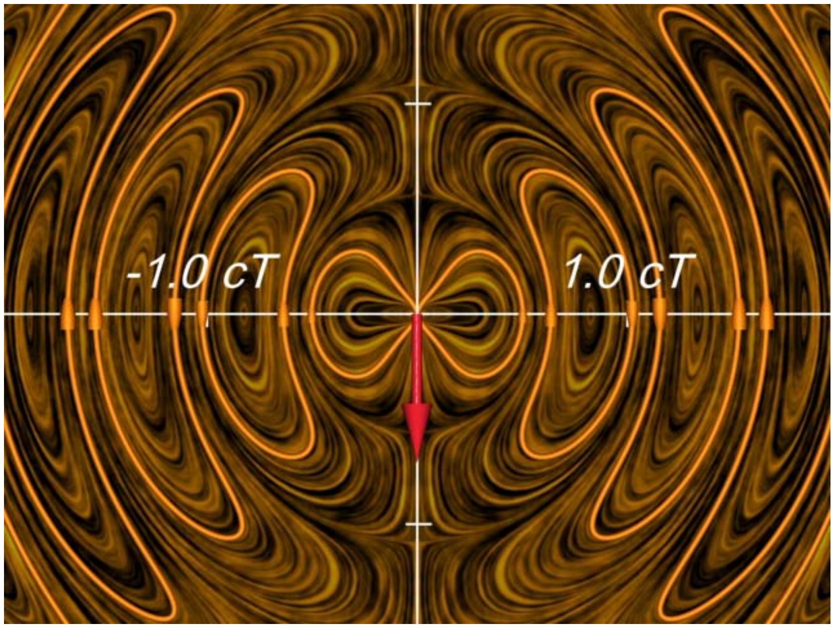


Figure 9-10: An oscillating dipole as the dipole moment approaches its downward maximum, in the far zone

9.4.4 Dipole moment increasing 50% over time *T*, in the near zone

Now I look at the behavior described by equation (9.4.2), in the case where the dipole magnitude increases by 50% over time T and then remains constant, first in the near zone.

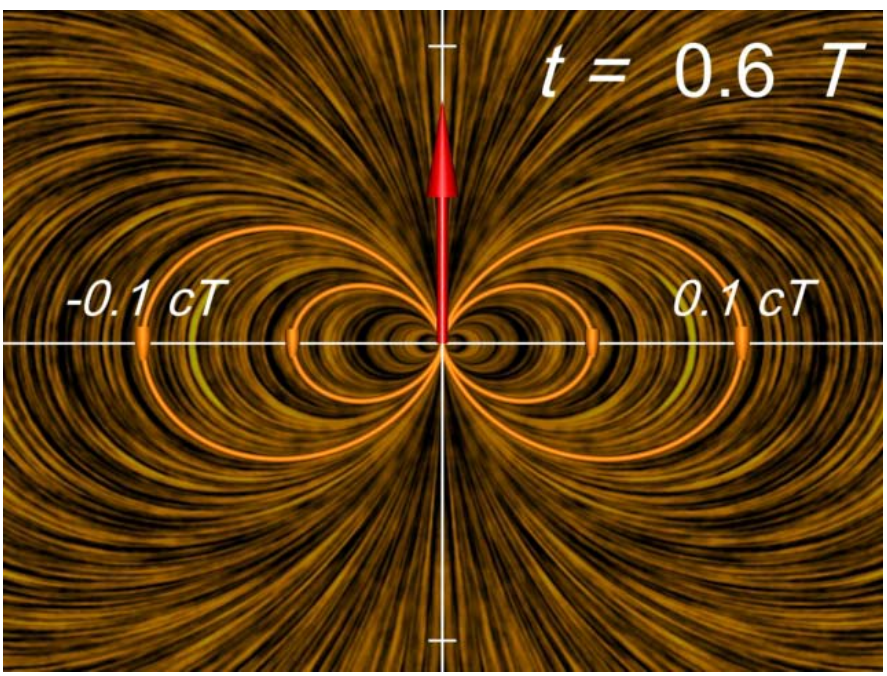


Figure 9-11: Dipole increasing over time *T* in the near zone

9.4.5 Dipole moment increasing over time *T*, in the intermediate zone

Now we look at the pattern in the intermediate zone.

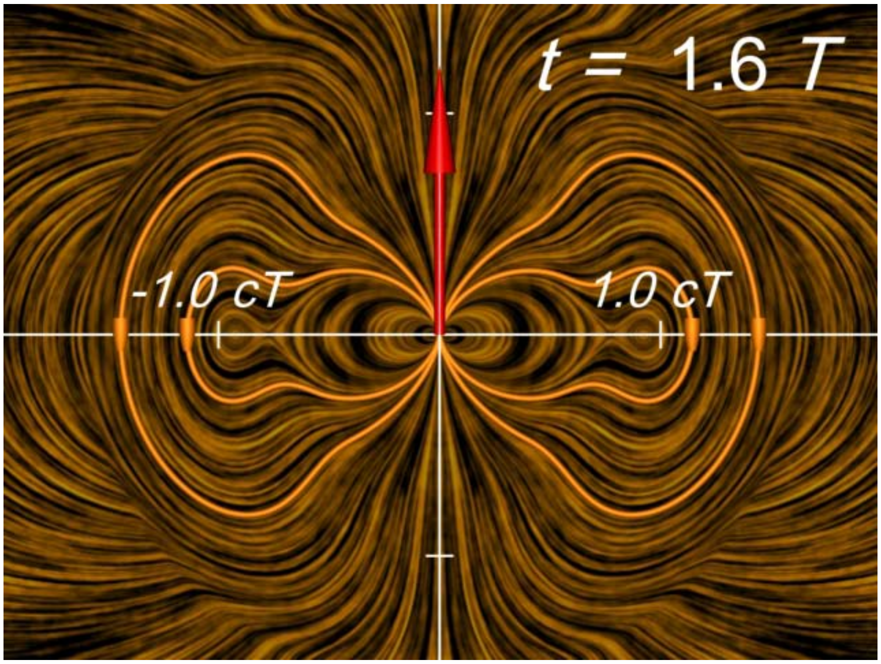


Figure 9-12: Dipole increasing over time *T* in the intermediate zone

9.4.6 Dipole moment decreasing by 33% over time *T*, in the near zone

Now we repeat the same sequence as above, except for the case that the dipole is decreasing in magnitude by 33% over time *T*.

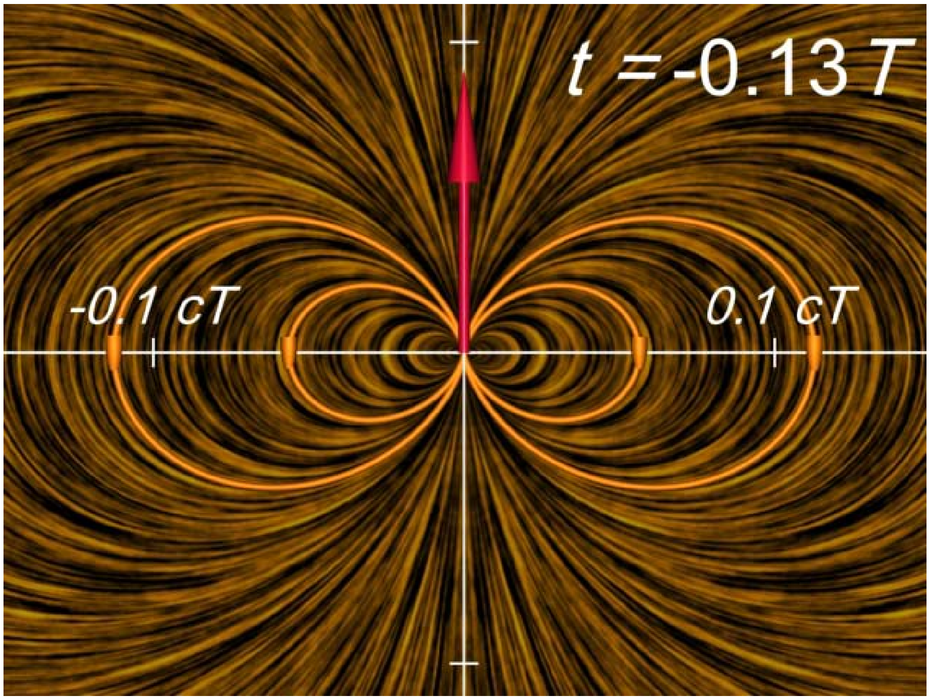


Figure 9-13: Dipole decreasing over time *T* in the near zone

9.4.7 Dipole moment decreasing over time *T*, in the intermediate zone

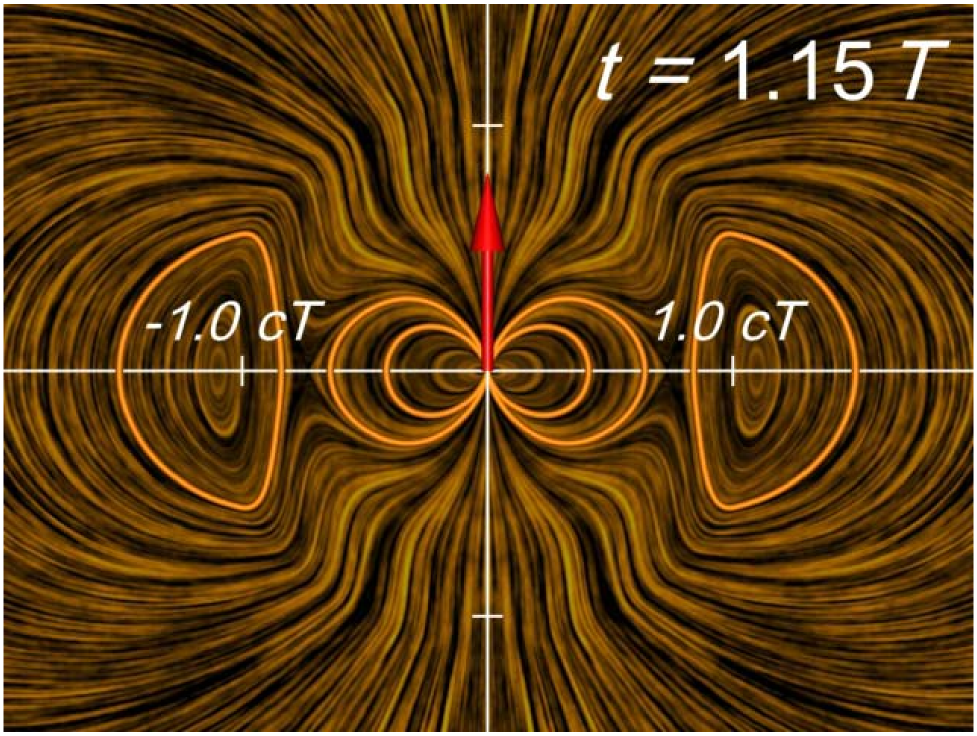


Figure 9-14: Dipole decreasing over time *T* in the intermediate zone.

9.5 Conservation of energy

Let us consider how the conservation of energy applies to our solutions above. First of all, our solutions above are only good for $r > d$, that is we are outside the source region containing charges and currents. Therefore in the region in which they apply, there are no sources or sinks of electromagnetic energy. We simply have energy flowing to fill up space with the local energy density of the electromagnetic field.

Second, we can calculate an expression for the flux of electromagnetic energy, that is rate at which total energy flows across a sphere of radius R_o per second. To get some feel for this, I consider the second type of time dependence for the dipole moment I considered above, when the dipole moment starts out at one value of the dipole moment, say p_1, and changes over a time T to another value of the dipole moment, say p_2. If we want to calculate the total amount of energy that has moved across a sphere of radius R_o in this process, we simply calculate the area integral over the surface of the sphere as follows:.

$$\text{Energy through } \mathrm{R}_o = \int_{-\infty}^{\infty} dt \int \frac{\mathbf{E}\times\mathbf{B}}{\mu_o}\cdot\hat{\mathbf{n}}\, da \tag{9.5.1}$$

If I refer to equations (9.2.5) and (9.2.6), I see that

$$\begin{aligned}\frac{\mathbf{E}\times\mathbf{B}}{\mu_o} &= \frac{\left(E_r\hat{\mathbf{r}} + E_\theta\hat{\boldsymbol{\theta}}\right)\times B_\varphi\hat{\boldsymbol{\varphi}}}{\mu_o} = \frac{\left(E_\theta B_\varphi\hat{\mathbf{r}} - E_r B_\varphi\hat{\boldsymbol{\theta}}\right)}{\mu_o} \\ &= \left[\hat{\mathbf{r}}\frac{\sin^2\theta}{(4\pi)^2\,\varepsilon_o}\left(\frac{\dot{p}}{r^2}+\frac{\ddot{p}}{c\,r}\right)\left(\frac{p}{r^3}+\frac{\dot{p}}{cr^2}+\frac{\ddot{p}}{c^2 r}\right) - \hat{\boldsymbol{\theta}}\frac{2\sin\theta\cos\theta}{(4\pi)^2\,\varepsilon_o}\left(\frac{p}{r^3}+\frac{\dot{p}}{cr^2}\right)\left(\frac{\dot{p}}{r^2}+\frac{\ddot{p}}{c\,r}\right)\right]\end{aligned} \tag{9.5.2}$$

Since we are considering a spherical surface of radius R_o, $\hat{\mathbf{n}} = \hat{\mathbf{r}}$, and we have

$$\text{Energy through } \mathrm{R}_o = \int_{-\infty}^{\infty} dt \int \frac{\mathbf{E}\times\mathbf{B}}{\mu_o}\cdot\hat{\mathbf{r}}\, r^2 d\Omega = \int_{-\infty}^{\infty} dt \int \frac{(\mathbf{E}\times\mathbf{B})_r}{\mu_o} r^2 d\Omega \tag{9.5.3}$$

With a little work (which you will do on a problem on Problem Set 4) this can be shown to be

$$\text{Energy through } \mathrm{R}_o = \frac{1}{12}\frac{\left(p_2^2 - p_1^2\right)}{4\pi\,\varepsilon_o R_o^3} + \int_{-\infty}^{\infty}\frac{\ddot{p}^2}{6\pi\,\varepsilon_o c^3}dt \tag{9.5.4}$$

Note that the second term on the right side of this equation is independent of R_o. This term represents the energy radiated away to infinity, and this is an irreversible process. We can easily see that the instantaneous rate at which energy is radiated away is

$$\text{Power in radiation (joules per sec)} = \frac{\ddot{p}^2}{6\pi\,\varepsilon_o c^3} \tag{9.5.5}$$

and this is known as Larmor's formula. What does the first term in (9.5.4) represent? Let's calculate the total amount of energy in an electrostatic dipole outside of a sphere of radius R_o. Using (7.4.1), you will show the following. The electrostatic energy of an electric dipole in the volume external to a sphere of radius R_o is given by

$$\text{Electrostatic energy of dipole outside } \mathrm{R}_o = \frac{1}{12}\frac{p^2}{4\pi\,\varepsilon_o R_o^3} \tag{9.5.6}$$

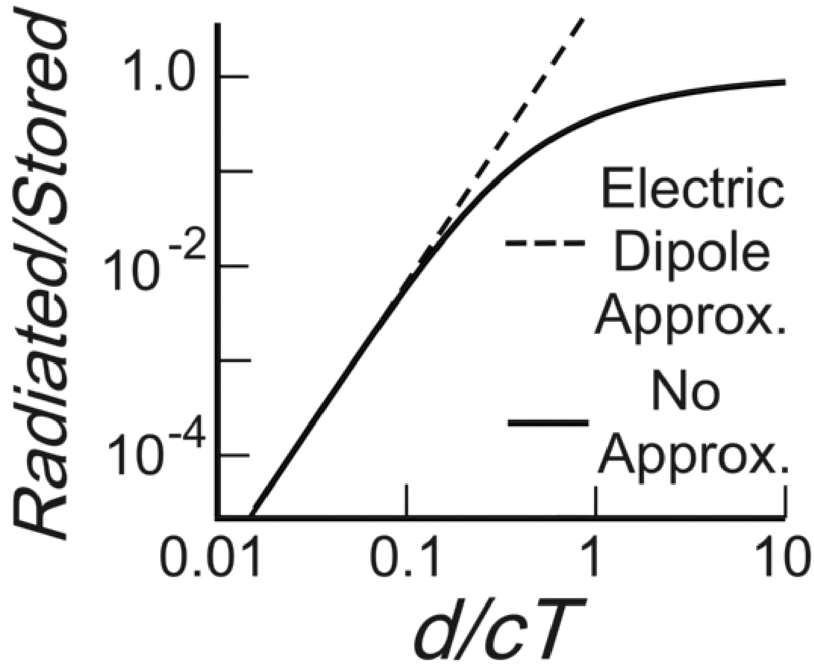
and thus we see that the first term in (9.5.4) represents the electrostatic energy needed to change the field from a dipole moment of p_1 to a dipole moment of p_2. Note that this can be positive or negative, depending on the relative sizes of these dipole moments, so that energy flows either inward or outward, depending on whether these quasi-electrostatic fields are being destroyed are created.

We can also compare the energy radiated away to the energy stored in or taken out of the electrostatic field, by looking at the ratio second to the first term in (9.5.4). If we take R_o to be the size of our source region, *d*, then we find that

$$\frac{\text{Radiated energy}}{\text{Stored energy}} \approx \frac{\ddot{p}^2 T}{6\pi\,\varepsilon_o c^3} \Big/ \frac{1}{12}\frac{\left(p_2^2 - p_1^2\right)}{4\pi\,\varepsilon_o d^3} \approx \frac{8d^3\left(p^2/T^4\right)T}{c^3\left(p_2^2 - p_1^2\right)} \approx \left[\frac{d}{cT}\right]^3 \tag{9.5.7}$$

Since our entire derivation above assumes the electric dipole approximation, that is $d/cT \ll 1$, we see that the radiated energy is always a small fraction of the energy that is stored or taken out of the electrostatic energy. Thus the irreversible energy loss due to radiation is small compared to the reversible energy storage in the electrostatic field.

Our course our whole expansion scheme rested on assuming $d < cT$ so that although you might thank that Eq. (9.5.7) implies that as *d* become larger than *cT* the radiated energy would exceed the stored energy by an arbitrarily large amount, in fact Eq. (9.5.7) is not valid when *d* becomes comparable to *cT*, so we can conclude nothing about the ratio of radiated to stored energy in such a case. A much more difficult and much more complicated calculation that does not make the electric dipole approximating in fact shows that when *d* becomes comparable to *cT* or greater, the radiated and stored energy are in fact exactly the same. I do not do that calculation here but I show a graph of the result of that calculation in Figure 9-15. For small values of *d/cT* we recover the electric dipole approximation result given in (9.5.7), but as *d/cT* approaches or exceeds one the radiated energy just equals the stored energy.



9-15: Energy radiated to stored energy as a function of *d/cT*

10 The General Form of Radiation E and B Fields

10.1 Learning Objectives

We stop and consider the general form of **E** and **B** radiation fields

10.2 General expressions for radiation E and B fields

I look at the terms in (8.3.10) and (8.3.14) which are a pure radiation field, that is the terms which carry energy off to infinity. These terms are the *1/r* terms, and are given for electric dipole radiation by

$$\mathbf{B}_{el\ dip}(\mathbf{r},t) = \frac{\mu_o}{4\pi}\left[\frac{\ddot{\mathbf{p}}\times\hat{\mathbf{n}}}{c\,r}\right]$$
$$\mathbf{E}_{el\ dip}(\mathbf{r},t) = \frac{1}{4\pi\,\varepsilon_o}\frac{1}{rc^2}(\ddot{\mathbf{p}}\times\hat{\mathbf{n}})\times\hat{\mathbf{n}} = c\,\mathbf{B}_{el\ dip}\times\hat{\mathbf{n}} \tag{10.2.1}$$

Before proceeding, I pause for a moment and consider the generalization of the form of the expressions in equation (10.2.1). As a general rule, whenever I deal only with radiation fields (terms falling off as *1/r*) of any kind (electric dipole, magnetic dipole, electric quadrupole, etc.), if we ignore everything except terms falling off as *1/r* , we will always have

$$\mathbf{B}_{rad} = +\frac{1}{c}\dot{\mathbf{A}}_{rad}\ \mathbf{x}\ \hat{\mathbf{n}} \quad \text{and} \quad \mathbf{E}_{rad} = (\dot{\mathbf{A}}_{rad}\ \mathbf{x}\ \hat{\mathbf{n}})\ \mathbf{x}\ \hat{\mathbf{n}} = c\,\mathbf{B}_{rad}\ \mathbf{x}\ \hat{\mathbf{n}} \tag{10.2.2}$$

These equations follow from assuming that the radiation part of the vector potential $\mathbf{A}(\mathbf{r},t)$ in (8.2.1) is *1/r* times some function of $t'_{ret} = t - r/c$, and dropping everything but *1/r* terms after taking derivatives. With this approach, it is clear that $\mathbf{B} = \nabla\times\mathbf{A}$ leads to the expression for $\mathbf{B}_{rad}$ in equation (10.2.2). What about the expression for $\mathbf{E}_{rad}$ in (10.2.2)? We appeal to the fact that outside the sources, where $\mathbf{J} = 0$, we have $\nabla\times\mathbf{B} = \mu_o\,\varepsilon_o\frac{\partial}{\partial t}\mathbf{E} = \frac{1}{c^2}\frac{\partial}{\partial t}\mathbf{E}$. Using the expression in (10.2.2) for the radiation $\mathbf{B}_{rad}$ field, and again dropping non-radiation terms, $\nabla\times\mathbf{B} = \frac{1}{c^2}\frac{\partial}{\partial t}\mathbf{E}$ tells us that we must have

$$\frac{1}{c^2}\left[\ddot{\mathbf{A}}_{rad}\ \mathbf{x}\ \hat{\mathbf{n}}\right]\ \mathbf{x}\ \hat{\mathbf{n}} = \frac{1}{c^2}\frac{\partial}{\partial t}\mathbf{E}_{rad} \tag{10.2.3}$$

Integrating (10.2.3) with respect to time, we recover the expression for the radiation $\mathbf{E}_{rad}$ field in (10.2.2) . We see that these equations are in particular appropriate for electric dipole radiation by inserting equation (18) into (23) and comparing to (22). Note that

$\mathbf{E}_{rad}$, $\mathbf{B}_{rad}$ and $\hat{\mathbf{n}}$ are all mutually perpendicular, and $cB_{rad} = E_{rad}$. This later equality means that the energy density in the magnetic field is equal to the energy density in the electric field.

10.3 Radiation of energy and momentum in the general case

What is the energy flux radiated into unit solid angle for the general radiation fields given in (10.2.2). From (10.2.2), we have for the radiation fields that

$$\frac{1}{\mu_o}(\mathbf{E}_{rad} \times \mathbf{B}_{rad}) = \frac{1}{\mu_o}\left(c\left(\mathbf{B}_{rad} \times \hat{\mathbf{n}}\right) \times \mathbf{B}_{rad}\right) = \hat{\mathbf{n}}\frac{c}{\mu_o}B_{rad}^2 - \frac{c}{\mu_o}\left(\mathbf{B}_{rad} \cdot \hat{\mathbf{n}}\right)\mathbf{B}_{rad} \tag{10.3.1}$$

Since $\mathbf{B}_{rad}$ and $\hat{\mathbf{n}}$ are perpendicular for the radiation fields, we have in general that

$$\frac{1}{\mu_o}(\mathbf{E}_{rad} \times \mathbf{B}_{rad}) = \hat{\mathbf{n}}\frac{c}{\mu_o}B_{rad}^2 \tag{10.3.2}$$

Now, suppose we take a very large sphere of radius *r* centered at the origin, and consider a surface element $\hat{\mathbf{n}}\,da$ on that sphere at a point (r,θ,ϕ), with

$$\hat{\mathbf{n}}\,da = r^2 d\Omega\,\hat{\mathbf{r}} = r^2 \sin\theta\; d\theta\, d\phi\,\hat{\mathbf{r}} \tag{10.3.3}$$

In the context of the energy conservation law that we have developed, we know that the quantity $\frac{(\mathbf{E}\,\mathbf{x}\,\mathbf{B})}{\mu_o}\cdot\hat{\mathbf{n}}\,da\;dt$ represents the amount of electromagnetic energy dW in joules flowing through $\hat{\mathbf{n}}\,da$ in a time dt :

$$dW = \frac{1}{\mu_o}(\mathbf{E}\,\mathbf{x}\,\mathbf{B})\cdot\hat{\mathbf{n}}\,da\;dt = \frac{c}{\mu_o}(\mathbf{E}\,\mathbf{x}\,\mathbf{B})\cdot(r^2 d\Omega\,\hat{\mathbf{r}})\;dt \tag{10.3.4}$$

or

$$\frac{dW}{d\Omega\;\,dt} = \frac{c}{\mu_o}(\mathbf{E} \times \mathbf{B})\cdot(r^2\,\hat{\mathbf{r}}) \tag{10.3.5}$$

Equation (10.3.5) is a general expression, good for any *r* and for any fields (quasi-static, induction, or radiation fields). However, it is clear that if we consider only the energy per second radiated to infinity, we need only include terms in **E** and **B** which fall off as *1/r*, since terms which fall off faster than this in the expression (10.3.5) for *dW* will vanish as *r* goes to infinity, and therefore carry no energy to infinity. Thus the electromagnetic energy radiated to infinity per unit time per unit solid angle is given by

$$\frac{dW_{rad}}{d\Omega\, dt} = \frac{r^2(\mathbf{E}_{rad}\,\mathbf{x}\,\mathbf{B}_{rad})\cdot\hat{\mathbf{n}}}{\mu_o} = \frac{c\,r^2 B_{rad}^{\,2}}{\mu_o} = \frac{r^2\left|\dot{\mathbf{A}}_{rad}\,\mathbf{x}\,\hat{\mathbf{n}}\right|^2}{\mu_o\, c} \qquad (10.3.6)$$

where we have used equation (10.2.2) and (10.3.2) to obtain the various forms in (10.3.6)

In the context of the momentum conservation law that we have developed, we know that the quantity $-\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}\,da\,dt$ represents the amount of electromagnetic momentum $d\mathbf{P}_{rad}$ flowing through $\hat{\mathbf{n}}\,da$ in a time *dt*. As above, the momentum radiated per unit time per unit solid angle is thus

$$\frac{d\mathbf{P}_{rad}}{dt\, d\Omega} = -r^2\,\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}} \qquad (10.3.7)$$

Again, if we are only interested in the momentum radiated to infinity, we see that we need only consider terms in $\overleftrightarrow{\mathbf{T}}$ which fall off as $1/r^2$, since terms which fall off faster than this will vanish as *r* goes to infinity, and therefore carry no momentum to infinity. Since $\overleftrightarrow{\mathbf{T}}$ involves the square of the fields, we need only keep the radiation terms in calculating the momentum radiated to infinity.

11 Another Example of Electric Dipole Radiation

11.1 Moving non-relativistic point charges

I first point out one general expression for the amount of power radiated in electric dipole radiation. Using equation (10.2.1) and equation (10.3.6), I have for electric dipole radiation that

$$\frac{dW_{el\,dip}}{d\Omega\, dt} = \frac{c\,r^2}{\mu_o}\left[\frac{\mu_o}{4\pi}\frac{\ddot{\mathbf{p}}\,\mathbf{x}\,\hat{\mathbf{n}}}{c\,r}\right]^2 = \frac{\mu_o\,\ddot{p}^{\,2}}{(4\pi)^2 c}\sin^2\theta \qquad (11.1.1)$$

where θ is the angle between $\ddot{\mathbf{p}}$ and $\hat{\mathbf{n}}$. If I integrate this expression over solid angle, taking $\ddot{\mathbf{p}}$ to lie along the z-axis for convenience, I obtain the expression for the total energy per second radiated in electric dipole radiation,

$$\frac{dW_{el\,dip}}{dt} = \frac{\mu_o}{4\pi}\frac{2\left|\ddot{\mathbf{p}}\right|^2}{3c} = \frac{1}{4\pi\,\varepsilon_o}\frac{2\left|\ddot{\mathbf{p}}\right|^2}{3c^3} \quad \text{or} \quad \frac{dW_{el\,dip}}{dt} = \frac{1}{4\pi\,\varepsilon_o}\frac{2}{3}\frac{q^2a^2}{c^3} \qquad (11.1.2)$$

In the last form in equation (11.1.2), we have given an expression appropriate for the specific case where the radiation is due to a single point charge of charge *q* which is at **R**(t) at time t. It is clear that in such a situation, $\rho(\mathbf{r}',t') = q\,\delta^3(\mathbf{r}' - \mathbf{R}(t'))$, and that therefore, using $\mathbf{p}(t') = \int \mathbf{r}'\rho(\mathbf{r}',t')d^3x'$, that $\mathbf{p}(t) = q\ \mathbf{R}(t)$ and $\ddot{\mathbf{p}}(t) = q\ \ddot{\mathbf{R}}(t) = q\,\mathbf{a}(t)$, where **a**(t) is the acceleration of the particle at time t.

Thus for a single particle, the instantaneous rate at which it radiates electric dipole energy is proportional the square of its charge and the square of its instantaneous acceleration. Note that if the particle is moving a speed V, it will travel across a region of length $d = VT$ in time T. Thus if T is the time it takes for the speed V to increase significantly, our requirement that d/cT be small compared to one for our expansion to be valid becomes for particle motion the requirement that $d/cT = VT/cT = V/c$ be small compared to one, i.e. that the particle speed be non-relativistic. Indeed, the radiation patterns for relativistic particles look nothing like the simple dipole and quadrupole radiation patterns we will derive here.

We now look at a specific example of electric dipole radiation, by looking at the fields of two point charges. We emphasize, however, that the methods we develop here can be applied in much more general situations than just individual point charge motions. All we need to do to apply them is to compute the overall moments of the charge and current distributions, as in equations (8.2.13) through (8.2.15).

Consider the following time varying source functions. We have two point charges, one at rest at the origin, with charge $-q_o$, and one moving up and down on the z-axis, with charge $+q_o$, and with its position described by

$$\mathbf{r}'(t') = \hat{\mathbf{z}}\ R_o \cos\omega_o t' \qquad (11.1.3)$$

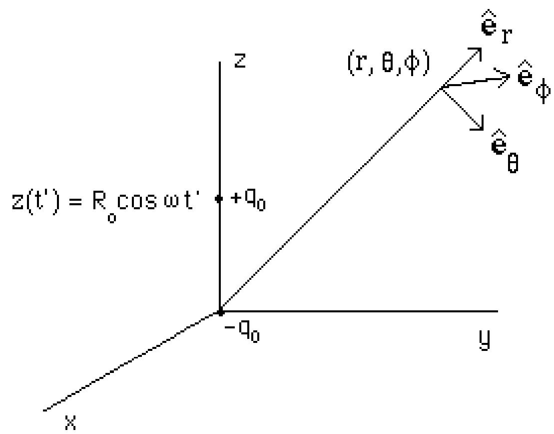


Figure 11-1: An electric dipole formed by moving point charges

We want to find the electric and magnetic fields appropriate to this source, including quasi-static, induction, and radiation fields. The charge density is given by the charge $\pm\, q_o$ times three dimensional delta functions at the positions of the charges. The charge density is thus

$$\rho(\mathbf{r}',t') = q_o\,\delta(x')\delta(y')\delta(z'-R_o\cos\omega_o t') - q_o\,\delta(x')\delta(y')\delta(z') \tag{11.1.4}$$

This charge distribution has no net charge. The electric dipole moment is (cf equation (8.2.14))

$$\begin{gathered}\mathbf{p}(t') = \int \mathbf{r}'\rho(\mathbf{r}',t')d^3x' = \\ q_o\int \mathbf{r}'\left[\delta(x')\delta(y')\delta(z'-R_o\cos\omega_o t') - \delta(x')\delta(y')\delta(z')\right]d^3x'\end{gathered} \tag{11.1.5}$$

or

$$\mathbf{p}(t') = q_o R_o \cos\omega_o t'\ \hat{\mathbf{z}} = p_o\cos\omega_o t'\ \hat{\mathbf{z}} \tag{11.1.6}$$

with $p_o = q_o R_o$ p_o. From (8.3.10) and (8.3.14), with this expression for **p**(t), and defining $k = \omega_o / c$, we have for **E** and **B** the expressions

$$\begin{aligned}\mathbf{B}(\mathbf{r},t) &= -\hat{\mathbf{e}}_\phi \frac{\mu_o}{4\pi}\frac{p_o c\,k^2\sin\theta}{r}\left[\cos\omega_o(t-r/c) + \frac{\sin\omega_o(t-r/c)}{kr}\right] \\ \mathbf{E}(\mathbf{r},t) &= \hat{\mathbf{e}}_r \frac{1}{4\pi\,\varepsilon_o}\frac{2p_o\cos\theta}{r^3}\left[\cos\omega_o(t-r/c) - kr\sin\omega_o(t-r/c)\right] \\ &+ \hat{\mathbf{e}}_\theta \frac{1}{4\pi\,\varepsilon_o}\frac{p_o\sin\theta}{r^3}\left[(1-k^2r^2)\cos\omega_o(t-r/c) - kr\sin\omega_o(t-r/c)\right]\end{aligned} \tag{11.1.7}$$

Terms in equation (11.1.7) like $\cos\omega_o(t-r/c)$ represent traveling waves moving away from origin with a frequency ω_o and a wave length $\lambda = 2\pi/k$, with period $T = 2\pi/\omega_o$. Note that our conditions in equation (8.2.9) are now $r >> d$ and $\lambda >> d$ (this requirement on λ is equivalent to the requirement that the maximum speed of the moving charge be small compared to the speed of light, as we saw above). Note however that we have made *no* requirement on r as compared to λ, only that both be much greater than d.

We now restrict ourselves to the radiation terms in equation (11.1.7), that is the terms that go to zero as $1/r$ as $r\to\infty$. In the limit that $kr >> 1$, the dominant terms in these equations are the radiation terms,

$$\begin{aligned}\mathbf{B}(\mathbf{r},t) &= -\hat{\mathbf{e}}_\phi \frac{\mu_o}{4\pi}\frac{c\ p_o\,k^2\sin\theta}{r}\cos\omega_o(t-r/c) \\ \mathbf{E}(\mathbf{r},t) &= -\hat{\mathbf{e}}_\theta \frac{\mu_o}{4\pi}\frac{c^2\ p_o\,k^2\sin\theta}{r}\cos\omega_o(t-r/c)\end{aligned} \tag{11.1.8}$$

where to get the form for the electric field in (11.1.8), we have used the fact that
$c^2 = \dfrac{1}{\mu_o\, \varepsilon_o}$

11.2 Energy and momentum flux

The energy radiated into a solid angle $d\Omega$ is just the Poynting flux into that solid angle, that is

$$\frac{dW_{rad}}{dt} = \left(\frac{\mathbf{E}_{rad}\, \mathbf{x}\, \mathbf{B}_{rad}}{\mu_o}\right)\cdot \hat{\mathbf{n}}\, r^2 d\Omega \tag{11.2.1}$$

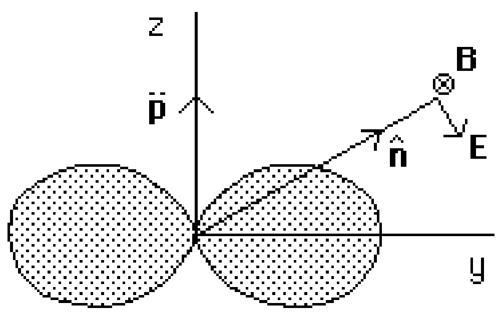
or

$$\frac{dW_{rad}}{d\Omega\, dt} = \frac{r^2(\mathbf{E}_{rad}\, \mathbf{x}\, \mathbf{B}_{rad})\cdot \hat{\mathbf{n}}}{\mu_o} \tag{11.2.2}$$

Using the radiation fields in (11.1.8) in (11.2.2), we have

$$\frac{dW_{rad}}{d\Omega\, dt} = +\frac{1}{(4\pi)^2 \varepsilon_o} c\, p_o^{\,2} k^4 \sin^2\theta\ \cos^2 \omega_o (t - r/c) \tag{11.2.3}$$

The radiation electric field **E** for electric dipole radiation is polarized in the plane of $\hat{\mathbf{n}}$ and **p**, and the radiation magnetic field **B** is out of that plane.



If we average over one period T and integrate over all solid angles, we find the total energy flux per second (in ergs/sec) is

$$\left\langle \frac{dW_{rad}}{dt} \right\rangle = \frac{c\, p_o^{\,2} k^4}{12\pi\, \varepsilon_o} = \frac{1}{4\pi\, \varepsilon_o} \frac{2}{3} \frac{q_o^{\,2} < a^2 >}{c^3} \tag{11.2.4}$$

where $<a^2>$ is equal to the value of the square of the acceleration averaged over one period (the average square of the acceleration is just one-half of the square of the peak acceleration). Compare equation (11.2.4) to equation (11.1.2) for the instantaneous rate at which energy is radiated.

So equation (11.2.4) gives the rate at which energy is radiated away. What about momentum? We use equation (10.3.7) We need to compute $-\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}$. The *j-th* component of $-\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}}$ is given by

$$(-\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}})_j = -T_{ji}n_i = -T_{jr} \quad \text{since } \hat{\mathbf{n}} = \hat{\mathbf{r}} \tag{11.2.5}$$

Since $B_r = 0$, we have

$$-\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{n}} = -\left[\varepsilon_o \mathbf{E}E_r - \frac{1}{2}\hat{\mathbf{e}}_r(\varepsilon_o E^2 + B^2/\mu_o)\right] \tag{11.2.6}$$

But $\mathbf{E}\,E_r$ is proportional to $1/r^3$, so we can drop this first term, and only keep radiation fields in the $\varepsilon_o E^2 + B^2/\mu_o$ term, so that equation(10.3.7) becomes

$$\frac{d\mathbf{P}_{rad}}{dt\,d\Omega} = +\frac{1}{(4\pi)^2\varepsilon_o}\hat{\mathbf{e}}_r\, k^4 p_o{}^2 \sin^2\theta\, \cos^2\omega_o(t - r/c) \tag{11.2.7}$$

This vector is radial and in magnitude is just $1/c$ times the energy per unit time passing through $\hat{\mathbf{n}}\,da$ (cf. equation (11.2.3)). A photon has energy $\hbar\omega$ and momentum $\hbar\omega/c$. We can time average $\frac{d\mathbf{P}_{rad}}{dt\,d\Omega}$ over one cycle, but if we integrate $\frac{d\mathbf{P}_{rad}}{dt\,d\Omega}$ over $d\Omega$ we get a net of zero (to do this must *first* express $\frac{d\mathbf{P}_{rad}}{dt\,d\Omega}$ in Cartesian components, and *then* integrate over $d\Omega$).

12 Magnetic Dipole and Electric Quadrupole Radiation

12.1 Learning Objectives

We now consider the properties of magnetic dipole radiation and electric quadrupole radiation.

12.2 Magnetic dipole radiation

In looking at electric dipole radiation, we have just scratched the surface of the radiation produced by time varying sources. Electric dipole radiation is the dominate mode of radiation, but if it vanishes there are other modes we now review. We only want to consider two other characteristic forms of radiation, which for $d/cT << 1$ turn out to be important only if the electric dipole moment vanishes. Consider the higher order terms in equation (8.2.11) for $\mathbf{A}(\mathbf{r},t)$:

$$\mathbf{A}(\mathbf{r},t) = \frac{\mu_o}{4\pi\, r^2}\int \left(\hat{\mathbf{n}}\cdot\mathbf{r}'\right)\mathbf{J}(\mathbf{r}',t')d^3x' + \frac{\mu_o}{4\pi\, r\, c}\int \left(\hat{\mathbf{n}}\cdot\mathbf{r}'\right)\frac{\partial}{\partial t'}\mathbf{J}(\mathbf{r}',t')\,d^3x' \qquad (12.2.1)$$

By using your results on Problem 3-1(c) of Problem Set 3, this can be written as

$$\mathbf{A}(\mathbf{r},t) = \frac{\mu_o}{8\pi\, r^2}\int\left[\left(\mathbf{r}'\,\mathbf{x}\,\mathbf{J}\right)\mathbf{x}\,\hat{\mathbf{n}}\right]d^3x' + \frac{\mu_o}{8\pi\, r\, c}\int\left[\left(\mathbf{r}'\,\mathbf{x}\,\frac{\partial}{\partial t'}\mathbf{J}\right)\mathbf{x}\,\hat{\mathbf{n}}\right]\, d^3x'$$
$$+\frac{\mu_o}{8\pi\, r^2}\int\left[\mathbf{r}'\,(\hat{\mathbf{n}}\cdot\mathbf{J}) + \mathbf{J}\,(\hat{\mathbf{n}}\cdot\mathbf{r}')\right]\, d^3x' + \frac{\mu_o}{8\pi\, r\, c}\int\left[\mathbf{r}'\,(\hat{\mathbf{n}}\cdot\frac{\partial}{\partial t'}\mathbf{J}) + \frac{\partial}{\partial t'}\mathbf{J}\,(\hat{\mathbf{n}}\cdot\mathbf{r}')\right]d^3x' \qquad (12.2.2)$$

Let us first treat the first two terms in (12.2.2), which will give us magnetic dipole radiation.. With the definition $\mathbf{m}(t') = \frac{1}{2}\int \mathbf{r}'\mathbf{x}\mathbf{J}(\mathbf{r}',t')d^3x'$ from (8.2.14), we have for the magnetic dipole part of **A**:

$$\mathbf{A}_{mag\ dip}(\mathbf{r},t) = \frac{\mu_o}{4\pi}\left[\frac{\mathbf{m}(t')}{r^2} + \frac{\dot{\mathbf{m}}(t')}{cr}\right]\mathbf{x}\,\hat{\mathbf{n}} \qquad (12.2.3)$$

The first term here is just static magnetic dipole vector potential. To get the full **B**, we must compute $\nabla\times\mathbf{A}$. This is messy, and we can avoid the work by noting that outside the source, we have from Maxwell's equations in vacuum that

$$\nabla\times\mathbf{B} = \frac{1}{c^2}\frac{\partial}{\partial t}\mathbf{E} \qquad \text{and} \qquad \nabla\times\mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B} \qquad (12.2.4)$$

In particular, the first equation on the left in (12.2.4) must be true for the electric dipole **B** and **E** given in (8.3.10) and (8.3.14) above. Thus

$$\nabla\times\frac{\mu_o}{4\pi}\left\{\left[\frac{\ddot{\mathbf{p}}}{c\,r} + \frac{\dot{\mathbf{p}}}{r^2}\right]\times\hat{\mathbf{n}}\right\} = \frac{1}{c^2}\frac{\partial}{\partial t}\left\{\frac{\left[3\hat{\mathbf{n}}(\mathbf{p}\cdot\hat{\mathbf{n}})-\mathbf{p}\right]}{4\pi\,\varepsilon_o r^3} + \frac{\left[3\hat{\mathbf{n}}(\dot{\mathbf{p}}\cdot\hat{\mathbf{n}})-\dot{\mathbf{p}}\right]}{4\pi\,\varepsilon_o c\, r^2} + \frac{(\ddot{\mathbf{p}}\,\mathbf{x}\,\hat{\mathbf{n}})\,\mathbf{x}\,\hat{\mathbf{n}}}{4\pi\,\varepsilon_o rc^2}\right\} \quad (12.2.5)$$

Since $\frac{\partial}{\partial t} = \frac{\partial}{\partial t'}$, we can integrate both sides of equation (12.2.5) with respect to t' to obtain

$$\nabla\,\mathbf{x}\left\{\frac{\mu_o}{4\pi}\left[\frac{\dot{\mathbf{p}}}{cr} + \frac{\mathbf{p}}{r^2}\right]\mathbf{x}\,\hat{\mathbf{n}}\right\} = \frac{1}{c^2}\left\{\frac{\left[3\hat{\mathbf{n}}(\mathbf{p}\cdot\hat{\mathbf{n}})-\mathbf{p}\right]}{4\pi\,\varepsilon_o\, r^3} + \frac{\left[3\hat{\mathbf{n}}(\dot{\mathbf{p}}\cdot\hat{\mathbf{n}})-\dot{\mathbf{p}}\right]}{4\pi\,\varepsilon_o c\, r^2} + \frac{(\ddot{\mathbf{p}}\,\mathbf{x}\,\hat{\mathbf{n}})\,\mathbf{x}\,\hat{\mathbf{n}}}{4\pi\,\varepsilon_o r\, c^2}\right\} \qquad (12.2.6)$$

If we just let **p** go to **m** this equation tells us what the curl of **A**(t') is in equation (12.2.3). Thus, we have the expression for **B** for terms proportional to **m** and its time derivatives:

$$\mathbf{B}(\mathbf{r},t) = \frac{\mu_o}{4\pi}\left\{\frac{1}{r^3}\left[3\hat{\mathbf{n}}(\mathbf{m}\cdot\hat{\mathbf{n}}) - \mathbf{m}\right] + \frac{1}{c\,r^2}\left[3\hat{\mathbf{n}}(\dot{\mathbf{m}}\cdot\hat{\mathbf{n}}) - \dot{\mathbf{m}}\right] + \frac{1}{rc^2}(\ddot{\mathbf{m}}\,\mathbf{x}\,\hat{\mathbf{n}})\,\mathbf{x}\,\hat{\mathbf{n}}\right\} \qquad (12.2.7)$$

quasi-static induction radiation

Again, we see that we have the quasi-static magnetic dipole term, plus induction and radiation terms.

What about **E** for the terms involving **m**? Well, we could go back and pull it out of $\mathbf{E}(\mathbf{r},t) = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$, but this is not necessary since we can again use Maxwell's equations. We know (12.2.5) is true, and it is just as true if we replace **p**(t') in that equation with **m**(t'). That is, we must have (using $c^2 = \frac{1}{\mu_o\varepsilon_o}$)

$$\nabla\,\mathbf{x}\,\frac{\mu_o}{4\pi}\left\{\left[\frac{\ddot{\mathbf{m}}}{c\,r} + \frac{\dot{\mathbf{m}}}{r^2}\right]\mathbf{x}\,\hat{\mathbf{n}}\right\} = \frac{\partial}{\partial t}\mu_o\left\{\frac{\left[3\hat{\mathbf{n}}(\mathbf{m}\cdot\hat{\mathbf{n}}) - \mathbf{m}\right]}{4\pi\,r^3} + \frac{\left[3\hat{\mathbf{n}}(\dot{\mathbf{m}}\cdot\hat{\mathbf{n}}) - \dot{\mathbf{m}}\right]}{4\pi\,c\,r^2} + \frac{(\ddot{\mathbf{m}}\,\mathbf{x}\,\hat{\mathbf{n}})\,\mathbf{x}\,\hat{\mathbf{n}}}{4\pi\,rc^2}\right\}$$

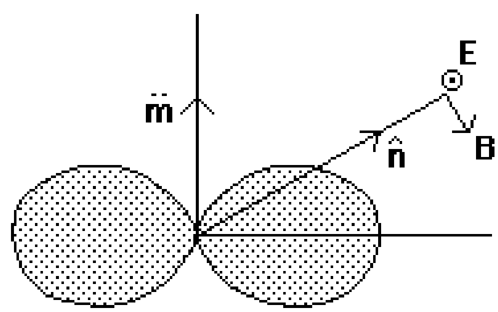
But the term in brackets on the right hand side of this equation is just the **B** field which involves magnetic dipole terms, so that

$$\nabla\,\mathbf{x}\left\{\frac{\mu_o}{4\pi}\left[\frac{\ddot{\mathbf{m}}}{c\,r} + \frac{\dot{\mathbf{m}}}{r^2}\right]\mathbf{x}\,\hat{\mathbf{n}}\right\} = \frac{\partial}{\partial t}\mathbf{B} \qquad (12.2.8)$$

But we know from Maxwell's equations that the **E** field must satisfy $\nabla\times\mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B}$.
Comparing this equation with equation (12.2.8), we see that the **E** field for terms involving **m** and its time derivatives must be given by

$$\mathbf{E}(\mathbf{r},t) = -\frac{\mu_o}{4\pi}\left[\frac{\ddot{\mathbf{m}}}{c\,r} + \frac{\dot{\mathbf{m}}}{r^2}\right]\mathbf{x}\,\hat{\mathbf{n}} \qquad (12.2.9)$$

If we abstract from equations (12.2.7) and (12.2.9) the radiation terms, we have the expressions for magnetic dipole radiation:



$$\mathbf{B}_{mag\ dip}(\mathbf{r},t) = \frac{\mu_o}{4\pi\ rc^2}(\ddot{\mathbf{m}}\,\mathbf{x}\,\hat{\mathbf{n}})\,\mathbf{x}\,\hat{\mathbf{n}}$$

$$\mathbf{E}_{mag\ dip}(\mathbf{r},t) = -\frac{\mu_o}{4\pi\,rc}\ddot{\mathbf{m}}\ \mathbf{x}\ \hat{\mathbf{n}}$$

(12.2.10)

For magnetic dipole radiation, **B** is in the plane of $\hat{\mathbf{n}}$ and $\ddot{\mathbf{m}}$ and **E** is perpendicular to that plane, *just the opposite of the situation for electric dipole radiation*. The angular distribution of the power radiated per unit solid angle is the same as for electric dipole radiation. That is, it goes as $\sin^2\theta$, where θ is the angle between $\hat{\mathbf{n}}$ and $\ddot{\mathbf{m}}$.

The total energy per second radiated is given by a form similar to equation (11.1.2), with $\ddot{\mathbf{p}}$ replaced by $\ddot{\mathbf{m}}/c$, that is

$$\frac{dW_{mag\ dip}}{dt} = \frac{\mu_o}{4\pi}\frac{2\left|\ddot{\mathbf{m}}\right|^2}{3c^3} \quad . \tag{12.2.11}$$

It is important to note that for particle motion, **J** in equation (8.2.14) defining **m** is a charge density times a velocity V of a particle, and simple dimensional analysis leads to the conclusion that:

For non-relativistic particle motion, we always have that the ratio of the power radiated into magnetic dipole radiation to that radiated into electric dipole radiation is $(V/c)^2$, unless $\left|\mathbf{p}\right|$ happens to be zero.

12.3 Electric quadrupole radiation

We now turn to electric quadrupole radiation, which we can obtain from the 3rd and 4th terms in equation (12.2.2). This is complicated mathematically, and let us start out by stating what the important points are.

The energy radiated into electric quadrupole will be down by a factor $(d/\lambda)^2$ compared to that radiated into electric dipole radiation. Thus, unless the electric dipole moment is zero, electric quadrupole radiation is an unimportant addition to the radiated energy for $d << \lambda$. Also, the frequency that emerges if we use the oscillating charge example of Section V above is *twice* the frequency ω_o with which the charge oscillates (see Section VIII below).

The following is an identity for the current density **J**

$$\int\left[\mathbf{r}'(\hat{\mathbf{n}}\cdot\mathbf{J}) + \mathbf{J}(\hat{\mathbf{n}}\cdot\mathbf{r}')\right]d^3x' = +\int\mathbf{r}'(\hat{\mathbf{n}}\cdot\mathbf{r}')\frac{\partial}{\partial t'}\rho(\mathbf{r}',t')\,d^3x' \tag{12.3.1}$$

Using this equation, we can write the electric quadrupole part of **A** in (12.2.2) as

$$\mathbf{A}_{el\ quad}(\mathbf{r},t) = \frac{\mu_o}{8\pi r^2}\int\left[\mathbf{r}'(\hat{\mathbf{n}}\cdot\mathbf{r}')\frac{\partial\rho(\mathbf{r}',t')}{\partial t'}\right]d^3x' + \frac{\mu_o}{8\pi c\, r}\int\left[\mathbf{r}'(\hat{\mathbf{n}}\cdot\mathbf{r}')\frac{\partial^2\rho(\mathbf{r}',t')}{\partial t'^2}\right]d^3x' \tag{12.3.2}$$

From now on, we drop all but the radiation terms in our expressions, since if we keep the full expansion, things get really messy. Thus we drop the first term in (12.3.2) So

$$\mathbf{A}_{el\ quad}(\mathbf{r},t)=\frac{\mu_o}{8\pi\, c\, r}\frac{d^2}{dt'^2}\int\left[\mathbf{r}'\,(\hat{\mathbf{n}}\cdot\mathbf{r}')\rho(\mathbf{r}',t')\right]\,d^3x' \tag{12.3.3}$$

For the moment, we define the 2nd rank tensor $\overleftrightarrow{\mathbf{H}}(t')$ by the equation

$$\overleftrightarrow{\mathbf{H}}(t')=\frac{d^2}{dt'^2}\int\left[\mathbf{r}'\,\mathbf{r}'\rho(\mathbf{r}',t')\right]d^3x'=\int\left[\mathbf{r}'\,\mathbf{r}'\ddot{\rho}(\mathbf{r}',t')\right]d^3x' \tag{12.3.4}$$

Then

$$\mathbf{A}_{el\ quad}(\mathbf{r},t)=\frac{\mu_o}{8\pi\, c\, r}\hat{\mathbf{n}}\cdot\overleftrightarrow{\mathbf{H}}(t') \tag{12.3.5}$$

and thus

$$\mathbf{B}_{el\ quad}(\mathbf{r},t)=\nabla\times\mathbf{A}_{el\ quad}=\nabla\times\frac{\mu_o\,\hat{\mathbf{n}}\cdot\overleftrightarrow{\mathbf{H}}(t')}{8\pi\, c\, r} \tag{12.3.6}$$

$$\left(\nabla\times\mathbf{A}_{el\ quad}\right)_j=\varepsilon_{jkl}\frac{\mu_o}{8\pi\, c}\frac{\partial}{\partial x_k}\frac{n_i H_{il}}{r} \tag{12.3.7}$$

$$\left(\nabla\times\mathbf{A}_{el\ quad}\right)_j=\varepsilon_{jkl}\frac{\mu_o}{8\pi\, c}\frac{n_i}{r}\frac{\partial}{\partial x_k}H_{il}+\varepsilon_{jkl}\frac{\mu_o}{8\pi\, c}H_{il}\frac{\partial}{\partial x_k}\frac{x_i}{r^2} \tag{12.3.8}$$

The second term on the right side of equation (12.3.8) is proportional to $1/r^2$, and since we are keeping only radiation terms, we drop it. Using our prescription for taking gradients of functions of $t' = t - r/c$, we have

$$B_j=-\varepsilon_{jkl}\frac{\mu_o}{8\pi\, c^2}\frac{n_k n_i}{r}\dot{H}_{il}\quad\text{or}\quad\mathbf{B}_{el\ quad}=-\hat{\mathbf{n}}\times\frac{\mu_o}{8\pi\, c^2}\frac{\left[\hat{\mathbf{n}}\cdot\dot{\overleftrightarrow{\mathbf{H}}}(t')\right]}{r} \tag{12.3.9}$$

Using the definition of $\overleftrightarrow{\mathbf{H}}(t')$,

$$\mathbf{B}_{el\ quad}(\mathbf{r},t)=-\left[\frac{\mu_o}{8\pi\, c^2}\frac{\hat{\mathbf{n}}}{r}\right]\times\left[\hat{\mathbf{n}}\cdot\int\left[\mathbf{r}'\,\mathbf{r}'\,\dddot{\rho}(\mathbf{r}',t')\right]d\tau'\right] \tag{12.3.10}$$

which can be written as

$$\mathbf{B}_{el\ quad}(\mathbf{r},t)=-\left[\frac{\mu_o}{8\pi\, c^2}\frac{\hat{\mathbf{n}}}{r}\right]\times\frac{\partial^3}{\partial t'^3}\left[\hat{\mathbf{n}}\cdot\int\left[\mathbf{r}'\,\mathbf{r}'-\frac{1}{3}\overleftrightarrow{\mathbf{I}}(r')^2\right]\rho(\mathbf{r}',t')d\tau'\right] \tag{12.3.11}$$

where we have added a term involving the identity tensor. The term that we have added is proportional to $\hat{\mathbf{n}} \times \hat{\mathbf{n}}$ and is therefore zero. Using (8.2.15) for the definition of the quadrupole moment, we have

$$\mathbf{B}_{el\ quad}(\mathbf{r},t) = -\frac{\mu_o}{24\pi c^2}\frac{1}{r}\hat{\mathbf{n}} \times \left[\hat{\mathbf{n}} \cdot \dddot{\overleftrightarrow{\mathbf{Q}}}\right] \tag{12.3.12}$$

For quadrupole radiation, then, using equation (10.3.6) and (12.3.12), the energy flux into unit solid angle is given by

$$\frac{dW_{el\ quad}}{d\Omega\, dt} = \frac{\mu_o}{(24)^2 \pi^2 c^3}\left|\hat{\mathbf{n}} \times \left[\hat{\mathbf{n}} \cdot \dddot{\overleftrightarrow{\mathbf{Q}}}\right]\right|^2 \tag{12.3.13}$$

A tedious integration of equation (12.3.13) over all solid angles gives the total radiated power

$$\frac{dW_{el\ quad}}{dt} = \frac{\mu_o}{720\pi c^3}\sum_{i=1}^{3}\sum_{j=1}^{3}\left| \dddot{Q}_{ij} \right|^2 \tag{12.3.14}$$

12.4 An Example Of Electric Quadrupole Radiation

We take the same problem as for the electric dipole example above, except now we compute the quadrupole radiation. What is $\overleftrightarrow{\mathbf{Q}}$? Well, using the definition in equation (8.2.15) and the charge density in equation **Error! Reference source not found.**, we have $Q_{xy} = Q_{yz} = Q_{xz} = 0$. Moreover,

$$\begin{aligned}
Q_{xx}(t') &= q_o \int \delta(x')\delta(y')\delta(z' - R_o \cos\omega_o t')\left[3x'^2 - (x'^2 + y'^2 + z'^2)\right] d^3x' \\
Q_{xx}(t') &= -q_o R_o^{\,2} \cos^2 \omega_o t' = Q_{yy}(t') \\
Q_{zz}(t') &= q_o \int \delta(x')\delta(y')\delta(z' - R_o \cos\omega_o t')\left[3z'^2 - (x'^2 + y'^2 + z'^2)\right] d^3x' \\
Q_{zz}(t') &= +2 q_o R_o^{\,2} \cos^2 \omega_o t'
\end{aligned}$$

If we use the trig identity $\cos^2 \omega_o t' = \frac{1}{2}(1 + \cos 2\omega_o t')$ we can write $\overleftrightarrow{\mathbf{Q}}$ as

$$\overleftrightarrow{\mathbf{Q}}(t') = q_o R_o^{\,2}(1 + \cos 2\omega_o t')\begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{12.4.1}$$

Note that we now have a ***time variation at a frequency of*** $2\omega_o$. This is because the quadrupole goes as the square of the position of the charge. Higher moments, which go as higher powers of the position of the charge, will for that reason exhibit time variations at higher multiples of ω_o. Note also that the trace of $\overleftrightarrow{\mathbf{Q}}$ is 0, as it must be. Taking the appropriate derivatives, we have

$$\dddot{\overleftrightarrow{\mathbf{Q}}}(t') = 8q_o R_o^{\,2}\omega_o^{\,3}\sin 2\omega_o t' \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{12.4.2}$$

Let's look at the angular distribution of this radiation. From (12.3.12), we see that we first need to compute $\hat{\mathbf{n}}\cdot\dddot{\overleftrightarrow{\mathbf{Q}}}$. First of all, $\hat{\mathbf{n}}$ is the unit vector in the radial direction, and in Cartesian coordinates that vector is $\hat{\mathbf{n}} = \hat{\mathbf{e}}_r = \sin\theta\cos\phi\,\hat{\mathbf{e}}_x + \sin\theta\sin\phi\,\hat{\mathbf{e}}_y + \cos\theta\,\hat{\mathbf{e}}_z$. Now,

$$\left[\hat{\mathbf{e}}_r\cdot\dddot{\overleftrightarrow{\mathbf{Q}}}\right]_i = \sum_{j=1}^{3} e_{rj}\dddot{Q}_{ji} = e_{ri}\dddot{Q}_{ii} \tag{12.4.3}$$

since $\dddot{\overleftrightarrow{\mathbf{Q}}}$ is diagonal, so that

$$\begin{aligned}\hat{\mathbf{e}}_r\cdot\dddot{\overleftrightarrow{\mathbf{Q}}}(t') &= C(t')\left[-\tfrac{1}{2}\sin\theta\cos\phi\,\hat{\mathbf{e}}_x - \tfrac{1}{2}\sin\theta\sin\phi\,\hat{\mathbf{e}}_y + \cos\theta\,\hat{\mathbf{e}}_z\right] \\ &\text{where } \mathrm{C(t')} = 8q_o R_o^{\,2}\omega_o^{\,3}\sin 2\omega_o t'\end{aligned} \tag{12.4.4}$$

We want to express this vector in spherical components, using the standard relationships:

$$\begin{aligned}\hat{\mathbf{e}}_x &= \sin\theta\cos\phi\,\hat{\mathbf{e}}_r + \cos\theta\cos\phi\,\hat{\mathbf{e}}_\theta - \sin\phi\,\hat{\mathbf{e}}_\phi \\ \hat{\mathbf{e}}_y &= \sin\theta\sin\phi\,\hat{\mathbf{e}}_r + \cos\theta\sin\phi\,\hat{\mathbf{e}}_\theta + \cos\phi\,\hat{\mathbf{e}}_\phi \\ \hat{\mathbf{e}}_z &= \cos\theta\,\hat{\mathbf{e}}_r - \sin\theta\;\hat{\mathbf{e}}_\theta\end{aligned} \tag{12.4.5}$$

This gives

$$\hat{\mathbf{e}}_r\cdot\dddot{\overleftrightarrow{\mathbf{Q}}}(t') = C(t')\begin{bmatrix} +\hat{\mathbf{e}}_r(\cos^2\theta - \frac{1}{2}\sin^2\theta\cos^2\phi - \frac{1}{2}\sin^2\theta\sin^2\phi) \\ +\hat{\mathbf{e}}_\theta(-\sin\theta\cos\theta - \frac{1}{2}\sin\theta\cos\theta\cos^2\phi - \frac{1}{2}\sin\theta\cos\theta\sin^2\phi) \\ +\hat{\mathbf{e}}_\phi(\frac{1}{2}\sin\theta\cos\phi\sin\phi - \frac{1}{2}\sin\theta\cos\phi\sin\phi) \end{bmatrix} \tag{12.4.6}$$

or

$$\hat{\mathbf{e}}_r\cdot\dddot{\overleftrightarrow{\mathbf{Q}}}(t') = C(t')\left[\hat{\mathbf{e}}_r(\cos^2\theta - \tfrac{1}{2}\sin^2\theta) - \hat{\mathbf{e}}_\theta\,\tfrac{3}{2}\sin\theta\cos\theta\right] \tag{12.4.7}$$

and finally we have

$$\begin{aligned}\hat{\mathbf{n}}\,\mathbf{x}\left[\hat{\mathbf{n}}\cdot\dddot{\overleftrightarrow{\mathbf{Q}}}\right] &= \hat{\mathbf{e}}_r\;\mathbf{x}\left[\hat{\mathbf{e}}_r\cdot\dddot{\overleftrightarrow{\mathbf{Q}}}\right] = -\tfrac{3}{2}\sin\theta\cos\theta\, C\,\hat{\mathbf{e}}_r\;\mathbf{x}\,\hat{\mathbf{e}}_\theta \\ \hat{\mathbf{n}}\,\mathbf{x}\left[\hat{\mathbf{n}}\cdot\dddot{\overleftrightarrow{\mathbf{Q}}}\right] &= -\tfrac{3}{2}\sin\theta\cos\theta\, C\,\hat{\mathbf{e}}_\phi\end{aligned} \tag{12.4.8}$$

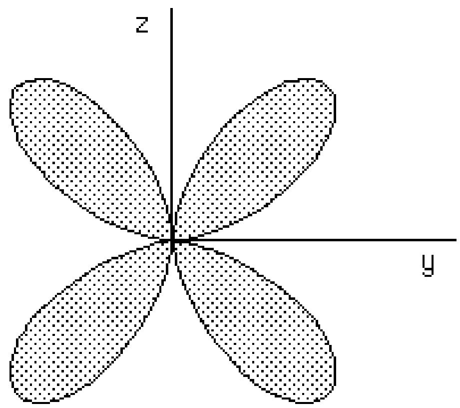
From (12.4.8), we see that **B** is in the $\hat{\mathbf{e}}_\phi$ direction, and thus **E** is in the $\hat{\mathbf{e}}_\theta$ direction, for this radiation field. From (12.3.13), we have that

$$\begin{aligned}\frac{dW_{el\ quad}}{d\Omega\, dt} &= \frac{\mu_o}{256\pi^2 c^3} C^2(t') \sin^2\theta \cos^2\theta \\ &= \frac{\mu_o c^3}{4\pi^2} k^6 q_o{}^2 R_o{}^4 \sin^2 2\omega_o t' \ \sin^2\theta \cos^2\theta\end{aligned} \tag{12.4.9}$$

Time averaging over one period gives the average power radiated per solid angle

$$\left\langle \frac{dW_{el\ quad}}{d\Omega\, dt} \right\rangle = \frac{\mu_o c^3}{8\pi^2} k^6 q_o{}^2 R_o{}^4 \sin^2\theta \cos^2\theta \tag{12.4.10}$$

The angular distribution of this radiation is shown in the sketch.



Note that if we compare this to the amount of power radiated into electric dipole radiation by this same system (equation (11.2.3)), we see that the ratio of quadrupole to dipole radiated power goes as $k^2 R_o{}^2 \approx R_o{}^2 / \lambda^2 << 1$, by assumption. So the power radiated into quadrupole radiation is unimportant under this assumption, unless the electric dipole moment is identically zero.

Note: in this example, the quadrupole radiation is emitted at an angular frequency of $2\omega_o$, and *not* ω_o, the frequency at which electric dipole radiation is emitted. In general, the oscillation of a charge as in this example here will result in radiation of ω_o, $2\omega_o$, $3\omega_o$ ...., that is all harmonics of ω_o. If the electric dipole approximation is not satisfied, the radiation emitted will emerge at higher and higher multiples of the fundamental. Sychrotron radiation of relativistic particles is a good example, where the radiation emitted extends up to $\gamma^3\omega_o$.

We go to no higher orders. To properly treat the expansion to all orders in d/λ, we need to introduce vector spherical harmonics (e.g., see Jackson, *Classical Electrodynamics*, Chapter 16).

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 $\mathbf{v}(t)=\Delta\mathbf{r}/\Delta t$. 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

$$\mathbf{v} = \mathrm{v}\,\hat{\mathbf{x}} \;. \tag{13.2.1}$$

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 $(\overline{x}_b, \overline{y}_b, \overline{z}_b, \overline{t}_b)$ .. 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 $(\overline{x}_b, \overline{y}_b, \overline{z}_b, \overline{t}_b)$ the coordinates of this particular space-time event, that is, we re-compute locations in space and time as $(\overline{x} - \overline{x}_b, \overline{y} - \overline{y}_b, \overline{z} - \overline{z}_b, \overline{t} - \overline{t}_b)$. 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 $t = \overline{t} = 0$.

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:

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

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

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

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

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,

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

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

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

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 $G(x,t)=G(x(\overline{x},\overline{t}),t(\overline{x},\overline{t}))=\overline{G}(\overline{x},\overline{t})$, where we have put a bar on *G* because the functional form of $\overline{G}(\overline{x},\overline{t})$ is different from that of $G(x,t)$, we have

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

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

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

Now, we similarly have that

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

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

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

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

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

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

With a little manipulation, this can be written as

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

or

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

Remember, the ratio $\frac{\bar{\omega}}{\bar{k}}$ 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

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

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

$$\frac{a_{10}}{a_{11}} = -\mathrm{v} \qquad (13.7.2)$$

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

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \overline{t}}\frac{\partial \overline{t}}{\partial t} + \frac{\partial}{\partial \overline{x}}\frac{\partial \overline{x}}{\partial t} = a_{00}\frac{\partial}{\partial \overline{t}} + a_{10}\frac{\partial}{\partial \overline{x}} \tag{13.7.3}$$

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial \overline{t}}\frac{\partial \overline{t}}{\partial x} + \frac{\partial}{\partial \overline{x}}\frac{\partial \overline{x}}{\partial x} = a_{01}\frac{\partial}{\partial \overline{t}} + a_{11}\frac{\partial}{\partial \overline{x}} \tag{13.7.4}$$

$$\frac{\partial}{\partial y} = \frac{\partial}{\partial \overline{y}} \qquad\qquad \frac{\partial}{\partial z} = \frac{\partial}{\partial \overline{z}} \tag{13.7.5}$$

which means that under this transformation

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

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

so that the wave equation becomes in the barred system is

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

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

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

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

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

These equations define the Lorentz transformations[3]. They reduce to the Galilean transformations for $\mathrm{v}/c << 1$. 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 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

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

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

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

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

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

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

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

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

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

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

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

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

Inserting (14.3.3) into (14.3.1) gives

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

which with a little rearrangement can be written as

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

or

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

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

$$\frac{\partial \overline{E}_x}{\partial \overline{x}}+\frac{\partial \overline{E}_y}{\partial \overline{y}}+\frac{\partial \overline{E}_z}{\partial \overline{z}}=\frac{\overline{\rho}}{\varepsilon_o} \qquad (14.3.7)$$

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

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

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

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

then

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

or

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

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

$$J^{\mu} = (c\rho, J_x, J_y, J_z) \tag{15.2.4}$$

Rewriting equation (14.2.4) slightly, we have

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

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

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

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

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

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 $\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$. 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

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

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

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

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

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

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

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

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

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

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

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

$$\partial_\mu \Theta^{\mu\nu} = -F^{\nu\sigma} J_\sigma \tag{15.7.5}$$

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

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

where

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

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):

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

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

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

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:

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

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

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

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

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

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

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

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

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

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

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

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

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

and

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

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

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

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 $\left|\,\mathbf{r}-\mathbf{X}(t')\right|$ 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

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

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

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

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

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

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

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

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

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

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

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

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))

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

or

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

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

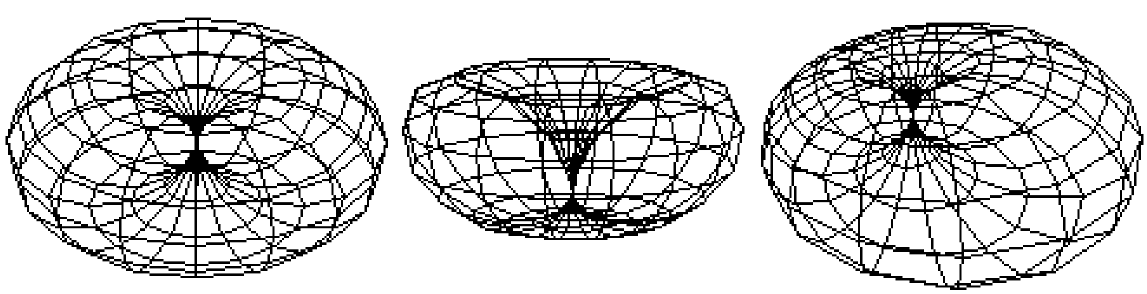
$$\frac{dW_{rad}}{d\Omega\,dt'} = \frac{dW_{rad}}{d\Omega\,dt}\frac{dt}{dt'} = \frac{q^2}{(4\pi)^2\,c\varepsilon_o}\frac{\left|\hat{\mathbf{n}}\times\left[(\hat{\mathbf{n}}-\boldsymbol{\beta})\times\dot{\boldsymbol{\beta}}\right]\right|^2}{(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})^5} \qquad (17.4.7)$$

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)

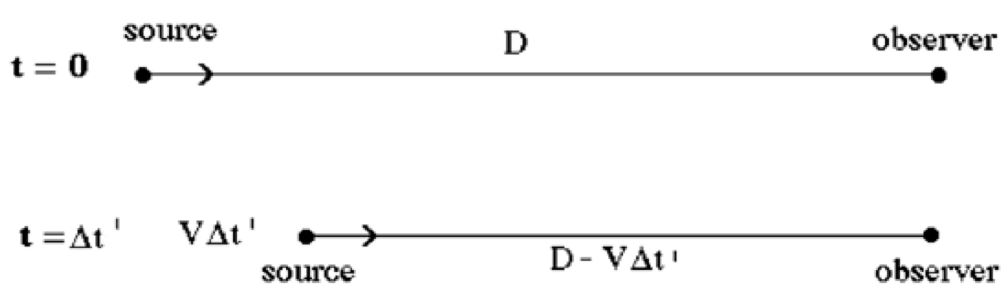


Figure 17-2: Beeps emitted by source as seen by the observer

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

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

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

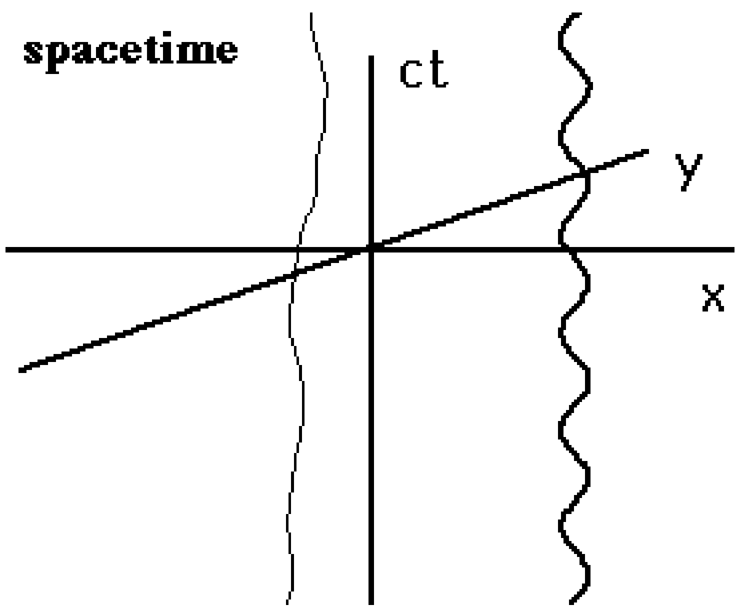


Figure 17-3: The space-time diagram

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

$$r = \sqrt{(x^2 + y^2)} = -ct \qquad (17.6.1)$$

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.

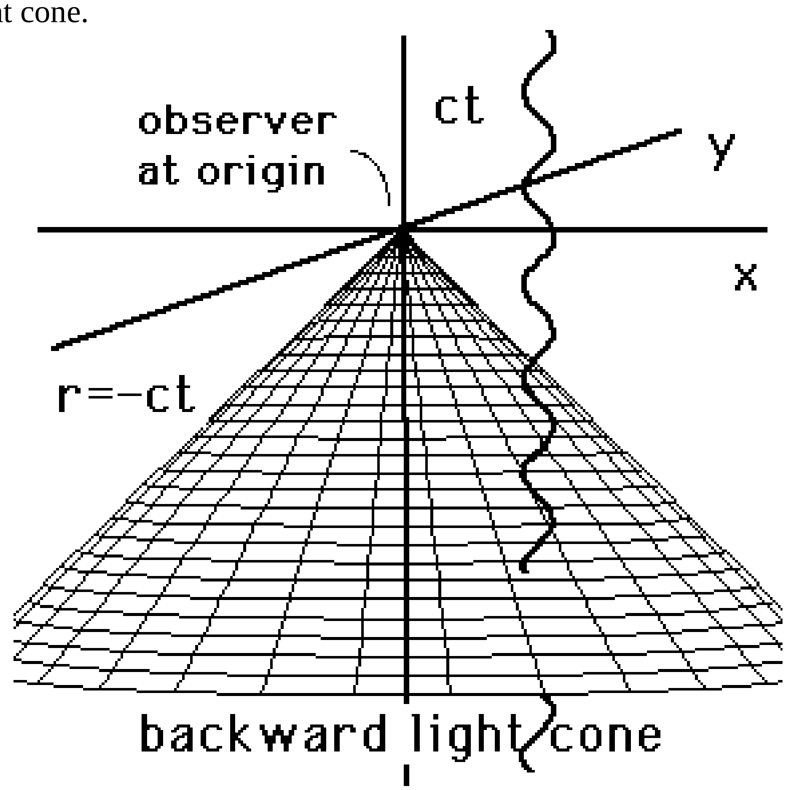


Figure 17-4: The observer's backward light cone.

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?

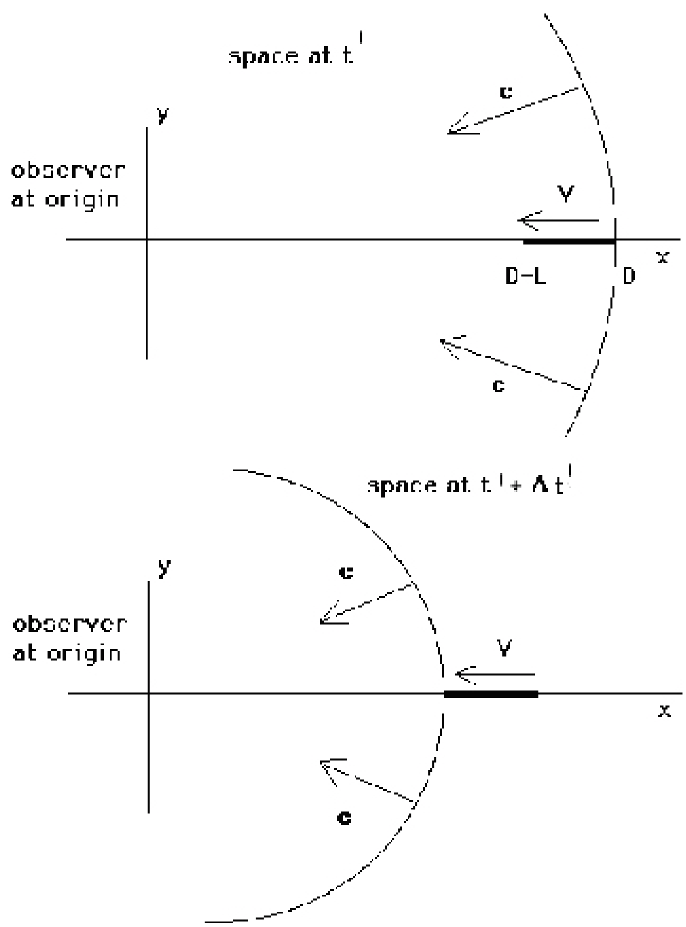


Figure 17-5: Sampling a line charge moving toward the observer.

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

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

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 $(1-\hat{\mathbf{n}}\cdot\boldsymbol{\beta})$.

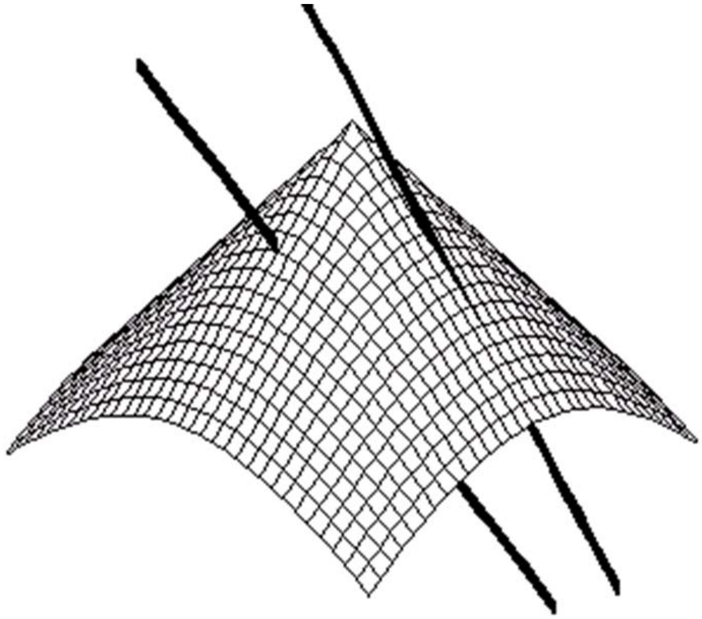


Figure 17-6: Two charges cutting the backward light cone of the observer.

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.

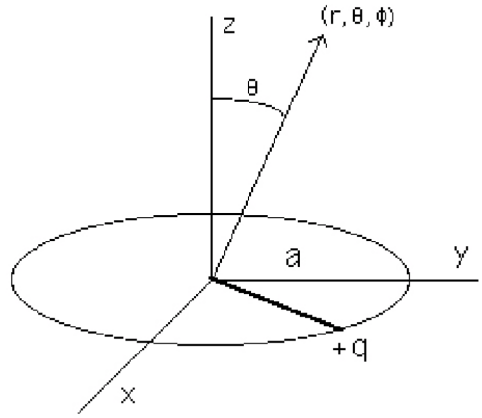


Figure 17-7: A charge moving in a circle in the *x*-*y* plane

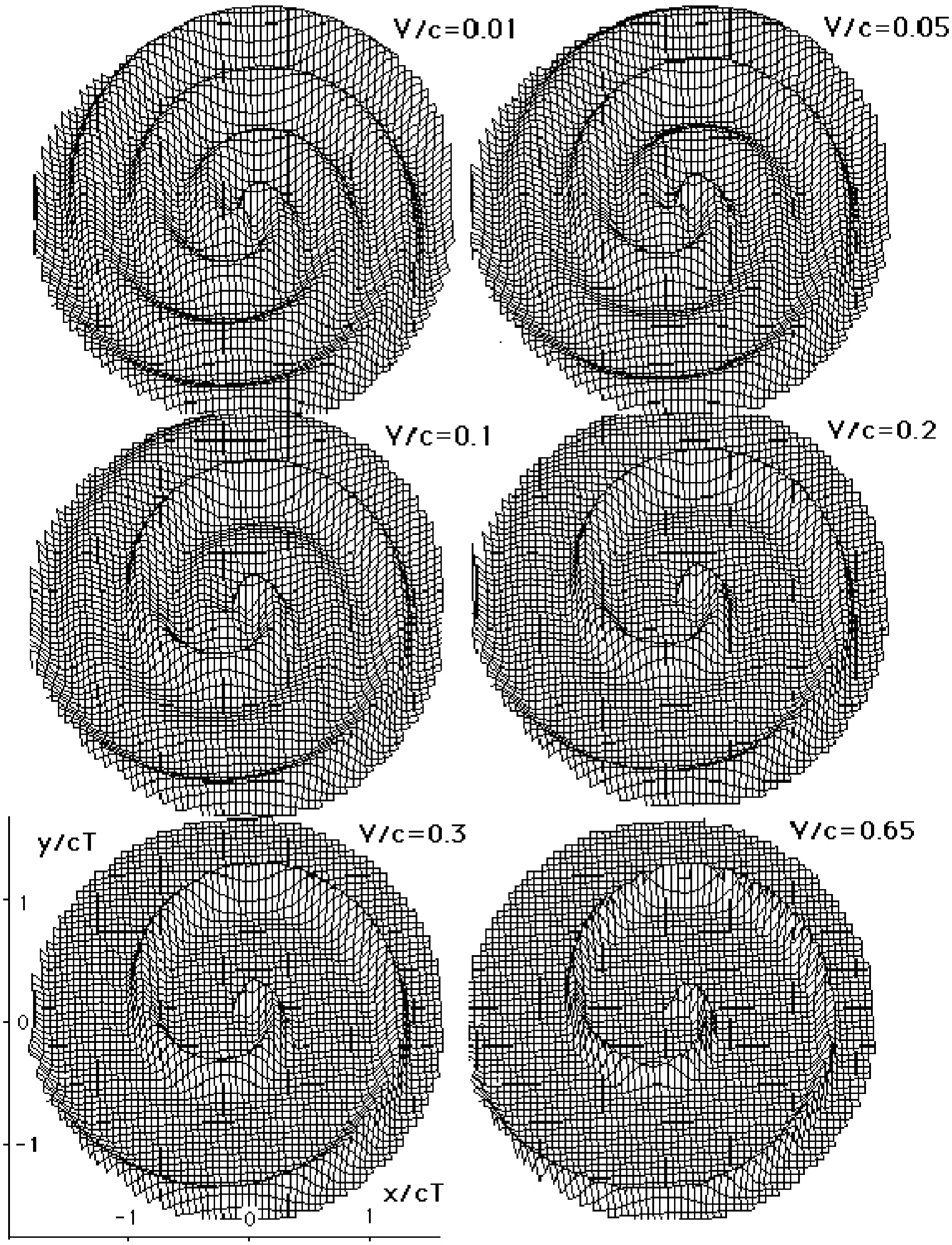


Figure 17-8: The spatial distribution of synchrotron radiation in the *x-y* plane.

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

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

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

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

If we differentiate equation (17.7.3), we find that

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

or that

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

This expression for $(1 - \hat{\mathbf{n}}\cdot\boldsymbol{\beta})$ is smallest when $\omega_o t'_{ret}$ is 3π/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.

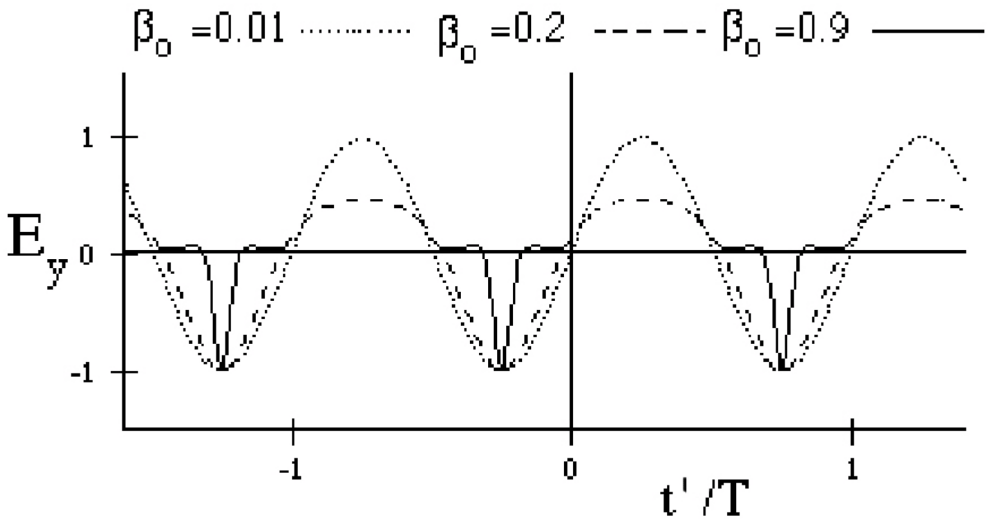


Figure 17-9: E_y as a function of time for various values of V_o/c.

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

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

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

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

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

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

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

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

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 $\mathbf{F}_{\text{radiation reaction}}$

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

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.

Charges in the application interact via the Coulomb force, with a Pauli repulsive force at close distances, plus a damping force proportional to velocity. The Pauli repulsive force goes at inverse radius to the sixth power, so it is very "stiff". That is it either dominates the interaction or it is more or less negligible compared to the Coulomb force. So as you watch the charges in the application interact, they "bounce" at close distances, when the Pauli repulsion very quickly overpowers any Coulomb attraction between charges. The damping proportional to the velocity allows our particles to settle down to a meta-stable configuration. If we did not have the damping there would be a continuous interchange of energy between the kinetic energy of the charges and the energy stored in the field. The damping allows us to drain away that kinetic energy so that the particles end up in a meta-stable state with zero kinetic energy and a local minimum in the electrostatic energy.

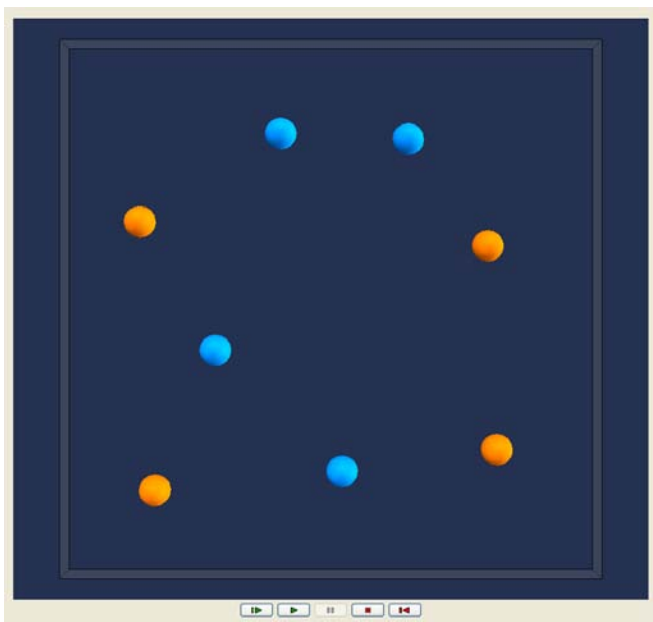


Figure 18-1: An applicaton showing charges interacting via Coulomb's Law
://web.mit.edu/viz/EM/visualizations/electrostatics/InteractingCharges/

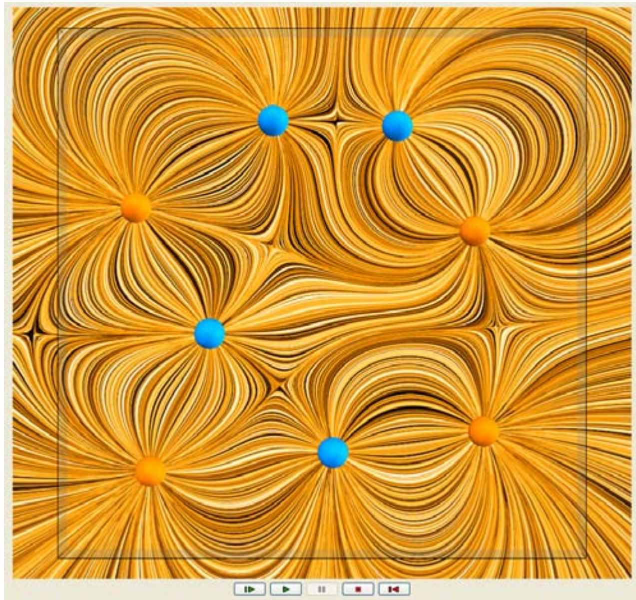


Figure 18-2: The same application as above except with a “grass seeds” representation of the electric fields

So the non-relativistic equations we are soving in this application, assuming we have N interacting charges $\{q_i\}_{i=1}^N$ with masses $\{m_i\}_{i=1}^N$ and frictional damping rates $\{\gamma_i\}_{i=1}^N$ located at positions $\{\mathbf{r}_i(t)\}_{i=1}^N$, with velocities $\{\mathbf{u}_i(t)\}_{i=1}^N$, are

$$\mathbf{u}_i(t) = d\,\mathbf{r}_i(t)/dt \qquad (18.2.1)$$

$$m_i \frac{d\,\mathbf{u}_i(t)}{dt} = \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{q_i\, q_j}{4\pi\varepsilon_o} \frac{\left(\mathbf{r}_i - \mathbf{r}_j\right)}{\left|\mathbf{r}_i - \mathbf{r}_j\right|^3} + \sum_{\substack{j=1 \\ j\neq i}}^{N} P \frac{\left(\mathbf{r}_i - \mathbf{r}_j\right)}{\left|\mathbf{r}_i - \mathbf{r}_j\right|^7} - m_i \gamma_i \mathbf{u}_i \qquad (18.2.2)$$

Coulomb force Pauli repulsion drag

This is a system of first order differential equations which in two (three) dimensions has 4 (6) dependent variables for each particle (a position vector and a velocity vector) and thus a total of 4N (6N) dependent variables. For the system shown in Figure 18-1, where we have eight particles, this is 32 dependent variables. Solving this set of coupled dependent variable equations as a function of the independent variable t can be carried out by standard numerical techniques, for example fifth order Runge-Kutta. The first term on the right in (18.2.2) is either repulsive or attractive depending on whether the signs of the ith particle and that of the jth particle are the same or opposite, and goes as inverse distance squared between the charges. The second term is always repulsive, and goes as inverse distance to the sixth between the charges. The third term is dissipative, that is it always drains kinetic energy from the charges unless they are stationary.

As can be easily explored by playing with this application, there is an enormous wealth of complex intereactons between charged particles when their interaction is governed by the Coulomb interaction. The particles initially rapidly try to form electric dipoles, that is a positive charge in close proximity to a negative charge. Once dipoles form, the forces between dipoles is greatly reduced, and therefore the time scale for changes in position is greatly reduced, because the dipole fields fall off as inverse distance cubed rather than inversed distance squared. Eventually the charges will aggregate into a clump, forming stable "crystals".

These behaviours mimic what we see in the real world, where the everyday interactions between matter are dominated by electrostatic forces. Even though we can solve for the dynamics of our charges in this application simpy using the Coulomb repulsion or attraction (along with the Pauli repulsion at small distances and the frictional damping), that calculation hides an enormous amount of the physics. What is not apparent when we these dynamical calculations is the complex exchange of energy between field and charges as these interactions proceed. We will really not be able to explain the complexity of this exchange until we also consider magnetostatics, because we need **B** as well as **E** to understand the flow of energy in these interactions via the Poynting flux $\mathbf{E}\times\mathbf{B}/\mu_o$. But that is where we are going.

First though we need to explore the electric fields of stationary particles, as we do in the present section. Then we will look at the magnetic fields generated when we allow electric charges to move. Then we consider the electric fields that are present whenever we see time changing magnetic fields—that is, Faraday's Law. Then will come back and try to understand the complexity of the interactions shown in the application. There is far more than meets the eye here, which we will discuss eventually, but first we look at electrostatics.

18.3 The electric field E and potential ϕ of a set of point charges

The force $\mathbf{F}$ on a test charge Q located at $\mathbf{r}$ due to a charge q located at $\mathbf{r}'$ is given by

$$\mathbf{F} = \frac{qQ}{4\pi\varepsilon_o}\frac{\mathbf{r}-\mathbf{r}'}{\left|\mathbf{r}-\mathbf{r}'\right|^3} = \frac{qQ}{4\pi\varepsilon_o}\frac{\hat{\mathbf{n}}}{\left|\mathbf{r}-\mathbf{r}'\right|^2} \tag{18.3.1}$$

where as always, the unit vector $\hat{\mathbf{n}}$ points from the source to the observation point, that is

$$\hat{\mathbf{n}} = \frac{\mathbf{r}-\mathbf{r}'}{\left|\mathbf{r}-\mathbf{r}'\right|} \tag{18.3.2}$$

This force is attractive if the two charges have the opposite sign. It is repulsive if the two charges have opposite signs.

We define the electric field **E** at $\mathbf{r}$ to be the ratio of the force on the test charge to the test charge *Q*. That is, the electric field is given by

$$\mathbf{E}(\mathbf{r}) = \frac{\mathbf{F}}{Q} = \frac{q}{4\pi\varepsilon_o}\frac{\mathbf{r}-\mathbf{r}'}{\left|\mathbf{r}-\mathbf{r}'\right|^3} \tag{18.3.3}$$

Since we can write $\frac{\mathbf{r}-\mathbf{r}'}{\left|\mathbf{r}-\mathbf{r}'\right|^3} = -\nabla\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|}$, we can also write (18.3.3) as

$$\mathbf{E}(\mathbf{r}) = -\nabla\left[\frac{q}{4\pi\varepsilon_o}\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|}\right] = -\nabla\phi(\mathbf{r}) \quad \text{where} \quad \phi(\mathbf{r}) = \frac{q}{4\pi\varepsilon_o}\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|} + \text{Constant} \tag{18.3.4}$$

where $\phi(\mathbf{r})$ is the electrostatic potential of a point charge. We will discuss the physical meaning of the electrostatic potential below. The unit of the electric field are thus Newtons/Coulomb.

If we have a set of charges $\{q_i\}_{i=1}^N$ located at positions $\{\mathbf{r}'_i\}_{i=1}^N$, then the electric field of this collection of charges is simply the sum of the electric field of each individual charge, that is

$$\mathbf{E}(\mathbf{r}) = \sum_{i=1}^N \frac{q_i}{4\pi\varepsilon_o}\frac{\mathbf{r}-\mathbf{r}'_i}{\left|\mathbf{r}-\mathbf{r}'_i\right|^2} = -\nabla\phi(\mathbf{r}) \qquad \text{where} \qquad \phi(\mathbf{r}) = \sum_{i=1}^N \frac{q_i}{4\pi\varepsilon_o}\frac{1}{\left|\mathbf{r}-\mathbf{r}'_i\right|} + \text{Constant} \tag{18.3.5}$$

We note that since the curl of any gradient of a scalar is zero, we have for electrostatic fields that

$$\nabla\times\mathbf{E}(\mathbf{r}) = 0 \tag{18.3.6}$$

18.4 The electric field E and potential ϕ of a continuous charge distribution

Suppose we have a continuous distribution of electric charge defined by a volume change density $\rho(\mathbf{r}')$. Then the amount of charge in an infinitesimal volume element d^3x' is $dq' = \rho(\mathbf{r}')\,d^3x'$, and our sums in (18.3.5) go into integrals, as follows.

$$\mathbf{E}(\mathbf{r}) = \int\limits_{all\,space} \frac{\rho(\mathbf{r}')\,d^3x'}{4\pi\varepsilon_o}\frac{\left(\mathbf{r}-\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|^3} = -\nabla\phi(\mathbf{r}) \tag{18.4.1}$$

$$\phi(\mathbf{r}) = \int\limits_{all\,space} \frac{\rho(\mathbf{r}')\,d^3x'}{4\pi\varepsilon_o}\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|} \tag{18.4.2}$$

We can get the second equation for electrostatics from (18.4.1) by taking the divergence of (18.4.1), which yields

$$\nabla\cdot\mathbf{E}(\mathbf{r}) = -\nabla^2\phi(\mathbf{r}) = -\nabla^2 \int_{all\,space} \frac{\rho(\mathbf{r}')\,d^3x'}{4\pi\varepsilon_o}\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|} = -\int_{all\,space} \frac{\rho(\mathbf{r}')\,d^3x'}{4\pi\varepsilon_o}\nabla^2\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|} \quad (18.4.3)$$

But of course we have $\nabla^2 \frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|} = -4\pi\delta^3\left(\mathbf{r}-\mathbf{r}'\right)$, and using this to do the integral in (18.4.3) yields

$$\nabla\cdot\mathbf{E}(\mathbf{r}) = \rho(\mathbf{r})/\varepsilon_o \quad (18.4.4)$$

18.5 The physical meaning of the electric potential ϕ

In an early problem in our problem sets, we showed that if $\mathbf{E}(\mathbf{r})$ is a vector function which has zero curl, then the scalar function $-\int_{\mathbf{r}_o}^{\mathbf{r}}\mathbf{E}(\mathbf{r}')\cdot d\,\mathbf{l}'$, which is the line integral of **E** along any path connecting a fixed reference location $\mathbf{r}_o$ and the observer's position **r**, is independent of the path taken. Therefore this scalar function is a single valued and unique function given E and the reference location $\mathbf{r}_o$. We also saw how to get **E** given $-\int_{\mathbf{r}_o}^{\mathbf{r}}\mathbf{E}(\mathbf{r}')\cdot d\,\mathbf{l}'$ by looking at, for example

$$-\int_{\mathbf{r}}^{\mathbf{r}+\Delta x\hat{\mathbf{x}}}\mathbf{E}(\mathbf{r}')\cdot d\,\mathbf{l}' \approx -\Delta x\,E_x \quad (18.5.1)$$

If we generalize (18.5.1) to all three components we have

$$\mathbf{E} = -\nabla\int_{\mathbf{r}_o}^{\mathbf{r}}\mathbf{E}(\mathbf{r}')\cdot d\,\mathbf{l} \quad (18.5.2)$$

Comparing (18.5.2) to (18.4.2) or (18.3.5), we see that we must have

$$\phi(\mathbf{r}) = -\int_{\mathbf{r}_o}^{\mathbf{r}}\mathbf{E}(\mathbf{r}')\cdot d\,\mathbf{l}' \quad (18.5.3)$$

where we are free to choose the "reference" point $\mathbf{r}_o$. If possible we will choose that point to be at infinity, but that is not possible in some situations. So we now know that $\mathbf{E}(\mathbf{r})$ and $\phi(\mathbf{r})$ are related not only by $\mathbf{E}(\mathbf{r}) = -\nabla\phi(\mathbf{r})$ but also by (18.5.3).

Equation (18.5.3) allows us to interpret the meaning of $\phi(\mathbf{r})$ in physical terms. To do so we must talk about moving a test electric charge Q around in the presence of fixed charges whose electric field is **E**. How much work do I do in moving that charge from one point to another, say from **a** to **b**? Well first of all I must assume that I do this very slowly in some sense, because I do not want to have a substantial amount of energy radiated away to infinity, but if I do this slowly enough I can make the radiated energy as small as I desire. I can thus ignore any radiation reaction term, and the force I must exert to move the charge around is simply the force I need to counter balance the force associated with the electric field that the test charge feels, plus a little bit more. That is I need to exert a force

$$\mathbf{F}_{me} = -Q\mathbf{E}(\mathbf{r}) \tag{18.5.4}$$

plus a little tiny bit more, to get the test charge to actually move. Again, I can make this additional "little bit more" force and its associated work arbitrarily small.

Thus the work I must do in moving our test charge from **a** to **b** is given by

$$W_{me}^{\mathbf{a}\rightarrow\mathbf{b}} = \int_{\mathbf{a}}^{\mathbf{b}} \mathbf{F}_{me}\cdot(\mathbf{r}')\cdot d\,\mathbf{l}' = -\int_{\mathbf{a}}^{\mathbf{b}} Q\mathbf{E}(\mathbf{r}')\cdot d\,\mathbf{l}' = -Q\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{E}(\mathbf{r}')\cdot d\,\mathbf{l}' \tag{18.5.5}$$

But using (18.5.3), we have

$$W_{me}^{\mathbf{a}\rightarrow\mathbf{b}} = -Q\left[\int_{\mathbf{r}_o}^{\mathbf{b}} \mathbf{E}(\mathbf{r}')\cdot d\,\mathbf{l}' - \int_{\mathbf{r}_o}^{\mathbf{a}} \mathbf{E}(\mathbf{r}')\cdot d\,\mathbf{l}'\right] = Q\left[\phi(\mathbf{b}) - \phi(\mathbf{a})\right] \tag{18.5.6}$$

or

$$\left[\phi(\mathbf{b}) - \phi(\mathbf{a})\right] = \frac{W_{me}^{\mathbf{a}\rightarrow\mathbf{b}}}{Q} \tag{18.5.7}$$

Thus the difference in electric potential between two points **a** and **b** is the amount of work I must do to move a unit test charge from **a** to **b** against the electric field. The units of ϕ are thus joules per coulomb, or *volts*. The units of electric field are Newtons per coulomb, which are also joules per coulomb per meter, so the units of electric field are also volts/meter, and this is how we most often quote the units of electric field.

18.6 The energy required to assemble charges

18.6.1 A set of point charges

Now we can calculate the amount of energy we need to do to assemble a set of point charges, bringing them in from infinitiy. We do this one by one. It takes no enerty to bring in the first charge q_1 from infinity to its final position $\mathbf{r}_1$, because there is no electric field to work against. Once the first particle is there we have a potential given by (see (18.3.4)) $\phi_1(\mathbf{r}) = \frac{q_1}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}-\mathbf{r}_1|}$, where have taken our reference point for zero potential at infinity.

Now if we bring in the second charge q_2 from infinity to its final position $\mathbf{r}_2$, we can calculate the work we need to do this by using the meaning of the potential we discussed above, to find that the work we need to do to bring in this charge is $q_2\phi_1(\mathbf{r}_2) = \frac{q_1 q_2}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|}$. Now our potential with these two charges present is

$$\phi_{1+2}(\mathbf{r}) = \frac{q_1}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}-\mathbf{r}_1|} + \frac{q_2}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}-\mathbf{r}_2|} \tag{18.6.1}$$

To bring in a third charge q_3 from infinity to its final position $\mathbf{r}_3$ requires energy $q_3\phi_{1+2}(\mathbf{r}_3)$, or

$$q_3\phi_{1+2}(\mathbf{r}_3) = \frac{q_3 q_1}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}_3-\mathbf{r}_1|} + \frac{q_3 q_2}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}_3-\mathbf{r}_2|} \tag{18.6.2}$$

So the total work we have done thus far is the sum of (18.6.2) and $\frac{q_1 q_2}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|}$, or

$$W_3 = \frac{q_1 q_2}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|} + \frac{q_3 q_1}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}_3-\mathbf{r}_1|} + \frac{q_3 q_2}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}_3-\mathbf{r}_2|} \tag{18.6.3}$$

It is fairly easy to see that if we do this N times, we do total work

$$W_N = \sum_{j<i}\sum_{i=1}^{N}\frac{q_i q_j}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} \tag{18.6.4}$$

If we put in a factor of ½ to account for double counting, we can write this as

$$W_N = \frac{1}{2}\sum_{j\neq i}^{N}\sum_{i=1}^{N}\frac{q_i q_j}{4\pi\varepsilon_o}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} \tag{18.6.5}$$

18.6.2 A continuous distribution of charges

If we go over to a continuous distribution of charges, our double sum in (18.6.5) becomes a double integral, as follows.

$$W = \frac{1}{2}\int\int \frac{\left(\rho(\mathbf{r})d^3x\right)\left(\rho(\mathbf{r}')d^3x'\right)}{4\pi\varepsilon_o\left|\mathbf{r}-\mathbf{r}'\right|} = \frac{1}{2}\int \rho(\mathbf{r})d^3x\int \frac{\rho(\mathbf{r}')d^3x'}{4\pi\varepsilon_o\left|\mathbf{r}-\mathbf{r}'\right|} \qquad (18.6.6)$$
$$W = \int \frac{1}{2}\rho(\mathbf{r})\phi(\mathbf{r})d^3x$$

where we have used (18.4.2) to get the last form in (18.6.6). To put (18.6.6) in a more familiar form, we use $\mathbf{E}(\mathbf{r}) = -\nabla\phi(\mathbf{r})$ and $\nabla\cdot\mathbf{E}(\mathbf{r}) = \rho(\mathbf{r})$ to obtain

$$W = \int \frac{\varepsilon_o}{2}(\nabla\cdot\mathbf{E})\phi(\mathbf{r})d^3x = \int \frac{\varepsilon_o}{2}\nabla\cdot\left(\phi(\mathbf{r})\mathbf{E}\right)d^3x - \int \frac{\varepsilon_o}{2}\mathbf{E}\cdot\left(\nabla\phi(\mathbf{r})\right)d^3x \qquad (18.6.7)$$
$$W = \int \frac{1}{2}\varepsilon_o E^2 d^3x$$

where in the second term in (18.6.7) we have converted the volume integral of a divergence to a surface integral at infinity and set it to zero because the integrand falls off faster than inverse *r* cubed. Thus we recover in this expression the term we have seen before for the energy density in the electric field, $\frac{\varepsilon_o E^2}{2}$.

Note that the expression in (18.6.7) is always positive, whereas the similar expression for point charges (18.6.5) can be positive or negative. The difference is that in our sum in (18.6.5) we explicity exclude the $i = j$ terms, which corresponds to the energy required to assemble the point charges themselves, which is infinite. In a more realistic model we would have the point charges having some small but finite radius *R*, and the energy to assemble them would be $W = \frac{q^2}{4\pi\varepsilon_o R}$ (see the next paragraph for a more quantitative justification of this). This of course blows up as *R* approaches zero, which is why we do not include it in the sum in (18.6.5). As long as we are not disassembling point charges, however, there is no harm in ignoring these infinite terms, because it is only the changes in energy we are interested in as the configuration changes, not the total energy.

18.7 Where is the energy really located in space?

If we look at (18.6.6) and (18.6.7) you might think that we have two equally good expressions for the energy density of the electromagnetic field, either $\frac{1}{2}\rho\phi$ or $\frac{1}{2}\varepsilon_o E^2$. These give very different spatial distributions of energy density however. To illustrate this, let us consider the energy necessary to assemble a spherical shell of charge of radius *R*, carrying total charge *Q*, distributed uniformly over the sphere in a surface change density $\sigma = Q/4\pi R^2$. We can easily use Gauss's Law to calculate the electric field everywhere (see next Section) and then use the electric field and (18.5.3) to calculate the electric potential from the electric field, to obtain the following results.

$$\mathbf{E}(\mathbf{r}) = \begin{cases} 0 & \text{r < R} \\ \hat{\mathbf{r}}\dfrac{Q}{4\pi\varepsilon_o r^2} & \text{r > R} \end{cases} \tag{18.7.1}$$

$$\phi(\mathbf{r}) = \begin{cases} \dfrac{Q}{4\pi\varepsilon_o R} & \text{r < R} \\ \dfrac{Q}{4\pi\varepsilon_o r} & \text{r > R} \end{cases} \tag{18.7.2}$$

Calculating the energy required to put this distribution of charge together using (18.6.7) is straightforward, giving

$$W = \int \frac{1}{2}\varepsilon_o E^2 d^3x = \int d\Omega \int_R^\infty dr \frac{1}{2}\varepsilon_o \left[\frac{Q}{4\pi\varepsilon_o r^2}\right]^2 r^2 = \frac{Q^2}{8\pi\varepsilon_o R} \tag{18.7.3}$$

This calculation is even simpler using (18.6.6), since $\rho(\mathbf{r})$ is only non-zero at $r = R$.

$$W = \int \frac{1}{2}\rho(\mathbf{r})\phi(\mathbf{r})d^3x = \frac{1}{2}\phi(R)\int \rho(\mathbf{r})d^3x = \frac{Q^2}{8\pi\varepsilon_o R} \tag{18.8.1}$$

These two methods of calculating *W* give the same result, as they must, even though the integrands are very different. So what is the correct expression for the energy density of the electric field? Electrostatics actually gives us little guidance in how to answer this question, but when we introduce time dependence, the answer is quite clear. If we go back to (4.4.2) for the differential form of the conservation of energy, we had

$$\frac{\partial}{\partial t}\left[\frac{1}{2}\varepsilon_o E^2 + \frac{B^2}{2\mu_o}\right] + \nabla \cdot \left(\frac{\mathbf{E}\times\mathbf{B}}{\mu_o}\right) = -\mathbf{E}\cdot\mathbf{J} \tag{18.8.2}$$

and this form clearly chooses $\frac{1}{2}\varepsilon_o E^2$ for the local energy density of the electrostatic field. Moreover if we look at the time dependent process by which the electrostatic energy is located at a given point in space, we clearly see the flow of energy from where we are doing work to create it (where the creation rate for electromagnetic energy, $-\mathbf{E}\cdot\mathbf{J}$, is non-zero) to where it resides in space, as indicated by the local value of $\frac{1}{2}\varepsilon_o E^2$, through the agency of the Poynting flux, $\mathbf{E}\times\mathbf{B}/\mu_o$.

To take a concrete example of what I mean by this, consider the following scenario. An electric field is created by an external agent who separates charges. We start out with five negative electric charges and five positive charges, all at the same point in space. Since there is no net charge, there is no electric field. Now the agent moves one of the positive charges at constant velocity from its initial position to a distance L away along the horizontal axis. After doing that, the agent moves the second positive charge in the same manner to the position where the first positive charge sits. The agent continues on with the rest of the positive charges in the same manner, until all of the positive charges are sitting a distance L from their initial position along the horizontal axis. We have color coded the "grass seeds" representation in the still below to represent the strength of the electric field. Very strong fields are white, very weak fields are black, and fields of intermediate strength are yellow.

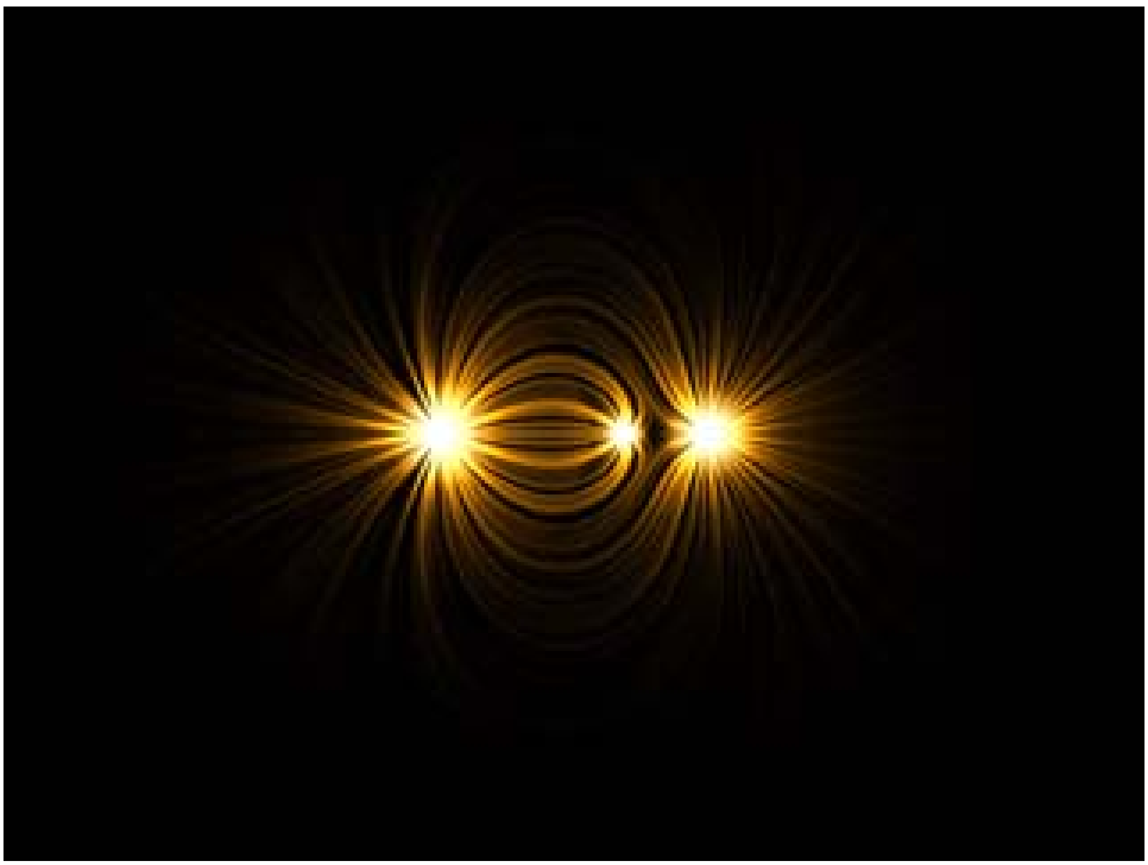


Figure 18-3: An external agent creating an electric field by separating + and – charges.

://web.mit.edu/viz/EM/visualizations/electrostatics/CreatingDestroyingEFields/

The field lines move in the direction of the energy flow of the electromagnetic field. Over the course of the animation, the strength of the electric field grows as each positive charge is moved into place. That energy flows out from the path along which the charges move, because that is where $-\mathbf{E}\cdot\mathbf{J} > 0$ is non-zero and positive, and nowhere else. That energy is being provided by the agent moving the charge against the electric field of the other charges. The work that this agent does to separate the charges against their electric attraction appears as energy in the electric field, and we can see it flow out from where it is created and take up its position in space.

When we do the reverse process, that is the external agent now moves the charges back to where they initially were, the energy stored in the electrostatic field moves back from where it is stored in space to the path of the particle, because that is where is non-zero and $-\mathbf{E}\cdot\mathbf{J} < 0$. It is then returned reversibly to the agent moving the charges back into their original positions. We are neglecting any energy radiated away in this process, which is fine as long as the charge speeds are non-relativisitic, so that the external agent recovers exactly the amount of energy he expended increating the electric fields in the first place.

The amazing thing about electromagnetism is that the fields contain energy, and we can see exactly how the electric energy is (how it is distributed in space) and how it got to where it is.

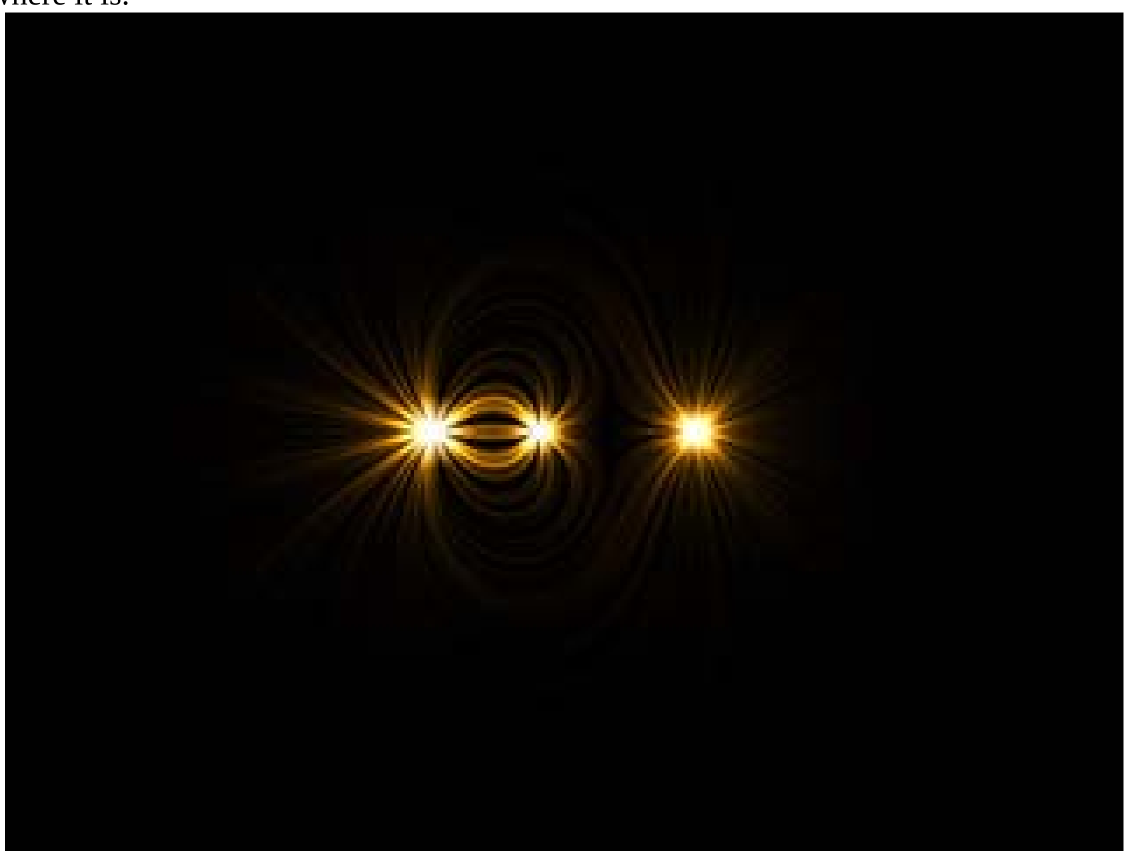


Figure 18-4: An external agent destroying an electric field by bringing together + and – charges.

18.8 Gauss's Law

18.8.1 The general statement

Gauss's Law is extremely useful in solving electrostatic problems with a high degree of symmetry, as we shall see. If we consider any closed surface surrounding a volume, we have from Gauss's Theorem that

$$\int_{volume} \nabla \cdot \mathbf{E}(\mathbf{r})\, d^3x = \int_{surface} \mathbf{E}(\mathbf{r}) \cdot \hat{\mathbf{n}}\ da \qquad (18.8.3)$$

If we use $\nabla \cdot \mathbf{E}(\mathbf{r}) = \rho(\mathbf{r}) / \varepsilon_o$ ((18.4.4), we have for any closed surface that

$$\int_{surface} \mathbf{E}(\mathbf{r}) \cdot \hat{\mathbf{n}}\ da = \frac{1}{\varepsilon_o} \int_{volume} \rho(\mathbf{r})\, d^3x = \frac{Q_{inside}}{\varepsilon_o} \qquad (18.8.4)$$

Equation (18.8.4) is always true, but it is not always useful in solving problems. As an example of this, in Figure 18-5, we show a "Gaussian cylinder" in the presence of two point charges. The surface integral of the electric field dotted into the normal to the surface for this cylinder times *da* is shown by the electric field on the surface of the cylinder evalualted at a number of ponts, where we also indicate the surface normal. In the scenario shown, there is zero flux through the cylinder because there is zero change in the cylinder, even though at every point on the surface there is an electric field. This is an example where Gauss's Law is true but useless in solving a problem.

18.8.2 The field of a point charge from Gauss's Law

To have Gauss's Law actually be useful for problem solving, we need a situation like that shown in Figure 18-6. Here we have a sphere centered on the charge, so that everywhere on the sphere of radius *r* the electric field is radially outward and thus parallel to the surface normal. It is also plausible to assume that the electric field magnitude is only a function of the radius *r*. Thus (18.8.4) becomes

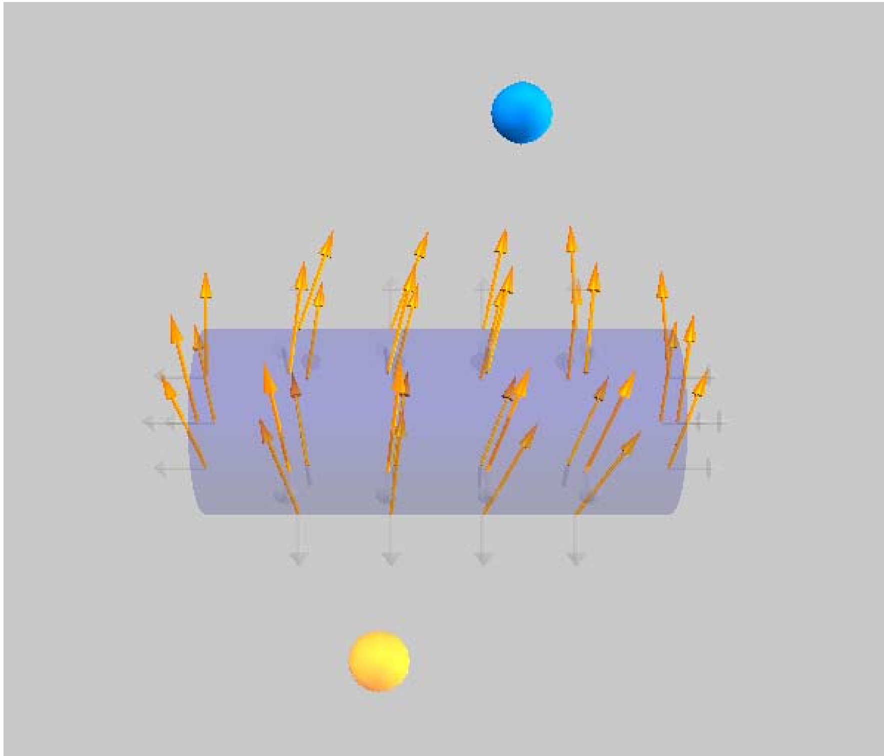


Figure 18-5: Gauss's Law application for point charges

://web.mit.edu/viz/EM/visualizations/electrostatics/GaussLawProblems/

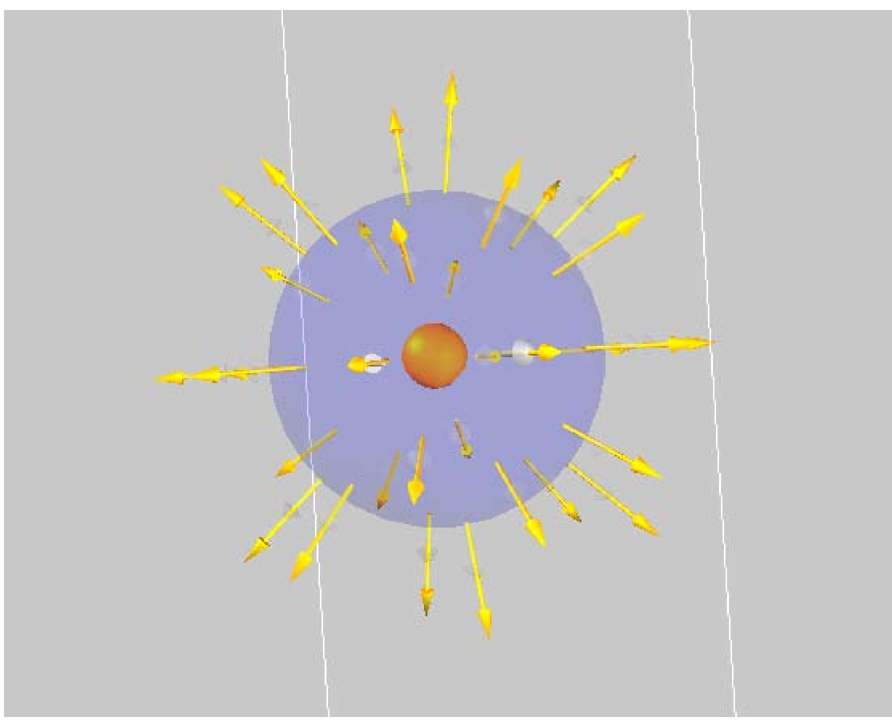


Figure 18-6: Gauss's Law application for a spherical Gaussian surface

$$\int_{surface} \mathbf{E}(\mathbf{r})\cdot\hat{\mathbf{n}}\; da = \frac{q}{\varepsilon_o} = \int_{surface} E(r)\; da = E(r)\int_{surface} da = E(r)4\pi r^2 \qquad (18.8.5)$$

and in this form Gauss's Law has actually allowed us to get the field of a point charge, that is it is radial and varies in magnitude as (18.8.5) prescribes, e.g. $E(r) = q/4\pi\varepsilon_o r^2$.

18.8.3 The field of a line charge from Gauss's Law

Lets do something more difficult with Gauss's Law—the field of an infinite line of charge. We can of course calculate this electric field using the procedure embodied in (18.4.1), but this turns out to be comparatively tedius, whereas the Gauss's Law procedure yields the answer in just a few steps. We must first assume that our electric field is along the cylindrically outward radial direction. Then for our imaginary Gaussian surface we choose a cylinder whose axis is the line of charge, with length *L* and cylindrical radius *r* (see Figure 18-7).

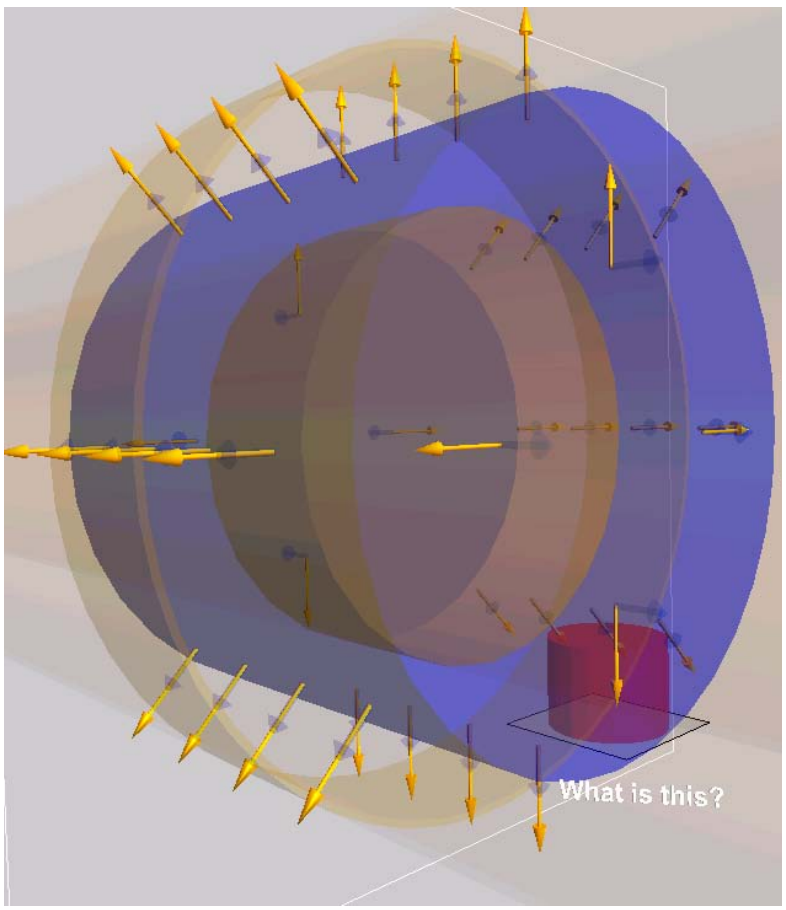


Figure 18-7: An imaginary Gaussian surface (blue cylinder) in a problem with cylindrical symmetry.

Since the cylinder is centered on the line charge, everywhere on the cylinder the electric field is radially outward in the cylindrical sense and thus parallel to the surface normal. It is also plausible to assume that the electric field magnitude is only a function of the cylindrical radius *r*. Thus (18.8.4) becomes (the integrals over the ends of the cylinder vanish because the electric field and the normal are perpendicular there)

$$\int_{sides} \mathbf{E}(\mathbf{r})\cdot\hat{\mathbf{n}}\ da = \frac{\lambda L}{\varepsilon_o} = \int_{sides} E(r)\ da = E(r)\int_{sides} da = E(r)2\pi r L \qquad (18.8.6)$$

And we easily recover from (18.8.6) that the field of a line charge is given in terms of the cylindrical

$$\mathbf{E}(\mathbf{r}) = \frac{\lambda}{2\pi\varepsilon_o r}\hat{\mathbf{r}} \qquad (18.8.7)$$

18.8.4 The field of a plane of charge from Gauss's Law

Finally, lets obtain the electric field of a plane of charge with charge per unit area σ. We choose a Gaussian surface as shown in Figure 18-8.

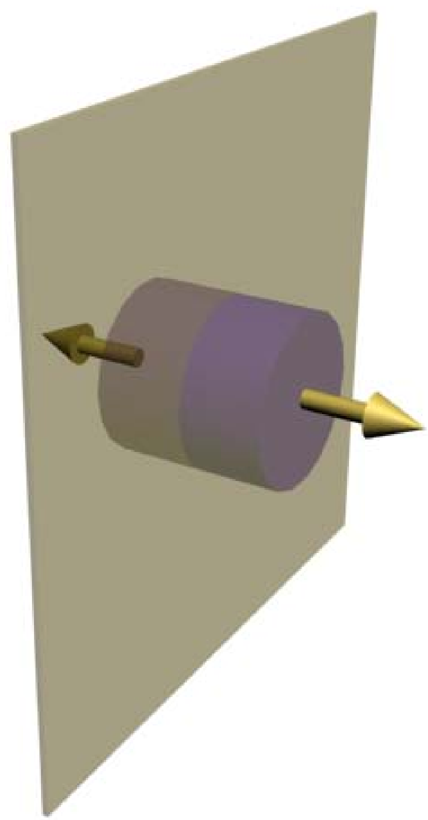


Figure 18-8: The Gaussian surface for a plane of charge.

With this choice and using (18.8.4), we find that

$$\mathbf{E}(\mathbf{r}) = \frac{\sigma}{2\varepsilon_o}\hat{\mathbf{x}}\, sign(x) \tag{18.8.8}$$

where the normal to the plane is $\hat{\mathbf{x}}$.

19 Boundary Value Problems in Electrostatics

19.1 Learning Objectives

We look at boundary value problems in electrostatics. First we define a typical boundary problem, and then we explain why these problems fall into the "hard" category of electromagnetism problems. We then discuss a classic method of solving these problems, the image charge method. The method only works for problems with a high degree of symmetry, so we move on to a method that has more general application, which is separation of variables in various coordinate systems. We only discuss this method for the case of spherical coordinates,

19.2 The Typical Boundary Value Problem

Figure 19-1: The typical boundary value problem

Figure Figure 19-1 shows a typical problem that we want to solve. We have a conductor of some shape, which is held at a fixed potential. We have a point charge with charge q located at some point in space outside the conductor. We want to find the electric field everywhere in space.

19.2.1 Poisson's and Laplace's Equations

Here is the way we go about solving such a problem. We know from (18.4.4) and (18.4.1) that we have

$$\mathbf{E}(\mathbf{r}) = -\nabla\phi(\mathbf{r}) \qquad \text{and} \qquad \nabla\cdot\mathbf{E}(\mathbf{r}) = \rho(\mathbf{r})/\varepsilon_o \tag{19.2.1}$$

These two equations immediately lead to

$$\nabla^2\phi(\mathbf{r}) = -\rho(\mathbf{r})/\varepsilon_o \tag{19.2.2}$$

This equation is known as Poisson's Equation (if $\rho(\mathbf{r}) = 0$ in the region outside of the conductor, it is also known as Laplace's Equation). So we want to find a function $\phi(\mathbf{r})$ which satisfies (19.2.2) for some specified "free" charge distribution $\rho(\mathbf{r})$ outside of the conductor ("free" means we control the location of those charges), with the additional requirement that the potential goes to a specified value on the surface of the conductor. This is a "hard" problem because we don't know where all the charges are. If we knew where all the charges were, we could simply use

$$\phi(\mathbf{r},t) = \frac{1}{4\pi\,\varepsilon_o}\int \frac{\rho\left(\mathbf{r}',t'_{ret}\right)}{\left|\mathbf{r}-\mathbf{r}'\right|}\,d^3x' \tag{19.2.3}$$

in the limit of no time dependence, which is

$$\phi(\mathbf{r}) = \frac{1}{4\pi\,\varepsilon_o}\int \frac{\rho\left(\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|}\,d^3x' \tag{19.2.4}$$

and we would be done. But we cannot use (19.2.4) to calculate the potential, because we ***do not know where all the charges are located***. All we know is where the "free" charges are, and that the charges in the conductor will have arranged themselves to make the conductor an equipotential, but we ***do not know*** *a priori* ***how they have done that***. This is an example of the "hard" electromagnetism: the charges move in response to the fields they create. In this kind of situation, even though (19.2.4) ***is always true***, it is useless in solving the problem, because we have unknowns on both sides of the equation.

19.2.2 Electric fields in and near isolated conductors

Below we will discuss various techniques for dealing with these kinds of problems, but first let's discuss the properties of electric fields in and near isolated conductors. First of all, the electric field inside an isolated conductor will vanish, because if there where any electric field, the charges in the conductor will move so as to cancel out that electric field. Since the electric field is zero inside the conductor, we can conclude using Gauss's Law for any volume inside the conductor, that the charge inside that volume must also be zero. Therefore any charge on a conductor must lie entirely on its surface.

Moreover, the surface of any conductor must be an equipotenial. If it were not an equipotential, that would mean that there would be an electric field on the surface of the conductor which is tangential to the surface of the conductor. But the charges in the conductor will move to cancel out any such tangential electric fields on the surface just as they move to cancel out any electric fields inside the conductor. Therefore an isolated conductor is an equipotential, and the electric field must everywhere be normal to the surface of the conductor. Using Gauss's Law for a small pillbox on the surface of any conductor, we can easily deduce that the component of the electric field normal to the surface, E_n, is related to the surface charge on the conductor by

$$E_n = -\frac{\partial \phi}{\partial n} = -\hat{\mathbf{n}} \cdot \nabla \phi = \frac{\sigma}{\varepsilon_o} \tag{19.2.5}$$

In (19.2.5) we have assumed that the normal $\hat{\mathbf{n}}$ points *out of* the conductor.

19.3 The Uniqueness Theorem

Before we go any further, we pause and prove the Uniqueness Theorem. The Uniqueness Theorem says that if we find by hook or crook any solution to the above problem that has the correct value of $\nabla^2\phi(\mathbf{r})$ in the space outside of the conductor and has the specified value on the surface of the conductor, then that is the solution, because there is one and only one such solution. The usefulness of this theorem is that it allows us to find a solution using any kind of trick we can think of (the method of images discussed below is such a trick), and we know that if we have one, we are done, it is the only possible solution.

Proof of the Uniqueness Theorem: Suppose we have two solutions to our boundary value problem, $\phi_1(\mathbf{r})$ and $\phi_2(\mathbf{r})$ that satisfy our requirements. That is,

$$\nabla^2\phi_1(\mathbf{r}) = -\rho(\mathbf{r}) / \varepsilon_o \quad \text{and} \quad \nabla^2\phi_2(\mathbf{r}) = -\rho(\mathbf{r}) / \varepsilon_o \tag{19.3.1}$$

$$\phi_1(\mathbf{r})\Big|_{\mathbf{r}\text{ on surface of condutor}} = \phi_2(\mathbf{r})\Big|_{\mathbf{r}\text{ on surface of condutor}} = \text{given function} \tag{19.3.2}$$

Define the function $W(\mathbf{r})$ to be the difference between these two solutions.

$$W(\mathbf{r}) = \phi_2(\mathbf{r}) - \phi_1(\mathbf{r}) \tag{19.3.3}$$

Then we must have that

$$\nabla^2 W = 0 \text{ outside of conductor} \quad \text{and} \quad \mathrm{W}(\mathbf{r})\Big|_{\text{r on surface of conductor}} = 0 \tag{19.3.4}$$

We have the general vector identity for any scalar potential that

$$\nabla \cdot (W \nabla W) = (\nabla W)^2 + W \nabla^2 W \tag{19.3.5}$$

We apply Gauss's Theorem to $\nabla \cdot (W \nabla W)$ for the volume outside of the conductor and the surface of the conductor, as follows

$$\int_{\text{vol where } \rho \text{ defined}} \nabla \cdot (W \nabla W) d^3x = \int_{\text{bounding surface}} (W \nabla W) \cdot \hat{\mathbf{n}}\, da \tag{19.3.6}$$

But we know that W is zero on the surface of the conductor, so we have

$$\int_{\text{vol where } \rho \text{ defined}} \nabla \cdot (W \nabla W) d^3x = \int_{\text{vol where } \rho \text{ defined}} \left[(\nabla W)^2 + W \nabla^2 W \right] d^3x = 0 \tag{19.3.7}$$

where we have used (19.3.5) to get the final form in (19.3.7). But we have that $\nabla^2 W = 0$ in the volume, so we must have

$$\int_{\text{vol where } \rho \text{ defined}} \left[(\nabla W)^2 \right] d^3x = 0 \tag{19.3.8}$$

Since $(\nabla W)^2$ is positive definite, we must have that everywhere in the volume ∇W =0. Thus W can only be a constant, and since it is zero on the surface of the conductor, that constant must be zero. Therefore $\phi_1(\mathbf{r})$ and $\phi_2(\mathbf{r})$ are the same, and the solution is unique.

19.4 The Method of Images

The method of images is basically a trick that takes advantage of uniqueness. It can only be used in situations with high degrees of symmetry. Basically you take the problem as stated, which we call Problem I, which has a conductor with some unknown surface charge induced by the presence of the known free charge. You find a ***different free space*** problem, which we call Problem II, ***which has no conductor***, where you arrange charges so that you get the proper value of the potential on the surface of the conductor from Problem I. We then take that part of Problem II relevant to Problem I and transfer it back to Problem I. Since it satisfies all the conditions for Problem I, we know it must be the solution to Problem I. The best way to illustrate how this works is to take examples, as follows.

19.4.1 Point Charge and a Conducting Plane

Suppose a point charge *q* is on the z-axis a distance *d* above the origin. The *xy* plane at $z = 0$ is an infinite grounded conducting plane. This is Problem I, as illustrated in Figure 19-2.

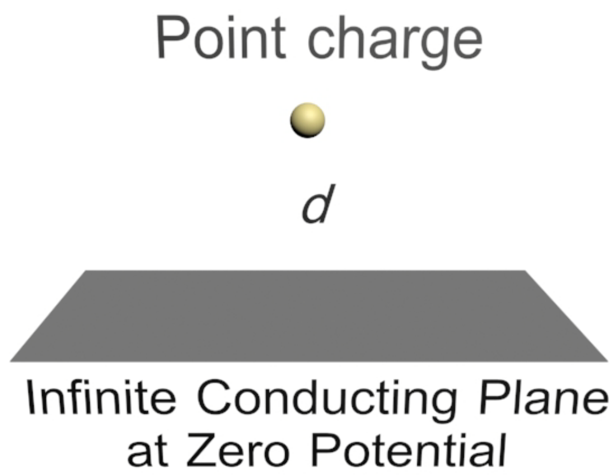


Figure 19-2: Problem I: A point charge above in infinite grounded couducting plane.

Here is Problem II, as illustrated in Figure 19-3. What we are doing with Problem II is to duplicate the charge distribution for $z > 0$ that we have in Problem I, and arrange charges in the region where the conductor is in Problem I ($z \leq 0$) so that we satisfy the boundary conditions at $z = 0$ from Problem I, that is zero potential at $z = 0$. We simply guess that if we put a charge of $-q$ in Problem II a distance *d* down the negative z-axis, we will get zero potential in Problem II at $z = 0$.

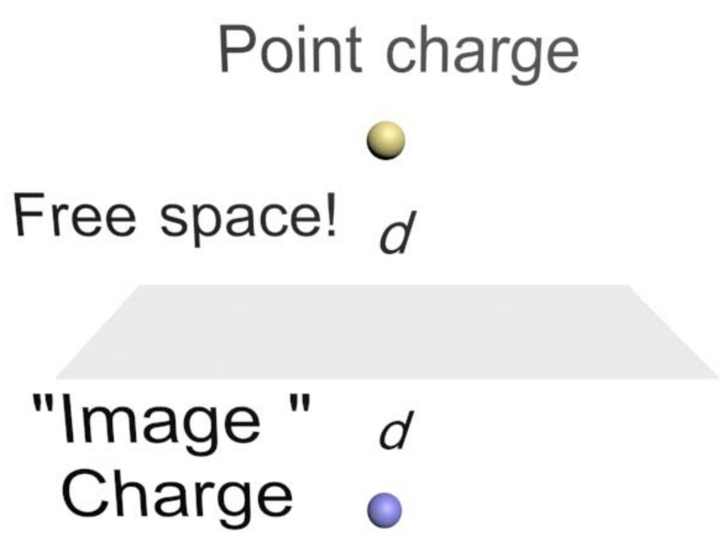


Figure 19-3: Problem II: A free space problem with two charges

And indeed this is the case. If we write down the solution for the potential for Problem II, it is

$$\phi_{II}(\mathbf{r}) = \frac{1}{4\pi\varepsilon_o}\left[\frac{q}{\sqrt{x^2+y^2+(z-d)^2}} - \frac{q}{\sqrt{x^2+y^2+(z+d)^2}}\right] \tag{19.4.1}$$

This potential has the right behavior at $z = 0$ (it is zero there), and

$$\nabla^2\phi_{II}(\mathbf{r}) = q\left[\delta^3(\mathbf{r}-d\,\hat{\mathbf{z}}) - \delta^3(\mathbf{r}+d\,\hat{\mathbf{z}})\right]/\varepsilon_o \tag{19.4.2}$$

so in the upper half plane it also has the right value for $\nabla^2\phi(\mathbf{r})$, a delta function at the position of the point charge q. It also has a delta function at the position of the negative point charge, but that is not in the upper half plane and we don't care about that one. That is why we are free to put any charges for z < 0 that we want to try to satisfy the boundary condition at $z = 0$—they don't count.

So the Uniqueness Theorem guarantees that if we simply take over to Problem I the solution to Problem II in the upper half plane, and put the potential for Problem I to zero for $z < 0$, then we have the unique solution to Problem I, which is

$$\phi_I(\mathbf{r}) = \begin{cases} \dfrac{1}{4\pi\varepsilon_o}\left[\dfrac{q}{\sqrt{x^2+y^2+(z-d)^2}} - \dfrac{q}{\sqrt{x^2+y^2+(z+d)^2}}\right] & z > 0 \\ 0 \quad z \le 0 \end{cases} \tag{19.4.3}$$

Once we have the solution to Problem I, we can go back and see what the mysterious surface charge induced on the conducting plane by the presence of the free charge q at $d\,\hat{\mathbf{z}}$ is. Using (19.2.5) we see that

$$\sigma = -\varepsilon_o \frac{\partial\phi_I}{\partial n} \qquad \text{or} \qquad \sigma = -\varepsilon_o \left.\frac{\partial\phi_I}{\partial z}\right|_{z=0+} \tag{19.4.4}$$

But

$$\left.\frac{\partial\phi_I}{\partial z}\right|_{z\ge 0} = \frac{1}{4\pi\varepsilon_o}\left[-\frac{q(z-d)}{\left(x^2+y^2+(z-d)^2\right)^{3/2}} + \frac{q(z+d)}{\left(x^2+y^2+(z+d)^2\right)^{3/2}}\right] \tag{19.4.5}$$

so

$$\sigma(x,y) = -\frac{1}{2\pi}\frac{qd}{\left(x^2+y^2+d^2\right)^{3/2}} \tag{19.4.6}$$

If we want to know the total induced charge on the plane, it is

$$\int_0^\infty dr \int_0^{2\pi} r d\phi \sigma = -\int_0^\infty \frac{dr\left(rqd\right)}{\left(r^2+d^2\right)^{3/2}} = +\frac{qd}{\left(r^2+d^2\right)^{1/2}}\Bigg|_0^\infty = -q \qquad (19.4.7)$$

19.4.2 Point Charge outside a Conducting Grounded Spherical Shell

We give one more example of using image charges. A point charge of charge q is located at position $\mathbf{r}'$ outside of a grounded conducting sphere of radius R. The point charge is a distance $r' > R$ from the center of the shell. Using the method of images, we find that the potential outside the sphere is the potential due to the point charge plus the potential due to an image charge with charge $-qR/r'$ located inside the shell at a distance R^2/r'. Therefore our solution is

$$\phi\left(\mathbf{r}\right) = \begin{cases} \dfrac{1}{4\pi\varepsilon_o}\left[\dfrac{q}{\left|\mathbf{r}-\mathbf{r}'\right|} - \dfrac{qR}{r'}\dfrac{1}{\left|\mathbf{r}-\dfrac{\mathbf{r}'R^2}{\left(r'\right)^2}\right|}\right] & r > R \\ 0 & r \le R \end{cases} \qquad (19.4.8)$$

19.5 Separation of Variables in Spherical Coordinates

Now we turn to another standard technique, which has much wider usage than the image charge method. What we do is look at solutions to Laplaces's equation, $\nabla^2\phi(\mathbf{r}) = 0$, and investiage the form of solutions to this equation in various coordinate systems. Although we could do this in cartesion and cylindrical coordinates, we will only look at spherical coordinates in these notes, as it is illustrative of the general technique. You might ask what good finding solutions to $\nabla^2\phi(\mathbf{r}) = 0$ is. But in many problems, $\nabla^2\phi(\mathbf{r})$ is zero almost everywhere, and we try to piece together a solution to a general problem, which charge located in limited regions, out of solutions to $\nabla^2\phi(\mathbf{r}) = 0$, appropriately chosen.

19.5.1 A Typical Problem Involving Seperation of Variables

Here is a typical boundary value problem where separation of variables is useful. We have a sphere of radius R, and on the surface of the sphere we know the values of the potential. That is, some one has given us the function $f\left(\theta,\phi\right)$, such that on the surface of the sphere

$$\phi(r,\theta,\phi)\big|_{r=R} = f(\theta,\phi) \tag{19.5.1}$$

We also know that the potential vanishes at infinity, and that is no free charge for $r > R$. Given all this, we want to find the potential for all $r \geq R$. Since our boundary condition is given in spherical coordinates, we see if we can find a solution by adding up many solutions to $\nabla^2\phi(\mathbf{r}) = 0$ in spherical coordinates. If we can satisfy all of the boundary conditions, then uniqueness tells us that out solution is the only solution, no matter how we come by it.

Laplace's equation in spherical polar coordinates is

$$\nabla^2\phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\phi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\phi}{\partial^2\phi} = 0 \tag{19.5.2}$$

We are going to assume that our solution $\phi(r,\theta,\phi)$ is separable, that is, that it can be written as

$$\phi(r,\theta,\phi) = R(r)P(\theta)W(\phi) \tag{19.5.3}$$

The reason we might think it is separable in these coordinates is that our boundary condition is in these coordinates. But the ultimate rationale is that we can find solutions when we make this assumption that satisfy our boundary conditions, and $\nabla^2\phi(\mathbf{r}) = 0$, and uniqueness tells us that this is the solution. If we insert (19.5.3) into (19.5.2) and divide by $\dfrac{r^2\sin^2\theta}{R(r)P(\theta)W(\phi)}$, we have

$$\frac{\sin^2\theta}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{\sin\theta}{P}\frac{d}{d\theta}\left(\sin\theta\frac{dP}{d\theta}\right) + \frac{1}{W}\frac{d^2W}{d^2\phi} = 0 \tag{19.5.4}$$

We have isolated all of the ϕ dependence in (19.5.4) in the last term, and the only way this can be true is if there is a constant m^2 such that

$$\frac{1}{W}\frac{d^2W}{d^2\phi} = -m^2 \tag{19.5.5}$$

The solutions to (19.5.5) are $\sin m\phi$ or $\cos m\phi$. In order for W to be single valued, we must have that m be an integer. Our remaining equation is

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)+\frac{1}{P\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dP}{d\theta}\right)-\frac{m^2}{\sin^2\theta}=0 \qquad (19.5.6)$$

where we have divided by $\sin^2\theta$ so as to again isolate all of the *r* dependence in the first term of (19.5.6). Again, this means that there must be a constant which we call $l(l+1)$ such that

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)=l(l+1) \qquad (19.5.7)$$

The solutions $R(r)$ are

$$R(r)=Ar^l+Br^{-l-1} \qquad (19.5.8)$$

This leaves the remaining equation for *P* as

$$\frac{1}{P\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dP}{d\theta}\right)+l(l+1)-\frac{m^2}{\sin^2\theta}=0 \qquad (19.5.9)$$

If we make the substitution $x=\cos\theta$, the equation for P becomes

$$\frac{d}{dx}\left[\left(1-x^2\right)\frac{dP}{dx}\right]+\left[l(l+1)-\frac{m^2}{1-x^2}\right]P=0 \qquad (19.5.10)$$

This is the generalized Legendre equation and its solutions are called the associated Legendre polynomials.

The Legendre functions are the solutions to the above equation with $m = 0$, that is

$$\frac{d}{dx}\left[\left(1-x^2\right)\frac{dP}{dx}\right]+l(l+1)P=0 \qquad (19.5.11)$$

If we impose the requirement that solutions to the above equation converge for $x^2\le 1$, then this requires that *l* be zero or a positive integer[5]. By convention our functions are normalized to have the value of unity at $x=+1$. The first few solutions, the Legendre polynomials, are given by

$$P_o(x)=1 \qquad P_1(x)=x \qquad P_2(x)=\frac{1}{2}\left(3x^2-1\right)$$
$$P_3(x)=\frac{1}{2}\left(5x^3-3x\right) \qquad P_4(x)=\frac{1}{8}\left(35x^4-30x^2+3\right) \qquad (19.5.12)$$

---

[5] See Jackson 2nd Edition, page 86.

In general one can show that

$$P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}\left(x^2 - 1\right)^l \tag{19.5.13}$$

This formula is known as *Rodrigues' formula*. From Rodriques' formula it can be shown that

$$\frac{dP_{l+1}}{dx} - \frac{dP_{l-1}}{dx} - (2l+1)P_l = 0 \tag{19.5.14}$$

This equation together with the differential equation can be use to show that

$$P_{l+1} = \frac{(2l+1)x\,P_l - lP_{l-1}}{\left(l+1\right)} \tag{19.5.15}$$

This recursion relation for Legendre polynomials is very useful for any numerical work with the Legendre polynomials because you need only know that $P_0(x) = 1$ and $P_1(x) = x$ and you can find the value of $P_l(x)$ at any *x* for $l > 1$ by applying this recursion relation $l-1$ times.

If the solutions to the associated Legendre equation, where *m* is not equal to zero, are to converge for $x^2 \le 1$, we find similarly that *l* must be zero or a positive integer and that the integer m can only take on the values –*l*, -(*l*-1),….,0….,(*l*-1), *l*. The associated Legendre function $P_l^m(x)$ is given by

$$P_l^m(x) = \left(-1\right)^m \left(1 - x^2\right)^{m/2} \frac{d^m}{dx^m} P_l\left(x\right) \tag{19.5.16}$$

Since the differential equation depends only on m^2, P_l^m and P_l^{-m} must be proportional, and it can be shown that

$$P_l^{-m}(x) = \left(-1\right)^m \frac{\left(l-m\right)!}{(l+m)!} P_l^m(x) \tag{19.5.17}$$

$$(\ell - m + 1)P_{\ell+1}^m(x) - (2\ell+1)xP_\ell^m(x) + (\ell+m)P_{\ell-1}^m(x) = 0 \tag{19.5.18}$$

$$P_{\ell+1}^m(x) = \frac{(2\ell+1)xP_\ell^m(x) + (\ell+m)P_{\ell-1}^m(x)}{(\ell - m + 1)} \tag{19.5.19}$$

$$(\ell - m + 1)P_{\ell+1}^m(x) + (1-x^2)^{1/2} P_\ell^{m+1}(x) - (\ell+m+1)xP_\ell^m(x) = 0 \tag{19.5.20}$$

For a given *m*, the set of functions $\left\{P_l^m(x)\right\}$ is a complete set of functions, with normalization given by

$$\int_{-1}^{1} P_l^m(x)P_{l'}^m(x)dx = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\delta_{ll'} \tag{19.5.21}$$

19.5.2 Spherical Harmonics

It is usual to combine the θ and ϕ to define the spherical harmonics $Y_{lm}(\theta,\phi)$ defined by

$$Y_{lm}(\theta,\phi) = \sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{+im\phi} \tag{19.5.22}$$

The Y_{lm}'s form a complete set of functions in θ and ϕ. The normalization and orthorgonality conditions are

$$\int_0^{\pi}\sin\theta d\theta\int_0^{2\pi} d\phi Y_{lm}^*(\theta,\phi)Y_{l'm'}(\theta,\phi) = \delta_{l'l}\delta_{m'm} \tag{19.5.23}$$

We can expand any function $g(\theta,\phi)$ as

$$g(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} A_{lm}Y_{lm}(\theta,\phi) \tag{19.5.24}$$

where

$$A_{lm} = \int_0^{\pi}\sin\theta d\theta\int_0^{2\pi} d\phi\, Y_{lm}^*(\theta,\phi)\, g(\theta,\phi) \tag{19.5.25}$$

With the definition of the Y_{lm}'s and (19.5.8), we thus see that the most general form of the solutions to $\nabla^2\phi(\mathbf{r}) = 0$ that is separable in spherical coordinates is given by

$$\phi(r,\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[A_{lm}r^l + B_{lm}r^{-l-1}\right]Y_{lm}(\theta,\phi) \tag{19.5.26}$$

19.5.3 The Solution to the Typical Problem

We now return to our "typical" problem above, and see if we can find a solution to that problem that is of the form given in (19.5.26). If we want to prevent the potential from blowing up at infinity, we must take all of our A_{lm} to be zero. To satisfy our boundary condition at the surface of the sphere (19.5.1) we must have

$$f(\theta,\phi) = \phi(r,\theta,\phi)\Big|_{r=R} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left[B_{lm}R^{-l-1}\right]Y_{lm}(\theta,\phi) \tag{19.5.27}$$

But since we know that the Y_{lm}'s are complete, we know this can be done, and in fact we can write down the coefficients as follows

$$B_{lm} = R^{l+1}\int_0^{\pi}\sin\theta d\theta\int_0^{2\pi}d\phi\, Y_{lm}^{*}(\theta,\phi)\, f(\theta,\phi) \tag{19.5.28}$$

And we are done. We have solved the "typical" problem involving separation of variables in spherical coordinates.

19.5.4 Azimuthal Symmetry

Let us suppose that we have a problem with azimuthal symmetry. Then our integer m above must be zero, and our complete solution given by (19.5.26) reduces to

$$\phi(r,\theta) = \sum_{l=0}^{\infty}\left(A_l r^l + \frac{B_l}{r^{l+1}}\right)P_l(\cos\theta) \tag{19.5.29}$$

Our normalization condition for the P_l's in this case (cf. (19.5.21) for $m = 0$) is

$$\int_{-1}^{1}P_l(x)P_{l'}(x)dx = \frac{2}{2l+1}\delta_{ll'} \tag{19.5.30}$$

It is clear that we can do any potential problem where the potential at $r = R$ is given as a function of θ using the same technique as above. But there are other interesting things to do as well.

Consider the following problem. We put a surface charge $\sigma(\theta) = \sigma_o P_n(\cos\theta)$ on the surface of a spherical shell. There are no charges inside or outside the shell other than these charges. What is the potential and the electric field everywhere? The lack of charges other than at the shell means that the potential must vanish at infinity and not blow up at the origin. Since we also know that the potential must be continous across $r = R$, we therefore we must have

$$\phi(r,\theta)=\begin{cases}\sum_{l=0}^{\infty}A_l\left(\frac{r}{R}\right)^l P_l(\cos\theta)\\ \sum_{l=0}^{\infty}A_l\left(\frac{R}{r}\right)^{l+1} P_l(\cos\theta)\end{cases} \tag{19.5.31}$$

Gauss's Law applied to a pill box on the surface of the sphere yields

$$E_{r=R+}-E_{r=R-}=\sigma(\theta)/\varepsilon_o=\left.\frac{\partial\phi}{\partial r}\right|_{r=R-}-\left.\frac{\partial\phi}{\partial r}\right|_{r=R+}=\sum_{l=0}^{\infty}\left(\frac{A_l l}{R}+\frac{A_l(l+1)}{R}\right)P_l(\cos\theta) \tag{19.5.32}$$

$$\sigma_0 P_n(\theta)/\varepsilon_o=\sum_{l=0}^{\infty}\frac{A_l(2l+1)}{R}P_l(\cos\theta) \tag{19.5.33}$$

Since the P_l's are orthogonal, we must there have all the A_l's are zero except A_n, with $A_n=\frac{\sigma_0 R}{(2n+1)\varepsilon_o}$. Therefore our solution (19.5.31) is

$$\phi(r,\theta)=\frac{\sigma_0 R}{\varepsilon_o(2n+1)}\begin{cases}\left(\frac{r}{R}\right)^n P_n(\cos\theta) & r<R\\ \left(\frac{R}{r}\right)^{n+1} P_n(\cos\theta) & r>R\end{cases} \tag{19.5.34}$$

As an example of (19.5.34), suppose we take $n = 0$, that is the surface charge density on the surface of the sphere is constant. Then we have for this case

$$\phi(r,\theta)=\frac{\sigma_0 R}{\varepsilon_o}\begin{cases}1 & r<R\\ \left(\frac{R}{r}\right) & r>R\end{cases}=\frac{\sigma_0 4\pi R^2}{4\pi\varepsilon_o}\begin{cases}\frac{1}{R} & r<R\\ \frac{1}{r} & r>R\end{cases}=\frac{q_o}{4\pi\varepsilon_o}\begin{cases}\frac{1}{R} & r<R\\ \frac{1}{r} & r>R\end{cases} \tag{19.5.35}$$

where $q_o=4\pi R^2$. This is what we expect for a uniformly charged spherical shell.

As a second example of (19.5.34), suppose we take $n = 1$, that is the surface charge density on the surface of the sphere goes as $\cos\theta$. Then we have for this case

$$\phi(r,\theta)=\begin{cases}\frac{\sigma_0}{3\varepsilon_o}\,r\cos\theta & r<R\\ \frac{\sigma_0 R^3}{3\varepsilon_o}\frac{\cos\theta}{r^2} & r>R\end{cases} \tag{19.5.36}$$

This potential gives the following electric field

$$\mathbf{E}(r,\theta)=\phi(r,\theta)=\begin{cases}-E_0\hat{\mathbf{z}} & r<R \\ \frac{p_0}{4\pi\varepsilon_o}\left(\frac{2\cos\theta}{r^3}\hat{\mathbf{r}}+\frac{\sin\theta}{r^3}\hat{\boldsymbol{\theta}}\right) & r>R\end{cases} \tag{19.5.37}$$

where $E_o=\frac{\sigma_0}{3\varepsilon_o}$ and $p_o=\frac{4\pi\sigma_0 R^3}{3}$. Thus we have a constant and downward field inside the sphere and a perfect dipole outside the sphere, if our surface charge goes as $\cos\theta$.

19.6 Boundary Conditions on the Electric Field

Before leaving electrostatics, we discuss the boundary conditions that must be true across any thin interface in electrostatics. From Gauss'a Law, $\nabla\cdot\mathbf{E}(\mathbf{r})=\rho(\mathbf{r})/\varepsilon_o$, we see that any change in the normal component of the electric field across a thin interface must be given by

$$E_{2n}-E_{1n}=\hat{\mathbf{n}}\cdot(\mathbf{E}_2-\mathbf{E}_1)=\sigma/\varepsilon_o \tag{19.5.38}$$

where the normal $\hat{\mathbf{n}}$ is assumed to point *from* 1 *to* 2.

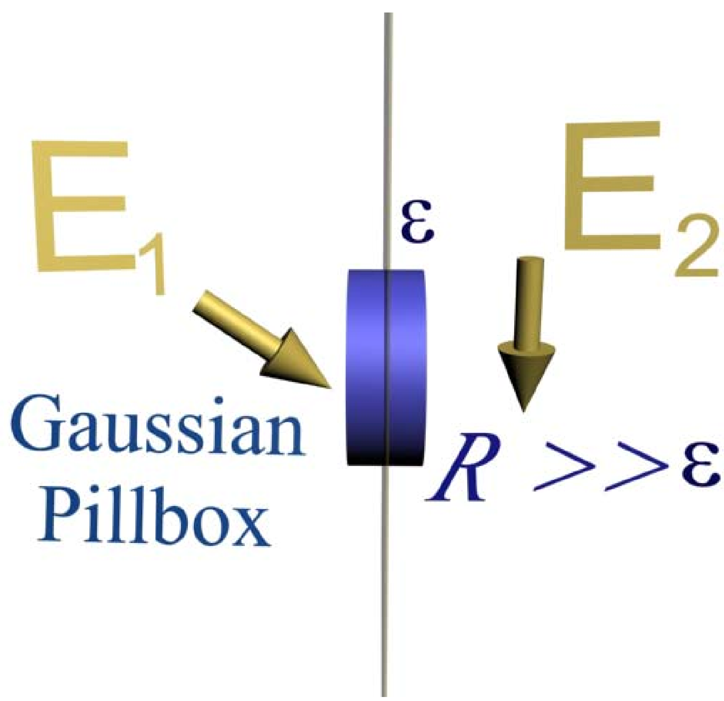


Figure 19-4: Boundary condition on the normal component of E across an interface

To make absolutely sure we understand where (19.5.38) comes from, consider the figure above. We apply Gauss's Law to the pillbox shown, making sure that the radius R of the pillbox is arbitrarily large compared to its height ε. The area of the sides of the pillbox is therefore given by $2\pi R\varepsilon$ and the area of the ends is given by πR^2. By making ε small enough compared to R, we can insure that any flux through the sides of the pillbox due to tangential components of **E** do not contribute to the integral over the surface area of the pillbox. Only the normal components of **E** contribute to the integral, and we easily obtain (19.5.38).

What about the tangential components of **E**? These must be continuous, even in a time varying situation. Take an amperean loop that spans the interface, with width ε perpendicular to the interface and length l tangential to the interface, as shown in the figure below. If we use Faraday's Law, and integrate around this loop, we have

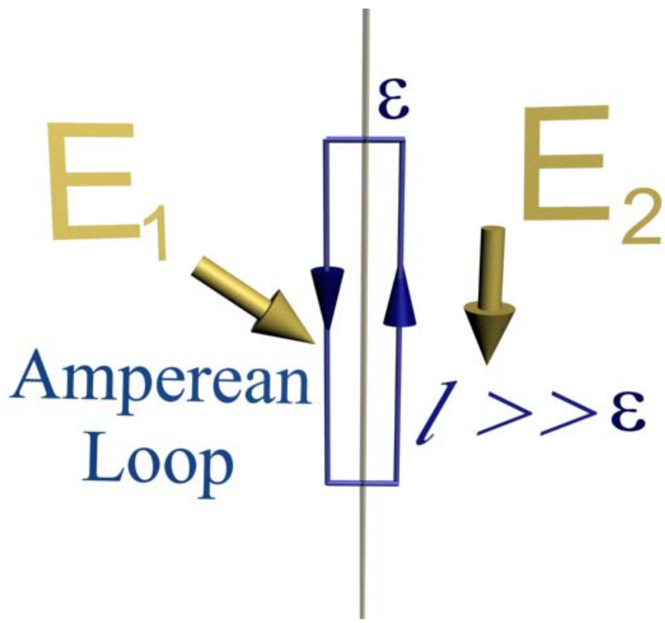


Figure 19-5: Boundary condition on the tangential component of E across an interface.

$$\oint \mathbf{E} \cdot d\,\mathbf{l} = \int_{\text{open surface}} \frac{\partial \mathbf{B}}{\partial t} \cdot \hat{\mathbf{n}}\, da \qquad (19.5.39)$$

If we make the width ε very small compared to the length l, the only component that will enter into the left hand side of (19.5.39) will be the tangential **E**, since any normal component will be multiplied by $\varepsilon << l$. And the magnitude of that component will be $E_t\, l$. If we look at the right hand side of (19.5.39), it involves an area $l\varepsilon$, and again if we make $\varepsilon << l$ we can make the area integral insignificant, even in a time varying situation. Therefore we will always have for the electric field that

$$\mathbf{E}_{2t} = \mathbf{E}_{1t} \tag{19.5.40}$$

20 Basic Magnetostatics

20.1 Learning Objectives

We now consider magnetostatics, and in particular concentrate on how to calculate magnetic fields using the Biot-Savart Law and Ampere's Law.

20.2 The relation between $\mathbf{J}\,d^3x$ and $I\;d\mathbf{l}$

In magnetostatics, we do a lot of switching back and forth between volume integrals involving $\mathbf{J}(\mathbf{r}')\,d^3x'$ and line integrals involving $I\,d\mathbf{l}'$, since many currents are carried by wires. The correspondence is as follows. If we have a wire carrying current *I*, then the current density **J** is always parallel to the local tangent to the wire, $I\,d\mathbf{l}$. Consider a small segment of the wire centered at point *Q* (see Figure 20-1). With

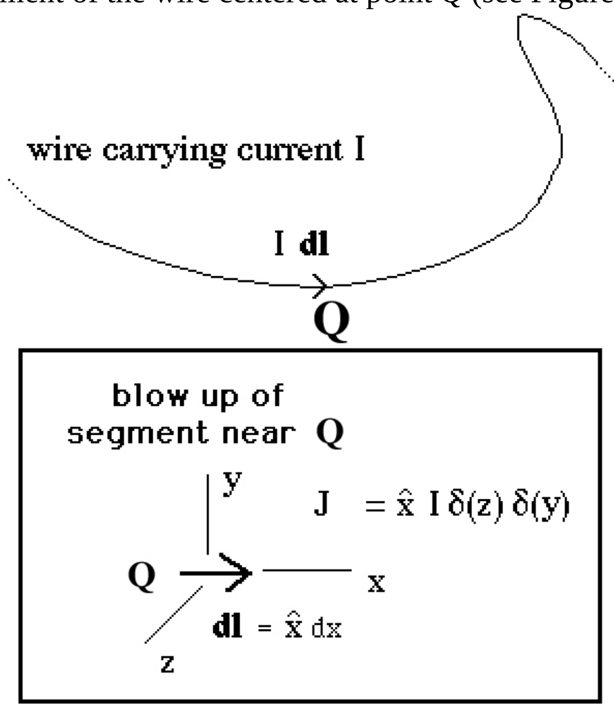


Figure 20-1: The relation between J and $I\,d\mathbf{l}$

no loss of generality, we can set up a local coordinate system where $d\mathbf{l}$ is along the x-axis. In that coordinate system, near the point *Q*, the current density **J** is given by $\mathbf{J} = \hat{\mathbf{x}}\,I\,\delta(z)\,\delta(y)$. Note that since the delta functions have dimensions of inverse argument, this expression for **J** has the correct units, that is, coulombs/m^{2}s. If we look at the integral of **J** over a small volume enclosing point *Q*, we have

$$\int_{volume} \mathbf{J}\,d^3x = \int_{volume} \hat{\mathbf{x}}\,I\,\delta(x)\,\delta(y)\,dx\,dy\,dz = \int_{line} I\,\hat{\mathbf{x}}\,dz = \int_{line} I\,d\,\mathbf{l} \tag{20.2.1}$$

where we have used the delta functions to do two of the spatial integrations, and in the last step we have used the fact that $\hat{\mathbf{x}}\, dx = d\,\mathbf{l}$. Thus the correspondence we want is that the volume integral of $\mathbf{J}\ d^3x$ goes over to a line integral of $I\,d\mathbf{l}$ when wires carry the current. Conversely, if we have a formula that involves a line integral of $I\,d\mathbf{l}$, we can generalize it to a volume integral by replacing $I\,d\mathbf{l}$ by $\mathbf{J}\ d^3x$ For example, in equation (8.2.14) we defined the magnetic dipole moment of a distribution of currents to be $\mathbf{m} = \frac{1}{2}\int \mathbf{r}'\times\mathbf{J}(\mathbf{r}')\,d^3x'$. We can use the prescription above to also write this as a line integral, $\mathbf{m} = \frac{1}{2} I \oint \left(\mathbf{r}'\times d\mathbf{l}'\right)$.

20.3 The Biot Savart Law

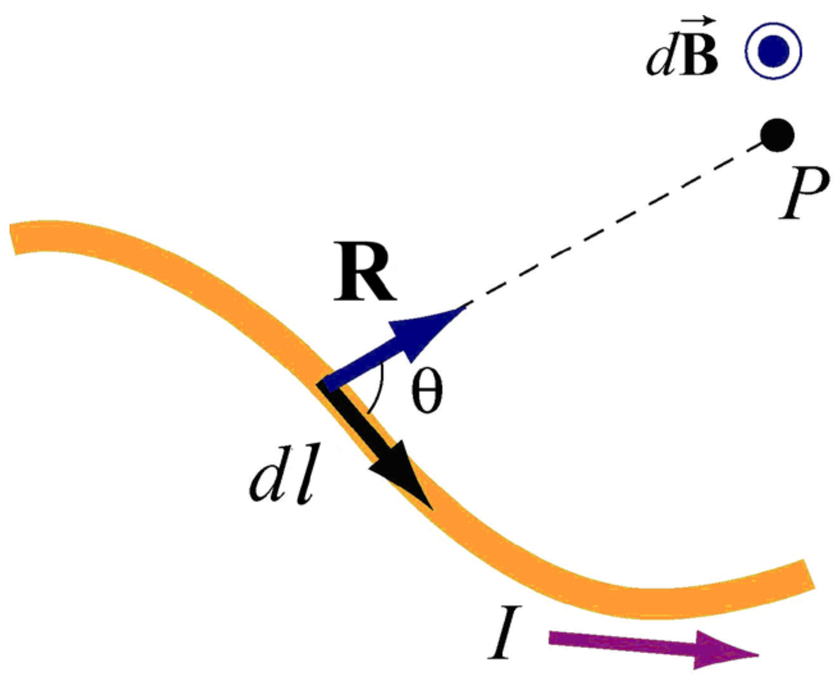


Figure 20-2: The Biot-Savart Law

The Biot-Savart Law tells us how magnetic fields are generated by steady currents. In reference to Figure 20-2, if **R** is the vector from the current carrying element $d\mathbf{l}$ to the point *P*, then the magnetic field d**B** generated at the point *P* by the current element $d\mathbf{l}$ is given by

$$d\mathbf{B} = \frac{\mu_0}{4\pi}\frac{I\,d\mathbf{l}\times\mathbf{R}}{R^3} \qquad (20.3.1)$$

If the current element $d\mathbf{l}$ is located at $\mathbf{r}'$ and the observer at point P is located at $\mathbf{r}$, then this equation is

$$d\mathbf{B}(\mathbf{r}) = \frac{\mu_o I}{4\pi} d\mathbf{l}' \times \frac{(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3} \tag{20.3.2}$$

where we have introduced a prime on $d\mathbf{l}$ to indicate that it is located at $\mathbf{r}'$. To compute the total magnetic field at point *P* due to the complete wire, we evaluate

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_o I}{4\pi} \oint d\mathbf{l} \times \frac{(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3} \tag{20.3.3}$$

Taking the cross-product in (20.3.2) is a particularly hard process to imagine in one's head, and I urge you to go to the Shockwave visualization show below to get a firm grasp of how this works.

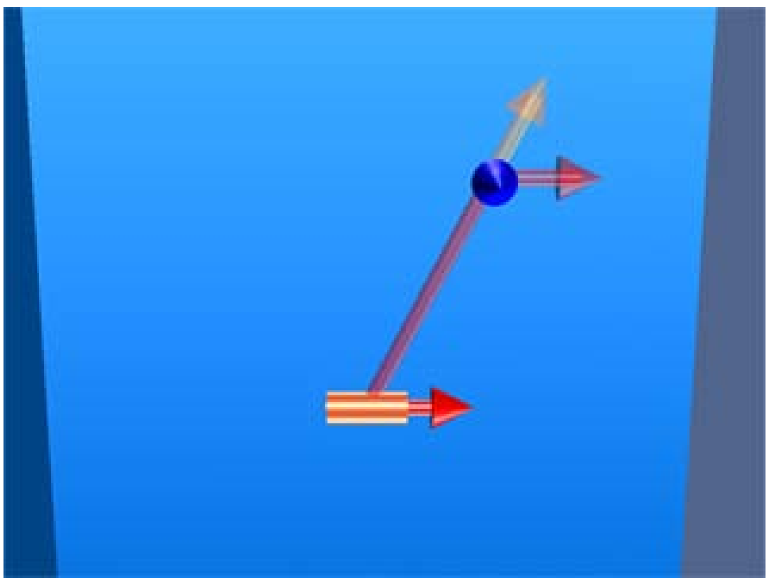


Figure 20-3: Shockwave visualization to illustrate the nature of Biot-Savart

://web.mit.edu/viz/EM/visualizations/magnetostatics/MagneticFieldConfigurations/CurrentElement3d/CurrentElement.htm

If we go over to a continuous distribution of current density, then our expression in (20.3.3) becomes

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_o}{4\pi} \int_{volume} d^3x' \, \mathbf{J}(\mathbf{r}') \times \frac{(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3} \tag{20.3.4}$$

If we go back to our proof of the Helmholtz Theorem, we see that we can write

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r}) \tag{20.3.5}$$

where

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_o}{4\pi} \int_{volume} d^3x' \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{20.3.6}$$

We can also deduce from (20.3.5) that

$$\nabla \cdot \mathbf{B}(\mathbf{r}) = 0 \tag{20.3.7}$$

and from (20.3.6) that

$$\nabla \times \mathbf{B}(\mathbf{r}) = \mu_o \mathbf{J} \tag{20.3.8}$$

20.4 The magnetic field on the axis of a circular current loop

A circular loop of radius R in the xy plane carries a steady current I, as shown in Figure 20-4. We want to use the Biot-Savart Law to calculate the magnetic field on the axis of the loop.

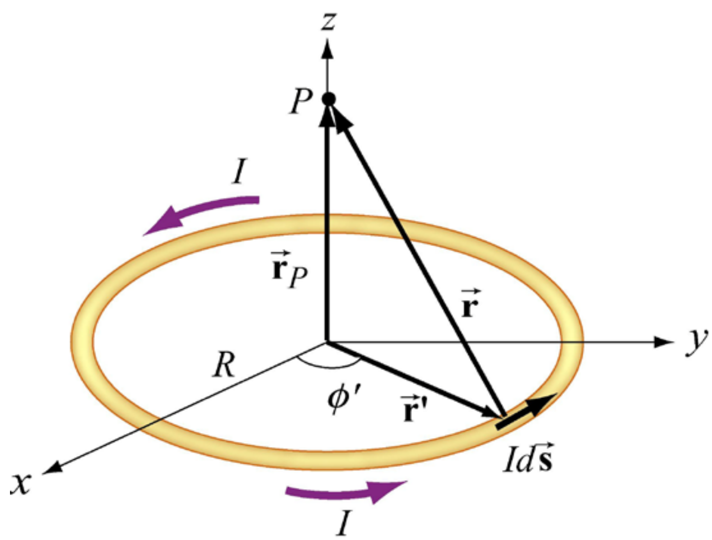


Figure 20-4: Calculating the magnetic field on the axis of a circular current loop

In Cartesian coordinates, the differential current element located at

$$\mathbf{r}' = R(\cos\phi'\,\hat{\mathbf{i}} + \sin\phi'\,\hat{\mathbf{j}}) \tag{20.4.1}$$

And $Id\mathbf{l}$ can be written as

$$I d\mathbf{l} = IRd\phi'(-\sin\phi'\,\hat{\mathbf{i}} + \cos\phi'\,\hat{\mathbf{j}}) \tag{20.4.2}$$

Since the field point *P* is on the axis of the loop at a distance *z* from the center, its position vector is given by $\vec{\mathbf{r}}_P = z\hat{\mathbf{k}}$. Thus we have

$$\mathbf{r}_P - \mathbf{r}' = -R\cos\phi'\,\hat{\mathbf{i}} - R\sin\phi'\,\hat{\mathbf{j}} + z\,\hat{\mathbf{k}} \tag{20.4.3}$$

and

$$|\,\mathbf{r}_P - \mathbf{r}'\,| = \sqrt{(-R\cos\phi')^2 + \left(-R\sin\phi'\right)^2 + z^2} = \sqrt{R^2 + z^2} \tag{20.4.4}$$

In (20.3.3), we need to compute the cross product $d\,\mathbf{l}' \times (\mathbf{r}_P - \mathbf{r}')$ which can be simplied as

$$\begin{aligned} d\,\mathbf{l}' \times (\mathbf{r}_P - \mathbf{r}') &= R\,d\phi'\left(-\sin\phi'\,\hat{\mathbf{i}} + \cos\phi'\,\hat{\mathbf{j}}\right) \times [-R\cos\phi'\,\hat{\mathbf{i}} - R\sin\phi'\,\hat{\mathbf{j}} + z\,\hat{\mathbf{k}}] \\ &= R\,d\phi'[z\cos\phi'\,\hat{\mathbf{i}} + z\sin\phi'\,\hat{\mathbf{j}} + R\,\hat{\mathbf{k}}] \end{aligned} \tag{20.4.5}$$

Using the Biot-Savart law, the contribution of the current element to the magnetic field at *P* is

$$\begin{aligned} d\vec{\mathbf{B}} &= \frac{\mu_0 I}{4\pi} \frac{d\,\mathbf{l} \times (\mathbf{r}_P - \mathbf{r}')}{|\,\mathbf{r}_P - \mathbf{r}'\,|^3} \\ &= \frac{\mu_0 IR}{4\pi} \left[\frac{z\cos\phi'\,\hat{\mathbf{i}} + z\sin\phi'\,\hat{\mathbf{j}} + R\,\hat{\mathbf{k}}}{(R^2 + z^2)^{3/2}} \right] d\phi' \end{aligned} \tag{20.4.6}$$

Carrying out the integration gives the magnetic field at *P* as

$$\mathbf{B} = \frac{\mu_0 IR}{4\pi} \int_0^{2\pi} \frac{\left[z\cos\phi'\,\hat{\mathbf{i}} + z\sin\phi'\,\hat{\mathbf{j}} + R\,\hat{\mathbf{k}}\right]}{(R^2 + z^2)^{3/2}} d\phi' \tag{20.4.7}$$

The *x* and the *y* components of $\vec{\mathbf{B}}$ can be readily shown to be zero, with the final result that

$$\mathbf{B} = \hat{\mathbf{z}} \frac{\mu_0 IR^2}{2(R^2 + z^2)^{3/2}} \tag{20.4.8}$$

Thus, we see that along the symmetric axis, B_z is the only non-vanishing component of the magnetic field. The conclusion can also be reached by using the symmetry arguments.

To calculate the field off axis is beyond our present mathematical capabilities, but we can see in principle what the field looks like. Figure 20-5 is a Shockwave visualization that shows how this is done.

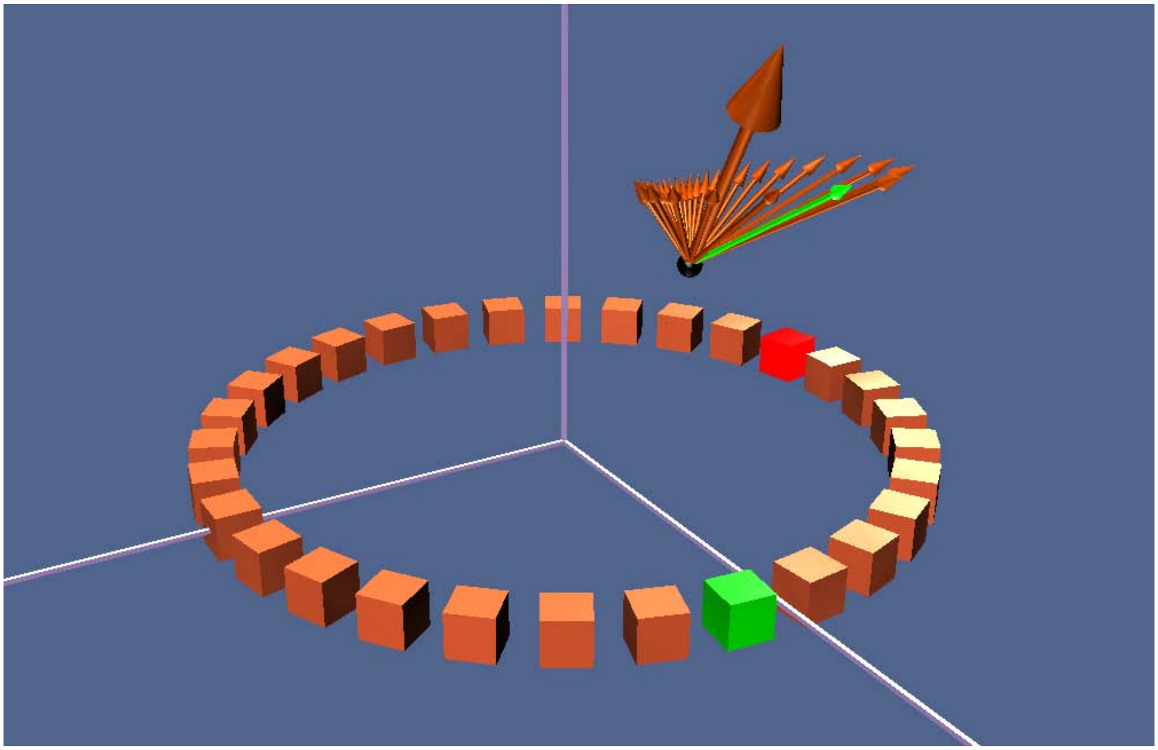


Figure 20-5: A shockwave visualization for constructing the field of a current ring

://web.mit.edu/viz/EM/visualizations/magnetostatics/calculatingMagneticFields/RingMagField/RingMagFieldFullScreen.htm

20.5 Ampere's Law

Ampere's Law is the integral form of (20.3.8). If we take any open surface *S* with contour *C*, then from (20.3.8) we have

$$\int_S \left(\nabla \times \mathbf{B}\right)\cdot\hat{\mathbf{n}}\,da = \mu_o \int_S \mathbf{J}\cdot\hat{\mathbf{n}}\,da \qquad (20.5.1)$$

We can use Stoke's theorem to convert the left hand side of (20.5.1) to a line integral, and the right hand side of (20.5.1) is simply the current through the surface S, with the positive direction of current defined right-handedly with respect to the direction of the contour integration. This gives us what is known as Ampere's Law.

$$\oint_C \mathbf{B}\cdot d\mathbf{l} = \mu_o I_{through} \qquad (20.5.2)$$

Ampere's Law can be used to do many straightforward problems which have a lot of symmetry. Examples are given on the problem set. As with Gauss's Law, Ampere's Law is always true, but it is usually useless for solving problems, unless there is a lot of symmetry. You can find a java applet on the web at the *url* below which illustrates what is going on with Ampere's Law in the general case without a lot of symmetry.

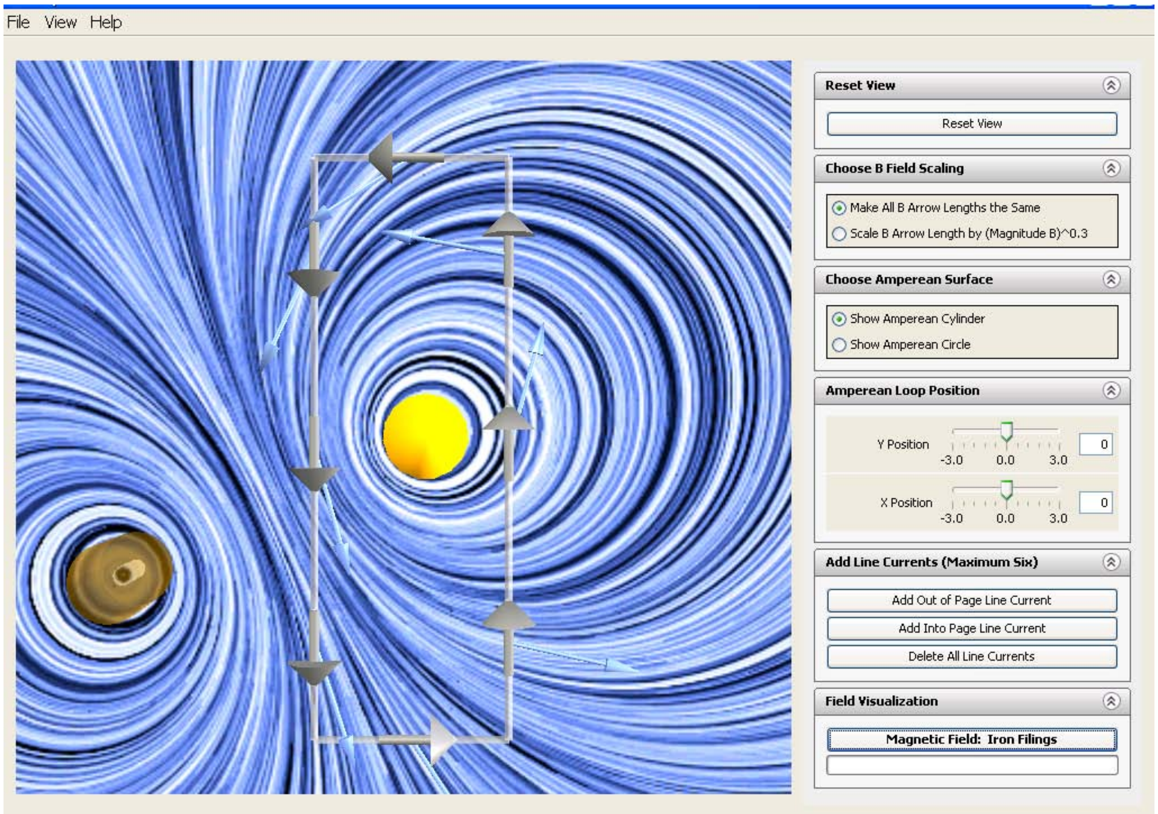


Figure 20-6: A Java applet illustrating Ampere's Law

://web.mit.edu/viz/EM/simulations/ampereslaw.jnlp

20.6 The Magnetic Potential

In general in magnetostatics the magnetic field cannot be derived from the gradient of a scalar function, since its curl is not zero ($\nabla\times\mathbf{B}(\mathbf{r})=\mu_o\mathbf{J}$). However if there are extensive regions where the current density **J** is zero, in those regions the curl of **B** is zero, and it can be derived from the gradient of the "magnetic" potential.

As an example of this, consider a spherical shell which carried a surface current $\kappa(\theta)$ in the azimuthal direction, that is

$$\mathbf{J}=\delta(r-R)\kappa(\theta)\hat{\phi} \tag{20.6.1}$$

In this situation we have that the curl of the magnetic field is zero everywhere in space except at $r = R$. Therefore the magnetic field can be written as the gradient of the following scalar potential

$$\phi_M(r,\theta) = \begin{cases} \sum_{l=0}^{\infty} A_l \left(\frac{r}{R}\right)^l P_l(\cos\theta) \\ \sum_{l=0}^{\infty} B_l \left(\frac{R}{r}\right)^{l+1} P_l(\cos\theta) \end{cases} \tag{20.6.2}$$

where we have imposed the conditions that the field not blow up at the origin and that it go to zero at infinity. We then determine the constant coefficients in (20.6.2) by imposing the boundary conditions that we must have on the field across $r = R$. These boundary conditions are discussed in the next section.

20.7 Boundary Conditions on the Magnetic Field

Before leaving magnetostatics, we discuss the boundary conditions that must be true across any thin interface in magnetostatics. We know that $\nabla \cdot \mathbf{B}(\mathbf{r}) = 0$, so we see that we must have no change in the normal component of the magnetic field across a thin interface, that is

$$B_{2n} = B_{1n} \tag{20.7.1}$$

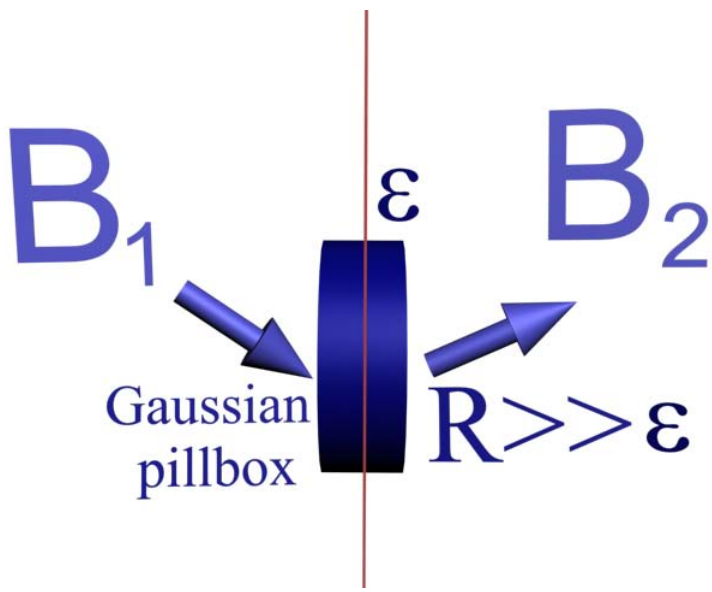


Figure 20-7: Boundary condition on the normal component of B across an interface

To make absolutely sure we understand where (20.7.1) comes from, consider the figure above. We Gauss's Law to the pillbox shown, making sure that the radius R of the pillbox is arbitrarily large compared to its height ε. The area of the sides of the pillbox is therefore given by $2\pi R\varepsilon$ and the area of the ends is given by πR^2. By making ε small enough compared to R, we can insure that any flux through the sides of the pillbox due to tangential components of **B** do not contribute to the integral over the surface area of the pillbox. Only the normal components of **B** contribute to the integral, and we easily obtain (20.7.1).

What about the tangential components of **B**? These can have a jump if there is a current sheet in the interface. To see this, take an amperean loop that spans the interface, with width ε perpendicular to the interface and length l tangential to the interface, with a loop normal that is parallel to the current sheet direction, as shown in the figure below. If we use Ampere's Law with the displacement term included, and integrate around this loop, we have

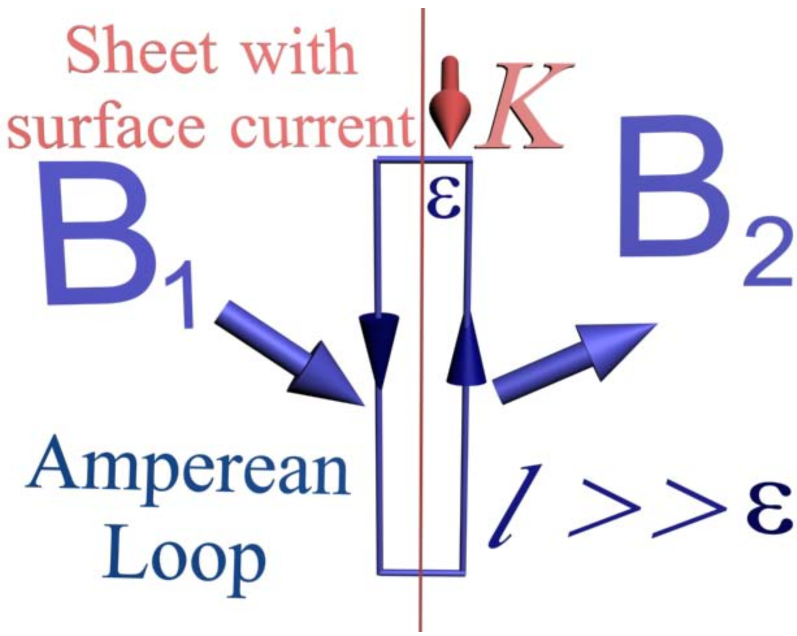


Figure 20-8: Boundary condition on the tangential component of B across an interface.

$$\oint \mathbf{B}\cdot d\,\mathbf{l} = \mu_o \int_{\text{open surface}} \mathbf{J}\cdot\hat{\mathbf{n}}\,da + \mu_o\varepsilon_o \int_{\text{open surface}} \frac{\partial \mathbf{E}}{\partial t}\,da \qquad (20.7.2)$$

If we make the width ε very small compared to the length l, the only component that will enter into the left hand side of (20.7.2) will be the tangential **B**, since any normal component will be multiplied by $\varepsilon << l$. And the magnitude of that component will be $B_t\, l$. If we look at the right hand side of (20.7.2), the first term will give $\mu_o\, l\, \kappa$, independent of ε, whereas the second term involves an area $l\varepsilon$. Again if we make $\varepsilon << l$ we can make the area integral of $\partial\mathbf{E}/\partial t$ insignificant. Therefore we will always have for the magnetic field, even in a time varying situation, that

$$\mathbf{B}_{2t} - \mathbf{B}_{1t} = \mu_o\, \boldsymbol{\kappa} \times \hat{\mathbf{n}} \qquad (20.7.3)$$

20.8 Biot-Savart as the Relativistic Transformation of the Rest Frame E Field

Before leaving magnetostatics, let us give a heuristic derivation of where the Biot-Savart Law comes from. If we think of our current source $I d\mathbf{l}$ as due to a moving point charge dq with velocity $\mathbf{v}$, then $I d\mathbf{l}$, which has units of charge times velocity, becomes $dq\mathbf{v}$, and Biot-Savart becomes

$$d\mathbf{B} = \frac{\mu_0}{4\pi}\frac{I\, d\mathbf{l}\times\mathbf{R}}{R^3} = \frac{\mu_0 dq}{4\pi}\frac{\mathbf{v}\times\mathbf{R}}{R^3} = \mu_0\varepsilon_o\mathbf{v}\times\left[\frac{dq}{4\pi\varepsilon_o}\frac{\mathbf{R}}{R^3}\right] = \frac{1}{c^2}\mathbf{v}\times\left[\frac{dq}{4\pi\varepsilon_o}\frac{\mathbf{R}}{R^3}\right] \quad (20.8.1)$$

We have written the last form in (20.8.1) so that we can discuss it in the context of (15.5.11), which we reproduce here.

$$\bar{B}_{\parallel} = B_{\parallel} \qquad \bar{\mathbf{B}}_{perp} = \gamma\left(\mathbf{B}_{perp} - \frac{1}{c^2}\mathbf{v}\times\mathbf{E}\right) \quad (20.8.2)$$

This tells us how to get the magnetic field in a frame moving at velocity $\mathbf{v}$ with respect to the laboratory frame. Let us suppose that in the laboratory frame we have a charge dq at rest. Then in that frame we have

$$d\mathbf{E} = \frac{dq}{4\pi\varepsilon_o}\frac{\mathrm{R}}{R^3} \qquad d\bar{\mathbf{B}} = 0 \quad (20.8.3)$$

We now use (20.8.2) to find the magnetic field in a frame moving at velocity $-\mathbf{v}$ with respect to the laboratory frame, where we assume that $\mathrm{v} << c$

$$\bar{\mathbf{B}} = \gamma\left(\mathbf{B}_{perp} - \frac{1}{c^2}\mathbf{v}\times\mathbf{E}\right) = \left(-\frac{1}{c^2}(-\mathbf{v})\times\frac{dq}{4\pi\varepsilon_o}\frac{\mathrm{R}}{R^3}\right) = \frac{1}{c^2}\mathbf{v}\times\left[\frac{dq}{4\pi\varepsilon_o}\frac{\mathrm{R}}{R^3}\right] \quad (20.8.4)$$

In this frame the charge dq appears to be moving at velocity $+\mathbf{v}$, and we see that the magnetic field is just that specified by Biot-Savart in the form of the last term in (20.8.1). Thus the magnetic field of a moving point charge is just the relativistic transformation of the Coulomb field in its rest frame to the frame in which the charge is moving.

20.9 Visualizations of the Magnetic Field of a Charge Moving at Constant Velocity

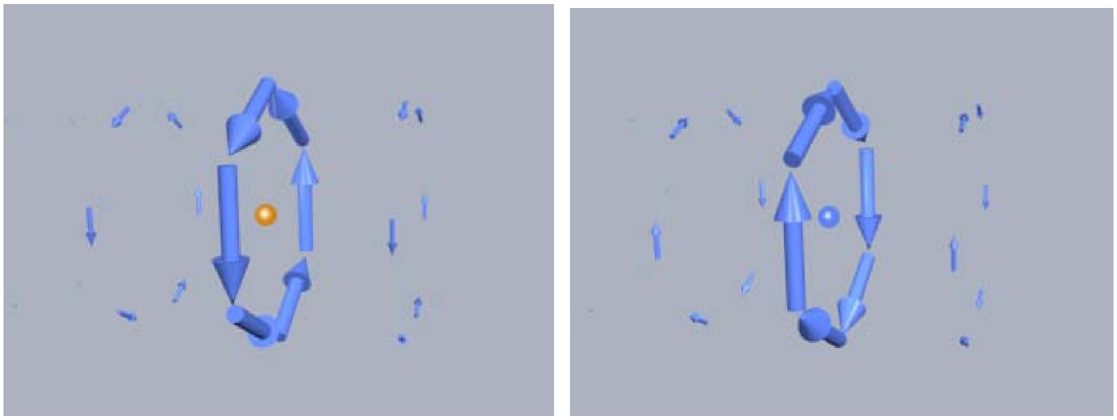


Figure 20-9: The Magnetic Field of a Positive and Negative Charge at Constant Speed

In Figure 20-9 we show the magnetic field of a positive and negative charge moving at constant speed. Movies of this can be found at

://web.mit.edu/viz/EM/visualizations/magnetostatics/MagneticFieldConfigurations/

21 Magnetic Force on a Moving Charge and on a Current Element

21.1 Learning Objectives

We look at the properties of the $q\mathbf{v}\times\mathbf{B}$ force and the corresponding form for a current carrying wire segment, $Id\mathbf{l}\times\mathbf{B}$.

21.2 $q\mathbf{E}$ and $q\mathbf{v}\times\mathbf{B}$ as the Result of Electric and Magnetic Pressure

The Lorentz force on a moving charge is given by

$$\mathbf{F} = q\left(\mathbf{E}+\mathbf{v}\times\mathbf{B}\right) \tag{21.2.1}$$

There are a number of things to be said about this equation, the first being that non-relativistically, the combination $\left(\mathbf{E}+\mathbf{v}\times\mathbf{B}\right)$ is the electric field as seen in the rest frame of the charge. We can see this by looking back at equation (15.5.10) for how $\mathbf{E}$ transforms, which we reproduce here.

$$\overline{E}_{\parallel} = E_{\parallel} \qquad \overline{\mathbf{E}}_{\perp} = \gamma\left(\mathbf{E}_{\perp}+\mathbf{v}\times\mathbf{B}\right) \tag{21.2.2}$$

There will be a number of occasions where we will identify $(\mathbf{E}+\mathbf{v}\times\mathbf{B})$ as the electric field in the rest frame of the charge q or of the moving current carrying element $d\mathbf{l}$.

Also, if we consider the Maxwell Stress Tensor, we see that the $q\mathbf{v}\times\mathbf{B}$ magnetic force can be understood as due to pressure and tension in the total magnetic field, just as $q\mathbf{E}$ can be understood as due to pressure and tension from the total electric field. In Figure 21-1 we show frames of movies that illustrate this for the electric field, and in Figure 21-2 we show frames of movies that illustrate this for the magnetic field, and we also give links to the respective movies. If we look at these movies with an eye towards the pressures and tensions given by the Maxwell Stress Tensor, we can get a qualitative feel for why these forces, both electric and magnetic, are in the directions that they are.

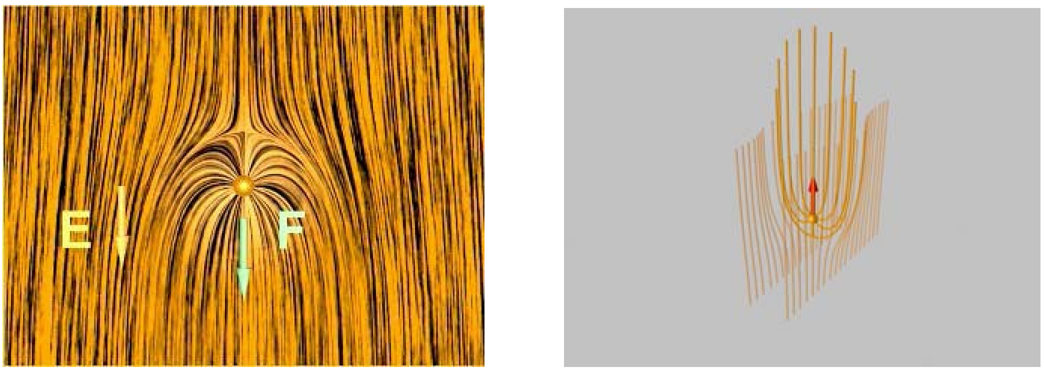


Figure 21-1: The qE Force as Due to Maxwell Stresses and Tensions

The figures are frames from movies which can be found at
://web.mit.edu/viz/EM/visualizations/electrostatics/ForcesOnCharges/forceq/forceq.htm
and
://web.mit.edu/viz/EM/visualizations/electrostatics/ForcesOnCharges/force_in_efield/force_in_efield.htm

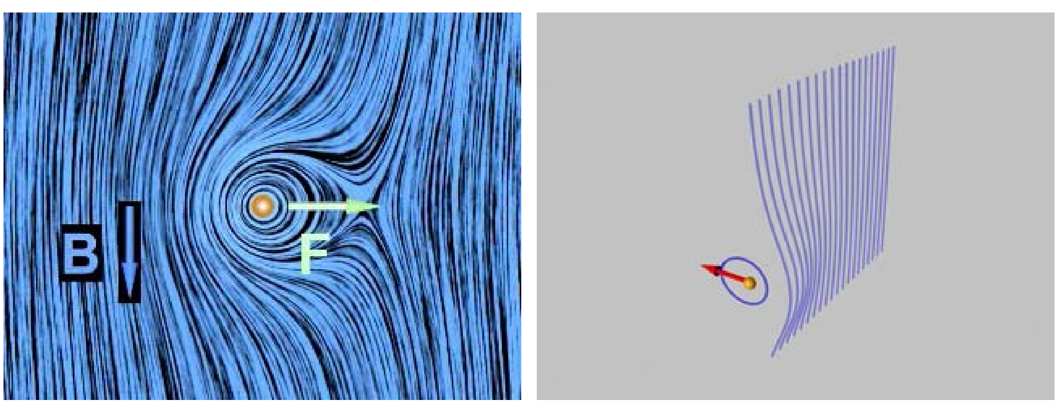


Figure 21-2: The $q\mathbf{v}\times\mathbf{B}$ Force as Due to Maxwell Stresses and Tensions

The figures are frames from movies which can be found at

://web.mit.edu/viz/EM/visualizations/magnetostatics/ForceOnCurrents/forcemovingq/forcemovingq.htm
and
://web.mit.edu/viz/EM/visualizations/magnetostatics/ForceOnCurrents/force_in_bfield/force_in_bfield.htm

21.3 Where Does the Momentum Go?

When a charge is moving in cyclotron motion in a magnetic field, its momentum is continually changing. How is momentum conserved in this situation? To answer this, we show in Figure 21-3 a charge entering an external magnetic field which is non-zero only over the circular segment shown in the Figure. The charge's velocity is such that it makes exactly a quarter of a cyclotron revolution before exiting the region where the external field is non-zero. We show only those magnetic field lines which are just outside the circle of revolution of the charge. The field shown is the sum of the external field and the field of the moving charge.

It is clear that the momentum change of the particle over this sequence is absorbed by the current segments which are producing the external field, contained within the solid circular segments at the top and bottom of the figures. If you look at the stress tensor near the currents producing the external field, you can clearly see a force which is pushing those current elements in a direction such as to conserve total momentum. The field is the conduit of the momentum exchange between the charge and the currents producing the external field, but it stores no momentum itself.

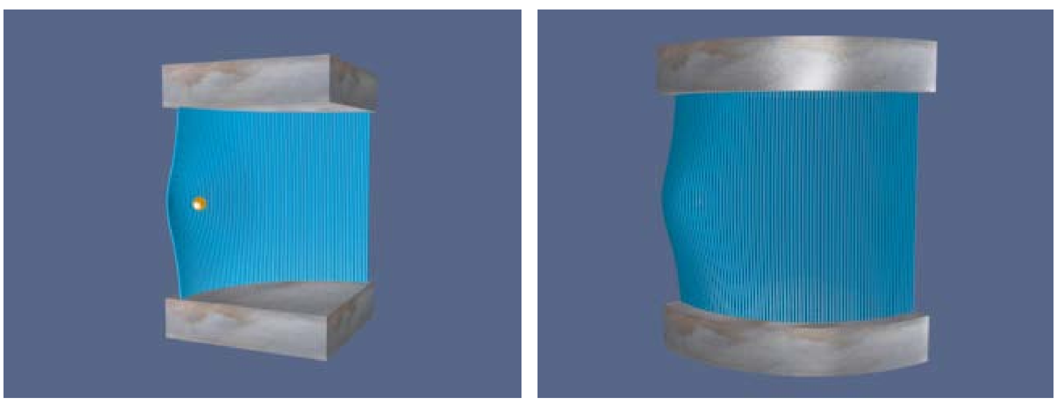


Figure 21-3: A Charge in a External Magnetic Field Transfering Momentum to the Sources of the External Field

The figures are frames from movies which can be found at
://web.mit.edu/viz/EM/visualizations/magnetostatics/ForceOnCurrents/MovingQinMagnet/MovingQinMagnetFront.htm
and
://web.mit.edu/viz/EM/visualizations/magnetostatics/ForceOnCurrents/MovingQinMagnet/MovingQinMagnetBack.htm

21.4 The $Id\mathbf{l}\times\mathbf{B}$ Force

Suppose we have n charges per unit volume moving at speed $\mathbf{v}$ in an external $\mathbf{B}$ field. Then the force per unit volume is the force on one times the number of charges per unit volume, or

$$\mathbf{F}_{\text{unit volume}} = qn\,\mathbf{v}\times\mathbf{B} = \mathbf{J}\times\mathbf{B} \tag{21.4.1}$$

where we have used the fact that $\mathbf{J}=qn\,\mathbf{v}$ to get the last form in (21.4.1). To get the total force on a given volume, we integrate over volume.

$$\mathbf{F} = \int \mathbf{J}\times\mathbf{B}\,d^3x \rightarrow \oint Id\mathbf{l}\times\mathbf{B} \tag{21.4.2}$$

where the last form in (21.4.2) is the total force on a current carrying wire in an external field. Thus we see that $Id\mathbf{l}\times\mathbf{B}$ is the force on a current carrying wire segment.

21.5 Force between Parallel Current Carrying Wires

It is easy to see using $Id\mathbf{l}\times\mathbf{B}$ that if we have two parallel current carrying wires, the magnetic force is attractive if the currents in the wires are in the same direction, and repulsive if the currents are in opposite directions. We can also see this from looking at the Maxwell Stress Tensor. If the current in the wires is in the same direction, the field from the two wires subtracts between the wires, and if the current in the wires is in opposite directions, the field adds between the wires. Thus the wires are either pulled together or forced apart by the pressures and tensions associated with the field. This is shown in the two movies in Figure 21-4.

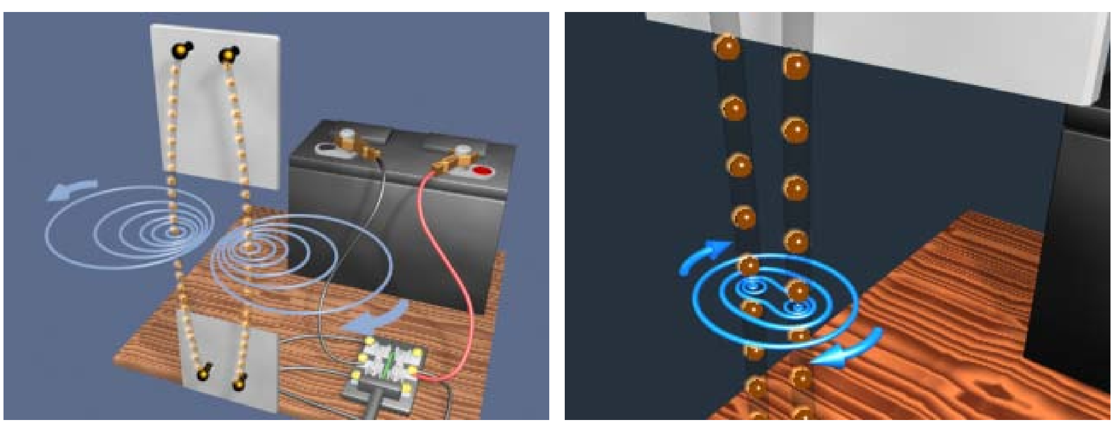


Figure 21-4: Magnetic fields between Parallel Current Carrying Wires

The figures are frames from movies which can be found at

://web.mit.edu/viz/EM/visualizations/magnetostatics/ForceOnCurrents/SeriesWires/SeriesWires.htm
and
://web.mit.edu/viz/EM/visualizations/magnetostatics/ForceOnCurrents/ParallelWires/ParallelWires.htm

21.6 Forces on Electric and Magnetic Dipoles in External Fields

21.6.1 The General Derivation

We consider the total force on an isolated and finite distribution of charges $\rho(\mathbf{r})$ and currents $\mathbf{J}(\mathbf{r})$ sitting in the electric field $\mathbf{E}^{ext}(\mathbf{r})$ and magnetic field $\mathbf{B}^{ext}(\mathbf{r})$ due to some external distribution of charges and currents. If the extent of our isolated distrition of charges is d, we will assume that any characteristic scale of variation L in the external fields is such that $d << L$. Then the force on our isolated distribution of charges and currents is given by

$$\mathbf{F} = \int_{volume\ V} \left[\rho\mathbf{E} + \mathbf{J}\times\mathbf{B}\right] d^3x = \int_{volume\ V} \left[\rho\mathbf{E}^{ext} + \mathbf{J}\times\mathbf{B}^{ext}\right] d^3x \tag{21.6.1}$$

We suppose that our isolated distribution of charges and currents are centered at the origin, and we expand our external fields in a Taylor series about the origin.

$$E_i^{ext}(\mathbf{r}) = E_i^{ext}(0) + x_j \frac{\partial}{\partial x_j} E_i^{ext}(0) + ... = E_i^{ext}(0) + \mathbf{r}\cdot\nabla E_i^{ext}(0) + ... \tag{21.6.2}$$

where $\frac{\partial}{\partial x_j} E_i^{ext}(0)$ is the gradient of the i^{th} component of **E** with respect to x_j, evaluated at the origin, and similarly for the magnetic field. With this expansion, (21.6.1) becomes to first zeroth order in d/L

$$\mathbf{F} = \mathbf{E}^{ext}(0)\int_V \rho d^3x + \left(\int_V \mathbf{J} d^3x\right)\times\mathbf{B}^{ext}(0) = Q\mathbf{E}^{ext}(0) \tag{21.6.3}$$

since in magnetostatics we have $\int_V \mathbf{J} d^3x = 0$, and where we are using

$$Q = \int_V \rho(\mathbf{r})\, d^3x \tag{21.6.4}$$

If we consider the first order terms in d/L in (21.6.1), using the Taylor series expansion (21.6.2), we have

$$\mathbf{F} = \left(\int_V \rho\, \mathbf{r}\, d^3x \right) \cdot \nabla \mathbf{E}^{ext}(0) + \left(\int_V \mathbf{J} \times \left[\mathbf{r} \cdot \nabla \mathbf{B}^{ext}(0) \right] d^3x \right) \tag{21.6.5}$$

and after some work (21.6.5) can be written as

$$\mathbf{F} = \mathbf{p} \cdot \nabla \mathbf{E}^{ext}(0) + \mathbf{m} \cdot \nabla \mathbf{B}^{ext}(0) \tag{21.6.6}$$

where

$$\mathbf{p} = \int_V \mathbf{r} \rho(\mathbf{r})\, d^3x \tag{21.6.7}$$

$$\mathbf{m} = \frac{1}{2} \int_V \mathbf{r} \times \mathbf{J}(\mathbf{r})\, d^3x' = \frac{I}{2} \oint \mathbf{r} \times d\mathbf{l} \tag{21.6.8}$$

Thus there is no force on an electric and magnetic dipole sitting in external fields unless the external fields have gradients, and if they do have gradients then the forces on the dipoles are given by (21.6.6).

21.6.2 A Specific Example of a $\mathbf{m} \cdot \nabla \mathbf{B}^{ext}(0)$ Force

It is easy to see why an electric dipole will field a net force in an inhomogeous electric field, but is not so obvious why a magnetic dipole will field a force in an inhomogeneous magnetic field. We give here an example of how this happens, using the case of two circular loops of current having the same axis, as shown in Figure 21-5.

Suppose we have two current loops sharing the same axis. Suppose furthermore that the sense of the current is the same in both loops as in Figure 21-5. Then these loops attract one another. This is a specific example of the forces on magnetic dipoles in non-uniform fields. To see why this is so, consider the magnetic field $\mathbf{B}_1$ of the lower loop as seen at the location of the upper loop. The presence of this **B** field will cause a force on the upper loop because of the $d\mathbf{F} = i\, d\mathbf{l} \times \mathbf{B}_1$ force. If we draw $d\mathbf{l}$ and $\mathbf{B}_1$ carefully, and take their cross-product, we get the result shown in the figure, that is a $d\mathbf{F}$ that has a radial component and a downward component. The radial component will cancel out when we integrate all away around the upper loop to find the net force, but the downward component of the force will not. Thus the upper loop feels a force *attracting* it to the lower loop.

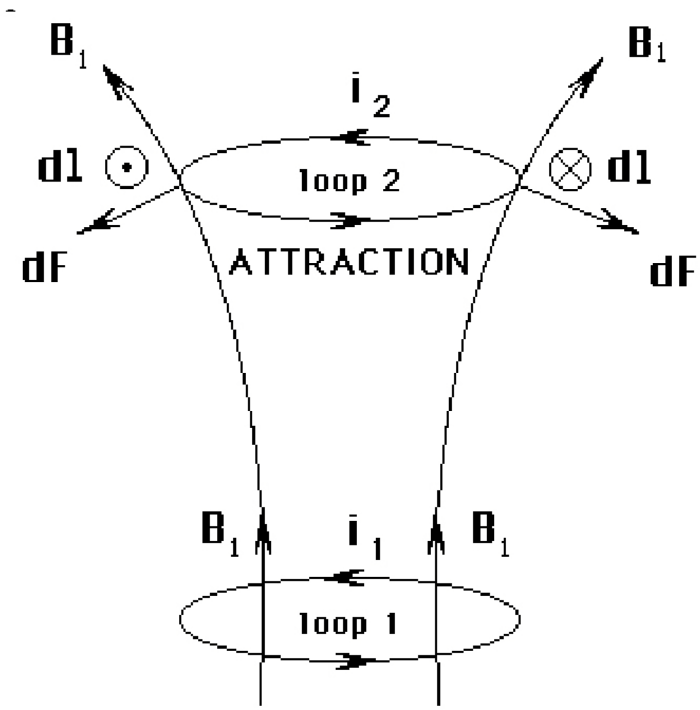


Figure 21-5: Co-axial current loops with currents in the same direction

If now we reverse the direction of current in the upper loop, so that now the currents in the two loops are in opposite senses, we find that the loops are repelled by each other, because the $d\mathbf{F}$'s have reversed from the situation above (see Figure 21-6).

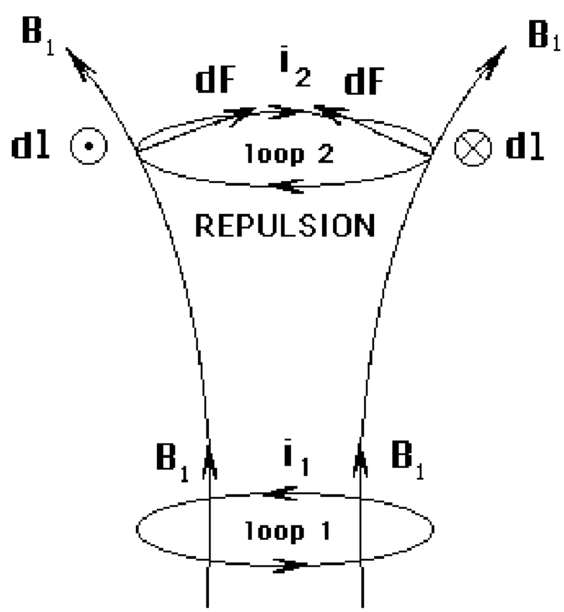


Figure 21-6: Co-axial current loops with currents in opposite directions

Note that this attraction or repulsion depending on the relative sense of the currents explains why north poles of permanent magnets attract the south poles of other magnetics and repel their north poles. We know that permanent magnets are caused by circulating atomic currents in the materials making up the magnet. The north pole of a permanent magnet is right-handed with respect to the its atomic currents--that is, if you curl the fingers of your right hand in the direction of atomic current flow, your thumb will point in the direction of the north pole of the permanent magnet.

In Figure 21-7, I show two permanent magnets, both with their north poles up, and the sense of their atomic currents for this orientation (the atomic currents of course flow on the surface or*inside* the magnets). The magnets in this orientation attract one another, for the same reason that the two co-axial current loops in Figure 21-5 attract one another--the currents are in the same sense, and the resulting $i\, d\mathbf{l} \times \mathbf{B}$ force on the atomic currents in one magnet due to the presence of the magnetic field of the other magnet results in attraction. We loosely say that the south pole of the top magnet "attracts" the adjacent north pole below.

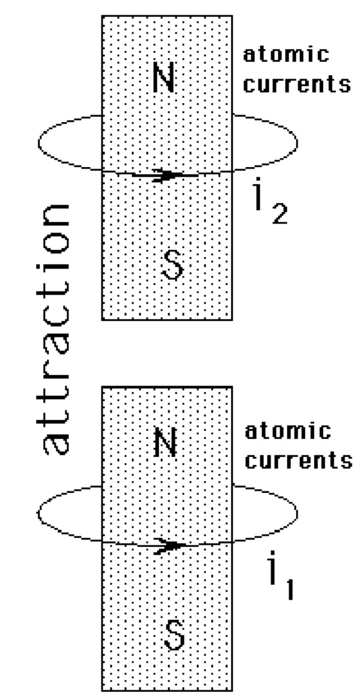


Figure 21-7: Two bar magnets attracting

If we turn the top magnet upside down, so that now the north poles of the two magnets are adjacent, we have reversed the sense of the currents, and therefore we get repulsion, for the same reasons the two current loops repel one another in Figure 21-6. We loosely say that the north pole of the top magnet "repels" the adjacent north pole below. All of these phenomena are simply the result of $q\mathbf{v} \times \mathbf{B}$ forces, which are the same as $i\, d\mathbf{l} \times \mathbf{B}$ forces.

21.7 Torques on Electric and Magnetic Dipoles in External Fields

21.7.1 The General Derivation

Now we consider the torque on an isolated and finite distribution of charges $\rho(\mathbf{r})$ and currents $\mathbf{J}(\mathbf{r})$ sitting in external fields. The torque is given by

$$\boldsymbol{\tau} = \int_V \mathbf{r}\times\left[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}\right]d^3x = \int_V \mathbf{r}\times\left[\rho\mathbf{E}^{ext}+\mathbf{J}\times\mathbf{B}^{ext}\right]d^3x \tag{21.7.1}$$

And after some manipulation this can be written as

$$\boldsymbol{\tau} = \mathbf{p}\times\mathbf{E}_{ext}(0)+\mathbf{m}\times\mathbf{B}_{ext}(0) \tag{21.7.2}$$

21.7.2 A Specific Example of a $\mathbf{m}\times\mathbf{B}_{ext}(0)$ Torque

Although it is easy to see where the torque on an electric dipole in an external electric field comes from, it is not so obvious where the torque on a magnetic dipole in an external magnetic field comes from, so we consider this topic further. Suppose we have a rectangular loop of wire with sides of length *a* and *b*, carrying current *i* (see Figure 21-8). It is free to rotate about the axis indicated in the figure. The normal to the plane of the rectangular loop of wire is **n**, where we take **n** to be right-handed with respect to the direction of the current flow (if you curl the fingers of your right hand in the direction of current flow, then your thumb is in the direction of **n**). The normal **n** makes an angle θ with a uniform external magnetic field $\mathbf{B}_{\text{ext}}$, as indicated. If we define the magnetic dipole moment **m** of the loop to be $iA\mathbf{n}$, where $A = ab$ is the area of the loop, then according to (21.7.2), the loop will feel a net torque $\boldsymbol{\tau}$ due to magnetic forces with $\boldsymbol{\tau} = \mathbf{m}\times\mathbf{B}_{ext}$. This torque will tend to align **m** with $\mathbf{B}_{\text{ext}}$.

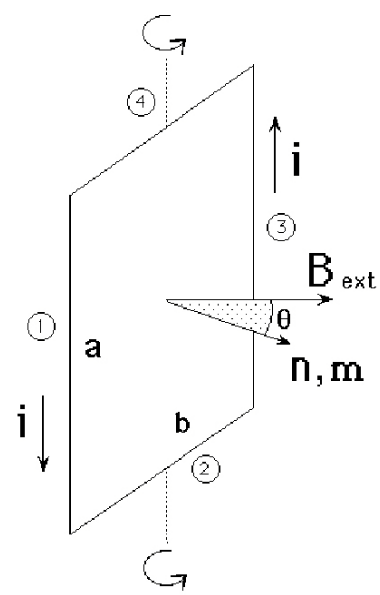


Figure 21-8: A rectangular loop of current in an external magnetic field

To demonstrate that $\boldsymbol{\tau} = \mathbf{m} \times \mathbf{B}_{ext}$ is the correct expression, consider the forces on each of the four sides of the loop, using the expression for the force on a segment of wire given by $d\mathbf{F} = Id\mathbf{l} \times \mathbf{B}$. The directions of these four forces are indicated in Figure 21-9 below.

In each case, these forces are perpendicular both to $\mathbf{B}_{ext}$ and to the direction of the current in the segment. The magnitude of the force $\mathbf{F}_2$ on side 2 (of length b) is $i\, b\, B_{ext} \cos\theta$, and is equal to the magnitude of the force $\mathbf{F}_4$ on side 4, although opposite in direction. These two forces tend to expand the loop, but taken together contribute nothing to the net force on the loop, and moreover they have zero moment arm through the center of the loop, and therefore contribute no net torque.

The forces $\mathbf{F}_1$ and $\mathbf{F}_3$ also have a common magnitude, $i\, a\, B_{ext}$, and are oppositely directed, so again they contribute nothing to the net force on the loop. However, they *do* contribute to a net torque on the loop, and it is obvious from studying the figure that the two forces tend to rotate the loop in a direction that tends to bring **m** into alignment with $\mathbf{B}_{ext}$. The forces both have a moment arm of $(b/2)\sin\theta$, and so the total torque is

$$\left|\boldsymbol{\tau}\right| = 2\left(iaB_{ext}\right)\frac{b}{2}\sin\theta = iab\sin\theta = \left|\mathbf{m} \times \mathbf{B}\right| \qquad (21.7.3)$$

which is the result we desired. In vector form, $\boldsymbol{\tau} = \mathbf{m} \times \mathbf{B}_{ext}$.

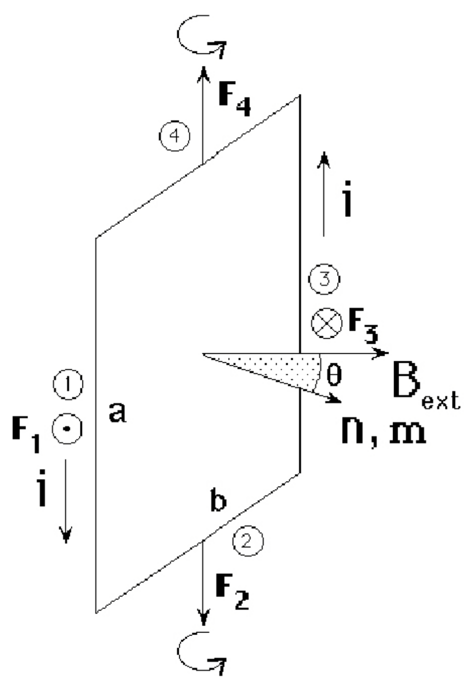


Figure 21-9: The forces on a rectangular loop in an external magnetic field

21.8 Free Dipoles Always Attract

Finally, we point out that when dipoles free to rotate and translate are allowed to interact, the combination of the torques and forces are such that the dipoles always attract. For example the torque will cause to magnetic dipoles to align so that their currents are in the same sense, and then the force of attraction in that configuration will pull them together. An example of this behavior is shown in Figure 21-10.

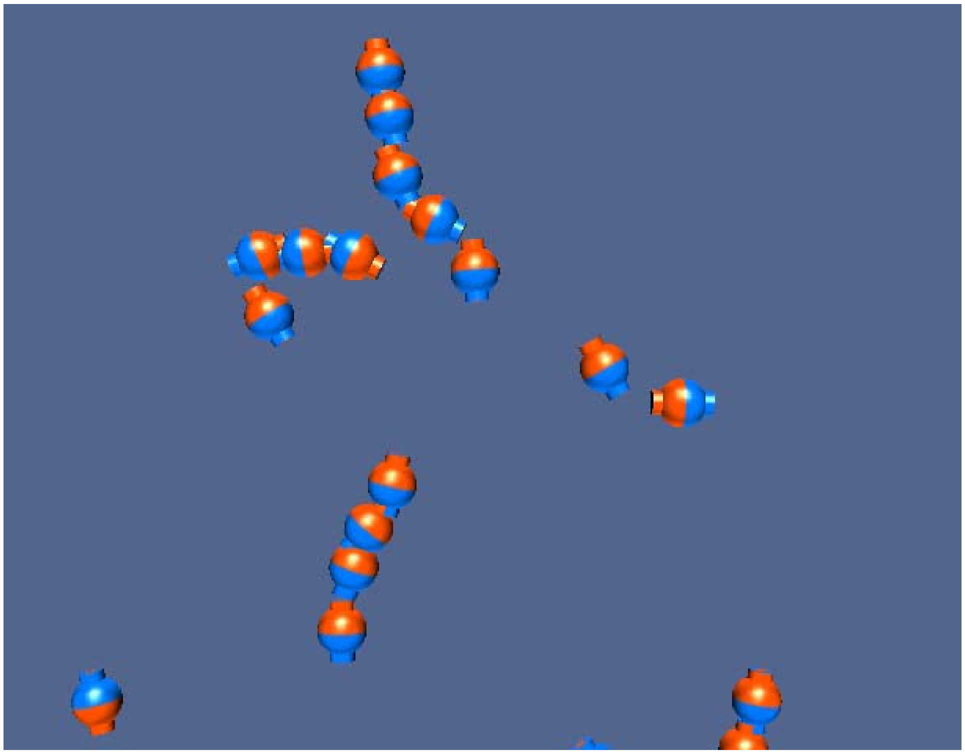


Figure 21-10: Interacting Dipoles Attracting

The application from which this frame is taken can be found at
://web.mit.edu/viz/EM/visualizations/electrostatics/ForcesTorquesOnDipoles/DipolesShock/DipolesShock.htm

22 Creating and Destroying Electromagnetic Energy and Angular Momentum

22.1 Learning Objectives

We begin our considerations of Faraday's Law, and consider it first in the context of the creation and destruction of magnetic fields, energy, and angular momentum. In later sections we will consider other uses, e.g. with respect to inductance in circuits, but first we look at the most fundamental implication of Faraday's Law.

22.2 Faraday's Law

Faraday's Law in differential form is

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{22.2.1}$$

This law is the basis for how we understand why it takes energy to create magnetic fields. The process by which we put energy into magnetic fields, and retrieve it, is fundamental enough that it bears emphasis, and it is a good illustration of the conservation laws we have developed. We have not discussed the energy in magnetic fields up until now because calculating the energy to create magnetic fields is an intrinsically time dependent process.

You may argue that calculating the energy to assemble electric charges is also an intrinsically time dependent process, and we used statics there and got an answer. The difference is that we must have an electric field to work against to evaluate the work done--magnetic fields do no work--and in electrostatics we know the electric fields we are working against to a first approximation. In magnetostatics there are no electric fields in the absolutely static limit, and up to now we have not known how to calculate the electric field produced by slowly varying magnetic fields, which is what we must work against to produce them.

So we are finally in a position to consider the forces that arise, and the work we must do against those forces, when we try to create a magnetic field. The best way to convince us intuitively that magnetic fields require energy to create is to set up a situation where *we* must do the work to create them (rather than let a battery do the work, for example). This is the basis of the example in Section

22.2.1 E x B drift of monopoles in crossed E and B fields

Before we get to the quantitative details, let us first look at a visualization of this process. Before we get to the visualization we need to discuss how we animate the magnetic field lines in this movie. Here is how it is done.

For an electric charge with velocity **v**, mass *m*, and electric charge *q*, the non-relativistic equation of motion in constant **E** and **B** fields is

$$\frac{d}{dt}m\mathbf{v} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{22.2.2}$$

If we define the $\mathbf{E} \times \mathbf{B}$ drift velocity for electric monopoles to be

$$\mathbf{V}_{d,E} = \frac{\mathbf{E} \times \mathbf{B}}{B^2} \tag{22.2.3}$$

and make the substitution

$$\mathbf{v} = \mathbf{v}' + \mathbf{V}_{d,E} \tag{22.2.4}$$

then (22.2.2) becomes (assuming **E** and **B** are perpendicular and constant)

$$\frac{d}{dt}m\mathbf{v}' = q\,\mathbf{v}' \times \mathbf{B} \tag{22.2.5}$$

The motion of the electric charge thus reduces to a gyration about the magnetic field line superimposed on the steady drift velocity given by (22.2.3). This expression for the drift velocity is only physically meaningful if the right-hand side is less than the speed of light. This assumption is equivalent to the requirement that the energy density in the electric field be less than that in the magnetic field.

For a hypothetical magnetic monopole of velocity **v**, mass *m*, and magnetic charge q_m, the non-relativistic equation of motion[i] is

(22.2.7) $$\frac{d}{dt}m\mathbf{v} = q_m(\mathbf{B} - \mathbf{v} \times \mathbf{E}/c^2) \tag{22.2.6}$$

If we define the $\mathbf{E} \times \mathbf{B}$ drift velocity for magnetic monopoles to be

$$\mathbf{V}_{d,B} = c^2\frac{\mathbf{E} \times \mathbf{B}}{E^2} \tag{22.2.7}$$

and make the substitution analogous to (22.2.3), then we recover (22.2.5) with **B** replaced by $-\mathbf{E}/c^2$. That is, the motion of the hypothetical magnetic monopole reduces to a gyration about the electric field line superimposed on a steady drift velocity given by (22.2.7). This expression for the drift velocity is only physically meaningful if it is less than the speed of light. This assumption is equivalent to the requirement that the energy density in the magnetic field be less than that in the electric field. Note that these drift velocities are independent of both the charge and the mass of the monopoles.

In situations where **E** and **B** are not independent of space and time, the drift velocities given above are still approximate solutions to the full motion of the monopoles as long as the radius and period of gyration are small compared to the characteristic length and time scales of the variation in **E** and **B**. There are other drift velocities that depend on both the sign of the charge and the magnitude of its gyroradius, but these can be made arbitrarily small if the gyroradius of the monopole is made arbitrarily small. The gyroradius depends on the kinetic energy of the charge as seen in a frame moving with the drift velocities. When we say that we are considering "low-energy" test monopoles in what follows, we mean that we take the kinetic energy (and thus the gyroradii) of the monopoles in a frame moving with the drift velocity to be as small as we desire.

The definition use to construct our electric field line motions is equivalent to taking the local velocity of an electric field line in electro-quasi-statics to be the drift velocity of low energy test magnetic monopoles spread along that field line. Similarly, the definition we use to construct our magnetic field line motions below is equivalent to taking the local velocity of a magnetic field line in magneto-quasi-statics to be the drift velocity of low energy test electric charges spread along that field line. These choices are thus physically based in terms of test particle motion, and have the advantage that the local motion of the field lines is in the direction of the Poynting vector.

22.2.2 A visualization of the creation and destruction process

Before we get to the quantitative details, let us consider a movie of this process. In Figure 22-1, we show one frame of a movie showing the creation process. You can find a link to this movie at

://web.mit.edu/viz/EM/visualizations/faraday/CreatingMagneticEnergy/

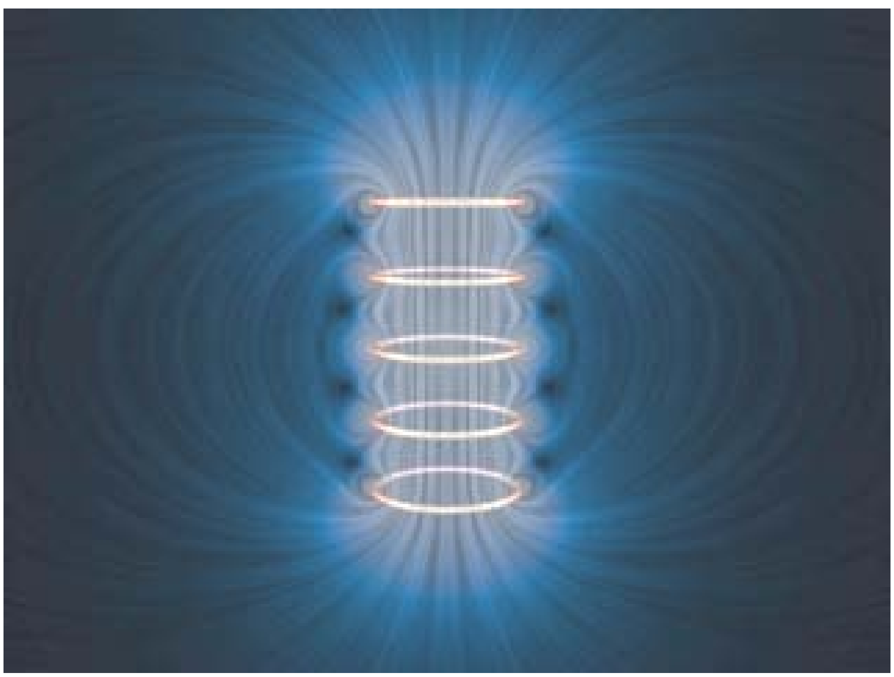


Figure 22-1: One frame of a movie showing the creation of magnetic fields

In the movie, we have five rings that carry a number of free positive charges that are not moving. Since there is no current, there is no magnetic field. Now suppose a set of external agents come along (one for each charge) and simultaneously spin up the charges counterclockwise as seen from above, at the same time and at the same rate, in a manner that has been pre-arranged. Once the charges on the rings start to accelerate, there is a magnetic field in the space between the rings, mostly parallel to their common axis, which is stronger inside the rings than outside. This is the solenoid configuration we shall consider in quantitative detail below in the next section.

As the magnetic flux through the rings grows, Faraday's Law tells us that there is an electric field induced by the time-changing magnetic field. This electric field is circulating clockwise as seen from above. The force on the charges due to this electric field is thus opposite the direction the external agents are trying to spin the rings up in (counterclockwise), and thus the agents have to do additional work to spin up the rings because they are charged. This is the source of the energy that is appearing in the magnetic field between the rings-the work done by the agents against the "back emf".

Over the time when the magnetic field is increasing in the animation, the agents moving the charges to a higher speed against the induced electric field are continually doing work. The electromagnetic energy that they are creating at the place where they are doing work (the path along which the charges move) flows both inward and outward. The direction of the flow of this energy is shown by the animated texture patterns. This is the electromagnetic energy flow that increases the strength of the magnetic field in the space between the rings as each positive charge is accelerated to a higher and higher velocity.

In the case of the destruction of magnetic fields, suppose we have the same five rings as above, this time carrying a number of free positive charges that are moving counter-clockwise. This current results in a magnetic field that is strong inside the rings and weak outside. Now suppose a set of external agents come along (one for each charge) and simultaneously spin down the charges as seen from above, at the same time and at the same rate, in a manner that has been pre-arranged. Once the charges on the rings start to deccelerate, the magnetic field begins to decrease in intensity.

As the magnetic flux through the rings decreases, Faraday's Law tells us that there is an electric field induced by the time-changing magnetic field. This electric field is circulating clockwise as seen from above. The force on the charges due to this electric field is thus opposite the direction the external agents are trying to spin the rings down in (counter-clockwise), and thus work is done on those agents.

As the strength of the magnetic field decreases, the magnetic energy flows from the field back to the path along which the charges move, and is now provided to the agents trying to spin down the moving charges. The energy provided to those agents as they destroy the magnetic field is exactly the amount of energy that they put into creating the magnetic field in the first place, neglecting radiative losses. This is a totally reversible process if we neglect such losses. That is, the amount of energy the agents put into creating the magnetic field is exactly returned to the agents as the field is destroyed.

22.3 The spinning cylinder of charge

22.3.1 The fields of a spinning cylinder

Now let us do the quantitative calculation that confirms the qualitative picture above. Suppose we create a magnetic field in the following way. We have a long cylindrical shell of non-conducting material which carries a surface charge fixed in place (glued down) of σ Coulombs per square meter. The length of the cylinder is *L*, which is much greater than its radius *R*. The cylinder is suspended in a manner such that it is free to revolve about its axis, without friction. Initially it is at rest, and there is no magnetic field. We come along and spin the cylinder up until the speed of the surface of the cylinder is V_O. After spinning it up, we will have a magnetic field inside the cylinder and outside the cylinder the field will be zero. The field inside the cylinder will be given by

$$\mathbf{B}_o = \mu_o \kappa_o \,\hat{\mathbf{z}} = \mu_o \sigma V_o \,\hat{\mathbf{z}} \qquad (22.3.1)$$

This is just our standard magnetostatic formula for the field inside a solenoid with surface current κ_o. Let's calculate the amount of work we have to do to create this magnetic field.

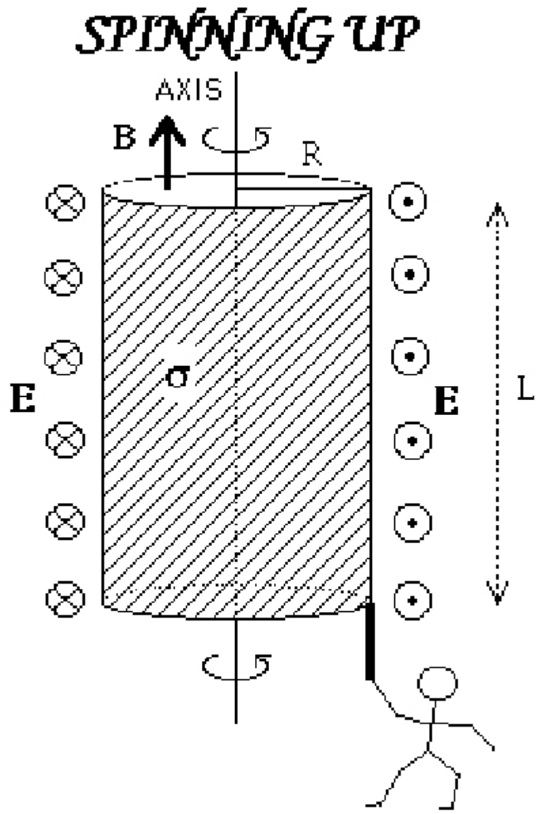


Figure 22-2: Spinning up a cylindrical shell of charge

Imagine we are in the middle of this process, and have gotten the cylinder up to some speed *V(t)* at time *t*, where *V(t)* is increasing with time and is less than the final speed V_o. At that point, there will be a magnetic field inside the cylinder that is also increasing with time, which is ***approximately*** given by

$$\mathbf{B}(\mathbf{r},t) \approx \mu_o \kappa(t)\,\hat{\mathbf{z}} = \mu_o \sigma V(t)\,\hat{\mathbf{z}} \tag{22.3.2}$$

You should immediately object that this is now a time varying situation, so the static solution is no longer correct. However we have seen in our studies of the various regions around an isolated time varying set of charges and currents (see Section 9.3) that it is ok to use the static solutions as long as the time T over which we spin the cylinder up significantly is much longer than the time that it takes light to cross the cylinder, R/c. If this is true, and we assume it is, and if we are within the Near Zone, that is, $r << cT$ then equation (22.3.2) is a good approximation to the actual magnetic field.

Now, $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$, with **B** in the z-direction, and let us assume that the electric field is a combination of an azimuthal electric field in addition to our cylindrical radial field associated with the charge density σ. That is, we assume that

$$\mathbf{E}(\mathbf{r},t) = \begin{cases} E_\phi \hat{\boldsymbol{\phi}} & r < R \\ \dfrac{\sigma}{\varepsilon_o}\hat{\mathbf{r}} + E_\phi \hat{\boldsymbol{\phi}} & r > R \end{cases} \tag{22.3.3}$$

If we apply Faraday's Law to a circle of cylindrical radius r about the axis of the cylinder, we find that the line integral of the electric field yields

$$\oint \mathbf{E}(\mathbf{r},t)\cdot d\mathbf{l} = 2\pi r E_\phi \tag{22.3.4}$$

The magnetic flux through the imaginary circle at time t is given by

$$\int_{surface} \mathbf{B}(\mathbf{r},t)\cdot\hat{\mathbf{n}}\,da = \begin{cases} \pi r^2 \sigma V(t) & r < R \\ \pi R^2 \sigma V(t) & r > R \end{cases} \tag{22.3.5}$$

Thus we find that that part of the electric field "induced" by the time-changing magnetic field is

$$E_\phi \hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\phi}} \begin{cases} \dfrac{\mu_o r\,\sigma}{2}\dfrac{dV(t)}{dt} & r < R \\ \dfrac{\mu_o R^2 \sigma}{2r}\dfrac{dV(t)}{dt} & r > R \end{cases} \tag{22.3.6}$$

We frequently refer to that part of the electric field associated purely with time-changing magnetic fields as *induced* fields, thus the origin of the terms "inductance", "inductors", and so on.

The sense of the electric field is shown in Figure 22-2. When we are trying to spin up the cylinder, the "induced" electric field will cause forces on the (glued on) charges on the cylinder walls which are such as to *resist* us spinning up the cylinder. This is why we have to do additional work to spin up a charged cylinder, because we have to do work to overcome the forces associated with the induced electric fields, which are in a direction such as to resist change. This is an example of Lenz's Law--the reaction of the system to change is always such as to resist the change.

22.3.2 The work done to spin up the cylinder

How much energy does it take us to do this? We must exert a force in the $\hat{\boldsymbol{\phi}}$ direction to increase the speed of the cylinder. In the following, we ignore the radial part of the electric field, since we do no work against it. To increase the speed of a little bit of charge $+dq$ on the cylinder, we must provide a force $\mathbf{F}_{me} = -dq\,\mathbf{E}$. That is, *we* must provide an additional force in the azimuthal direction that the charge is moving that *balances* the retarding force due to the induced electric field (plus a little teeny bit more, to actually increase the speed of the charge, but we can neglect that little teeny bit). Thus the rate at which we do work, $\frac{dW}{dt} = \mathbf{F}_{me} \cdot \mathbf{V}$, is positive, since $\mathbf{F}_{me} = -dq\,\mathbf{E}$ is in the direction of ***V***. The work to increase the speed of a little $+dq$ is thus

$$\frac{dW}{dt} = \mathbf{F}_{me} \cdot \mathbf{V} = -dqV(t)E_\phi = +dqV(t)\left[\frac{\mu_o R\sigma}{2}\frac{dV(t)}{dt}\right] \tag{22.3.7}$$

$dW/dt = +\,dq\,V(\mu_o \sigma R/2)dV/dt$, where we have used our expression in (22.3.6) for E_ϕ. Since the total charge on the cylinder is $Q = \sigma\left[2\pi rL\right]$, the total rate at which we do work to spin the cylinder up is just found by replacing the *dq* in (22.3.7) by Q, giving

$$\frac{dW}{dt} = \sigma\left[2\pi rL\right]V(t)\left[\frac{\mu_o R\sigma}{2}\frac{dV(t)}{dt}\right] \tag{22.3.8}$$

which after some manipulation and using $B = \mu_o \sigma V(t)$ can be written as

$$\frac{dW}{dt} = \pi R^2 L \frac{d}{dt}\left[\frac{B^2}{2\mu_o}\right] \tag{22.3.9}$$

To get the total energy to spin up the cylinder, we just integrate this expression with respect to time to obtain $\pi R^2 L\left[\dfrac{B_o^2}{2\mu_o}\right]$. Since $\pi R^2 L$ is the volume of our cylinder, this implies an energy density in the magnetic field of $\dfrac{B_o^2}{2\mu_o}$. It is clear where the energy to create this field came from--*us!*

Moreover, this is process is totally reversible--we can get the energy right back out of the magnetic field. To do this, we just grab hold of the spinning cylinder and spin it down. When we try to spin the cylinder down, we are trying to decrease the magnetic flux, and thus we will have a induced electric field that will try to *keep the cylinder spinning*, that is it will reverse direction from the situation above. This is exactly what we expect from Lenz's Law--the induced electric field is in a direction so as to keep things the same. The rate at which we do work, $\dfrac{dW}{dt} = \mathbf{F}_{me} \cdot \mathbf{V}$, will now be negative, since $\mathbf{F}_{me} = -dq\,\mathbf{E}$ reverses sign with the electric field. That means that work is being done on us, and we can use that work to take a free ride, or whatever.

In any case, we are getting energy back out of the process as we spin down the cylinder, and thereby destroying the magnetic fields. With a little thought, it is clear that we will get back exactly the amount of energy we put in the first place to create the magnetic field. Of course we are neglecting radiation losses when we make this statement, but from our discussions in Section 9.5, equation (9.5.7), we know that the energy loss compared to the stored energy is of order $\left(R/cT\right)^3$, and we can make this as small as we desire by simply doing things more and more slowly.

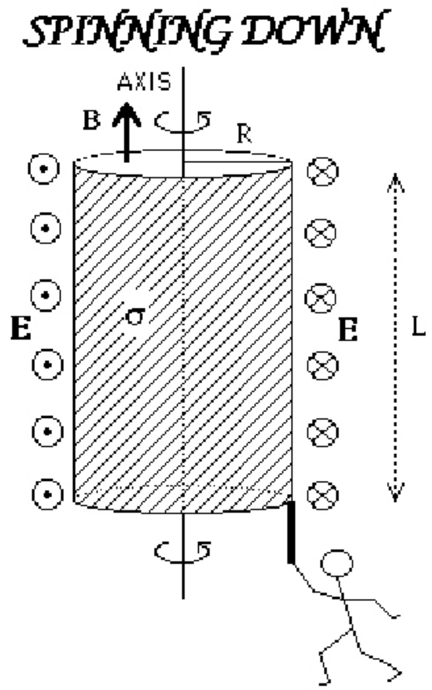


Figure 22-3: Spinning down a cylindrical shell of charge

What we have presented above, based on forces, is an intuitive approach to understanding the origin of energy in magnetic fields. It is makes manifest that to create a magnetic field you must do work against the induced electric fields that are associated with the increasing magnetic flux. Conversely, when you destroy a magnetic field, work is done on you by the induced electric field associated with the decreasing magnetic flux.

22.3.3 Energy flow in spinning up the cylinder

Now, consider this entire process from the point of view of the conservation of energy law given in (4.4.2) in integral form, that is,

$$\frac{\partial}{\partial t}\int_{volume}\left[\frac{1}{2}\varepsilon_o E^2 + \frac{B^2}{2\mu_o}\right]d^3x + \int_{surface}\left(\frac{\mathbf{E}\times\mathbf{B}}{\mu_o}\right)\cdot\hat{\mathbf{n}}\,da = \int_{volume}\left[-\mathbf{E}\cdot\mathbf{J}\right]d^3x \qquad (22.3.10)$$

When we have a surface current, the term

$$\int_{volume}\left[-\mathbf{E}\cdot\mathbf{J}\right]\boldsymbol{\kappa}\,d^3x = \int_{surface}\left[-\mathbf{E}\cdot\quad\right]da \qquad (22.3.11)$$

where $\boldsymbol{\kappa} = \sigma V\hat{\phi}$. This is the term we evaluated above. We are doing work and therefore depositing energy at the circumference of the cylinder. That energy, deposited in a thin cylindrical shell of radius *R*, then flows inward toward the center of the cylinder at a rate given by the Poynting vector $\frac{\mathbf{E}\times\mathbf{B}}{\mu_o}$. We can easily see that the total Poynting flux calculated just inside the surface of the cylinder *r* a little less that *R* is given by

$$\int_{surface}\left(\frac{\mathbf{E}\times\mathbf{B}}{\mu_o}\right)\cdot\hat{\mathbf{n}}\,da=\int_{surface}\left(-\frac{R\sigma}{2}\frac{dV(t)}{dt}B\,\hat{\mathbf{r}}\right)\cdot\hat{\mathbf{r}}\,da=-\pi R^2L\frac{d}{dt}\left[\frac{B^2}{2\mu_o}\right] \tag{22.3.12}$$

Thus we create the energy at $r = R$, and it flows inward to reside in the magnetic field for $r < R$.

22.3.4 Electromagnetic angular momentum for the spinning cylinder

In you think a minute you realize that we should also be creating electromagnetic angular momemtum when we spin up the cylinder, because we have to apply an additional torque to spin up the cylinder because of the "back" emf. But if you look at the density of electromagnetic angular momentum after we are finished, $\mathbf{r}\times\left[\varepsilon_o\mathbf{E}\times\mathbf{B}\right]$, it is zero everywhere. In actual fact you can indeed calculate the creation rate of angular momentum at the shell and the flux of angular momentum outward for $r > R$, and it is *not* zero, but it flows to infinity and is stored in the fringing fields at infinity, which are not accessible to us because we have assumed an infinitely long cylinder. To see the electromagnetic angular momemtum we put into spinning up a distribution of charge, we must take a finite, not infinitely long body, which we do in the next section for a sphere.

22.4 The Spinning Sphere of Charge

22.4.1 The fields of a spinning sphere

We now carry out the same calculations as above except we look at a more realistic situation, which will allow us to see where the electromagnetic angular momentum is stored. Instead of an infinitely long spinning cylinder of charge, we look at a spherical shell of radius *R* that carries a uniform surface charge σ. Its total charge *Q* is $4\pi R^2\sigma$, and its Coulomb electric field is

$$\mathbf{E}_{coulomb}=\begin{cases}0 & r<R\\ \dfrac{Q}{4\pi\varepsilon_o r^2}\hat{\mathbf{r}} & r>R\end{cases} \tag{22.4.1}$$

We begin spinning the sphere at a angular velocity $\omega(t)$ with $\omega R \ll c$. The motion of the charge glued onto the surface of the spinning sphere results in a surface current

$$\boldsymbol{\kappa}(t)=\sigma\,\omega(t)\,R\sin\theta\;\hat{\boldsymbol{\phi}}=\kappa(t)\sin\theta\,\hat{\boldsymbol{\phi}} \tag{22.4.2}$$

where $\kappa(t) = \sigma\,\omega(t)\,R$.

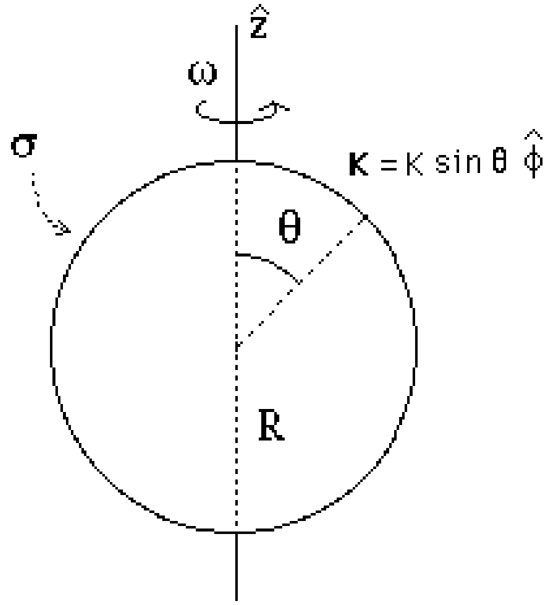


Figure 22-4: A spinning sphere of charge with surface current κ

We can use the quasi-static approximation to get a good approximation to the time dependent solution for **B** (good for variations in *κ(t)* with time scales $T \approx \dfrac{\kappa}{d\kappa/dt} \gg \dfrac{R}{c}$)
If we define

$$m(t) = \frac{4\pi\, R^3}{3}\kappa(t) \qquad B(t) = \frac{2\mu_o}{3}\kappa(t) \tag{22.4.3}$$

so that

$$B(t) = \frac{\mu_o m(t)}{2\pi R^3} \tag{22.4.4}$$

then our quasi-static solution for **B** can be shown to be (see Griffiths Example 5.11 page 236-237)

$$\mathbf{B}(\mathbf{r},t) = \begin{cases} \dfrac{\mu_o m(t)}{4\pi\, r^3}\left(2\cos\theta\ \hat{\mathbf{r}} + \sin\theta\ \hat{\boldsymbol{\theta}}\right) & (r > R) \\ \hat{\mathbf{z}}\ B(t) & (r < R) \end{cases} \tag{22.4.5}$$

We thus have from (22.4.5) that

$$\frac{d\mathbf{B}(\mathbf{r},t)}{dt}=\begin{cases}\dfrac{\mu_o}{4\pi\, r^3}\left(2\cos\theta\,\hat{\mathbf{r}}+\sin\theta\,\hat{\boldsymbol{\theta}}\right)\dfrac{dm(t)}{dt} & (r>R)\\[2ex] \hat{\mathbf{z}}\,\dfrac{d}{dt}B(t) & (r<R)\end{cases} \tag{22.4.6}$$

Given (22.4.6), we can find the induction electric field everywhere in space, as follows.

For $r < R$, take we take a circle whose normal is along the z-axis, whose center is located a distance $r\cos\theta$ up the z-axis, and whose radius is $r\sin\theta$. If we apply Faraday's Law in integral form to that circle, we have

$$2\pi r\sin\theta E_\phi=-\pi\left(r\sin\theta\right)^2\frac{dB}{dt} \tag{22.4.7}$$

or

$$\mathbf{E}_{\text{induction}}=-\hat{\boldsymbol{\phi}}\frac{r\sin\theta}{2}\frac{dB}{dt}=-\hat{\boldsymbol{\phi}}\frac{r\mu_o}{3}\frac{d\kappa}{dt}\sin\theta=-\hat{\boldsymbol{\phi}}\frac{r\mu_o\sin\theta}{4\pi R^3}\frac{dm}{dt} \tag{22.4.8}$$

where we have used (22.4.3).

For $r > R$, if we assume $\mathbf{E}_{induction}=E_\phi\hat{\boldsymbol{\phi}}$, then

$$\nabla\times\mathbf{E}=\hat{\mathbf{r}}\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta E_\phi\right)-\hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial}{\partial r}\left(rE_\phi\right) \tag{22.4.9}$$

Comparing this expression to (22.4.6), we see that from Faraday's Law we have for $r > R$

$$\mathbf{E}_\phi=-\frac{\mu_o\sin\theta}{4\pi r^2}\frac{dm}{dt} \tag{22.4.10}$$

Our complete electric field, Coulomb plus induction field, is thus given by

$$\mathbf{E}=\mathbf{E}_{coulomb}+\mathbf{E}_{induction}=\begin{cases}-\hat{\boldsymbol{\phi}}\dfrac{r\mu_o\sin\theta}{4\pi R^3}\dfrac{dm}{dt} & r<R\\[2ex] \dfrac{Q}{4\pi\varepsilon_o r^2}\hat{\mathbf{r}}-\hat{\boldsymbol{\phi}}\dfrac{\mu_o\sin\theta}{4\pi r^2}\dfrac{dm}{dt} & r>R\end{cases} \tag{22.4.11}$$

It is instructive to compare the induction term for $r > R$ to the Coulomb term

$$\frac{E_{induction}}{E_{coulomb}}=\frac{4\pi\varepsilon_o r^2}{Q}\frac{\mu_o\sin\theta}{4\pi r^2}\frac{dm}{dt}=\frac{\sin\theta}{Qc^2}\frac{d}{dt}\left[\frac{4\pi R^3}{3}\omega R\frac{Q}{4\pi R^2}\right] \tag{22.4.12}$$

$$\frac{E_{induction}}{E_{coulomb}} = \frac{\sin\theta R^2}{3c^2}\frac{d\omega}{dt} = \frac{\omega\sin\theta R^2}{3c^2}\frac{1}{\omega}\frac{d\omega}{dt} \approx \frac{\omega R^2}{3c^2 T} \approx \frac{1}{3}\left(\frac{\omega R}{c}\right)\left(\frac{R}{cT}\right) \qquad (22.4.13)$$

To get the final form in (22.4.13) above we have used the time scale T for changes in the angular velocity defined by

$$\frac{1}{T} = \frac{1}{\omega}\frac{d\omega}{dt} \qquad (22.4.14)$$

Our final result in (22.4.13) above shows that the ratio is the product of two terms, both of which we are assuming to be small, so the ratio of the induction field to the Coulomb is second order small in small quantities.

22.4.2 The total magnetic energy of the spinning sphere

For our purposes below, we want to calculate the magnetic energy outside of R and inside of R. Outside of the sphere

$$\int_{r>R}\left[\frac{B^2}{2\mu_o}\right]d^3x = 2\pi\int_{-1}^{1} d\left(\cos\theta\right)\int_R^\infty\left[\frac{B^2}{2\mu_o}\right]r^2 dr \qquad (22.4.15)$$

and using (22.4.5)

$$\begin{aligned}\int_{r>R}\left[\frac{B^2}{2\mu_o}\right]d^3x &= 2\pi\int_{-1}^{1} d\left(\cos\theta\right)\int_R^\infty\left[\frac{1}{2\mu_o}\left(\frac{\mu_o m}{4\pi\, r^3}\right)^2\left(4\cos^2\theta + \sin^2\theta\right)\right]r^2 dr \\ &= 2\pi\int_{-1}^{1}\left(3\cos^2\theta + 1\right)d\left(\cos\theta\right)\int_R^\infty\left[\frac{\mu_o m^2}{32\pi^2}\frac{1}{r^4}\right]dr \qquad (22.4.16) \\ &= \frac{\mu_o m^2}{12\pi R^3}\end{aligned}$$

In contrast, inside the sphere,

$$\int_{r<R}\left[\frac{B^2}{2\mu_o}\right]d^3x = \frac{B^2}{2\mu_o}\frac{4\pi R^3}{3} = \frac{2\pi R^3}{3\mu_o}\left[\frac{2\mu_o}{3}\kappa(t)\right]^2 = \frac{\mu_o m^2}{6\pi R^3} \qquad (22.4.17)$$

so the magnetic energy inside the sphere is twice that outside the sphere. The total energy we have at time t in the magnetic field is given by the sum of the two terms we calculated above, or $\frac{\mu_o m^2}{6\pi R^3} + \frac{\mu_o m^2}{12\pi R^3}$, which is $\frac{\mu_o m^2}{4\pi R^3}$.

22.4.3 The creation rate of magnetic energy

The total rate at which electromagnetic energy is being created as the sphere is being spun up, $\int\limits_{\text{all space}} -\mathbf{J}\cdot\mathbf{E}\,d^3x$, is given by

$$\int\limits_{\text{all space}} -\mathbf{J}\cdot\mathbf{E}\,d^3x = \int\limits_{\text{all space}} -\left[\hat{\boldsymbol{\phi}}\kappa(t)\sin\theta\delta(r-R)\right]\cdot\left[-\hat{\boldsymbol{\phi}}\frac{\mu_o\sin\theta}{4\pi r^2}\frac{dm}{dt}\right]d^3x$$
$$= \frac{\mu_o\kappa(t)}{2}\frac{dm}{dt}\int_{-1}^{1}\sin^2\theta\, d\left(\cos\theta\right) = \frac{2\mu_o\kappa(t)}{3}\frac{dm}{dt} = \frac{\mu_o m}{2\pi\, R^3}\frac{dm}{dt} = \frac{d}{dt}\left[\frac{\mu_o m^2}{4\pi\, R^3}\right] \quad (22.4.18)$$

so from (22.4.18) we are doing work at just the rate needed to increase the total magnetic energy density we have at a given time. You might worry that we have not taken into account the energy required to create the inductive electric field, but if we estimate how big that energy is, using (22.4.10), we have

$$\frac{1}{2}\varepsilon_o\mathbf{E}_\phi^2 \approx \frac{1}{2}\varepsilon_o\left(\frac{\mu_o}{4\pi R^2}\frac{m}{T}\right)^2 \approx \frac{1}{2}\varepsilon_o\left(\frac{R^3}{4\pi R^2}\frac{B}{T}\right)^2 \approx \left(\frac{R}{cT}\right)^2\frac{B^2}{2\mu_o} \quad (22.4.19)$$

So we can neglect the energy in the induction electric field compared to that in the magnetic field, since it is second order small. If you are spinning up the sphere, it is you who are creating this energy by the additional work you must do to offset the force associated with the induction electric field.

22.4.4 The flow of energy in the spinning sphere

The energy is being created at $r = R$, and we can see if flowing away from its creation site by using the Poynting vector to calculate the flux of electromagnetic energy $\int\limits_{surface}\left[\frac{\mathbf{E}\times\mathbf{B}}{\mu_o}\right]\cdot\hat{\mathbf{r}}\,da$ through a spherical surface of radius r for r a little greater than R and also for r a little smaller than R. First let's do this for r a little greater than R. There we have

$$\mathbf{E}\times\mathbf{B} = \left[\frac{Q}{4\pi\varepsilon_o r^2}\hat{\mathbf{r}} - \hat{\boldsymbol{\phi}}\frac{\mu_o\sin\theta}{4\pi r^2}\frac{dm}{dt}\right]\times\left[\frac{\mu_o m(t)}{4\pi\, r^3}\left(2\cos\theta\,\hat{\mathbf{r}} + \sin\theta\,\hat{\boldsymbol{\theta}}\right)\right]$$
$$= \frac{\mu_o m(t)}{4\pi\, r^3}\left[\frac{Q}{4\pi\varepsilon_o r^2}\sin\theta\left(\hat{\mathbf{r}}\times\hat{\boldsymbol{\theta}}\right) - \frac{\mu_o\sin\theta}{4\pi r^2}\frac{dm}{dt}2\cos\theta\left(\hat{\boldsymbol{\phi}}\times\hat{\mathbf{r}}\right) - \frac{\mu_o\sin^2\theta}{4\pi r^2}\frac{dm}{dt}\left(\hat{\boldsymbol{\phi}}\times\hat{\boldsymbol{\theta}}\right)\right]$$
$$== \frac{\mu_o m(t)}{4\pi\, r^3}\left[\frac{Q}{4\pi\varepsilon_o r^2}\sin\theta\,\hat{\boldsymbol{\phi}} - \frac{\mu_o 2\sin\theta\cos\theta}{4\pi r^2}\frac{dm}{dt}\hat{\boldsymbol{\theta}} + \frac{\mu_o\sin^2\theta}{4\pi r^2}\frac{dm}{dt}\hat{\mathbf{r}}\right]$$
(22.4.20)

So that the energy moving away from the sphere at a distance a little greater than R is

$$\int_{surface} \left[\frac{\mathbf{E}\times\mathbf{B}}{\mu_o}\right]\cdot\hat{\mathbf{r}}\,da = \int_{surface}\left[\frac{1}{\mu_o}\frac{\mu_o m(t)}{4\pi\, R^3}\frac{\mu_o \sin^2\theta}{4\pi R^2}\frac{dm}{dt}\right]da = \frac{\mu_o m(t)}{6\pi\, R^3}\frac{dm}{dt} = \frac{d}{dt}\left[\frac{\mu_o m^2(t)}{12\pi\, R^3}\right] \tag{22.4.21}$$

This is exactly the energy flow outward we need to increase the magnetic energy outside the sphere at a given time. Now let us consider the energy flow across a sphere of radius $r < R$. There we have

$$\mathbf{E}\times\mathbf{B} = \left[-\hat{\boldsymbol{\phi}}\frac{r\mu_o\sin\theta}{3}\frac{d\kappa}{dt}\right]\times\left[\hat{\mathbf{z}}\;B(t)\right] = -\frac{r\mu_o}{3}\frac{d\kappa}{dt}\sin\theta B(t)\left(\hat{\boldsymbol{\phi}}\times\hat{\mathbf{z}}\right)$$

$$\mathbf{E}\times\mathbf{B} = -\frac{r\left(\mu_o\right)^2}{9}\frac{d\kappa^2}{dt}\sin\theta\left(\hat{\boldsymbol{\phi}}\times\left(\hat{\mathbf{r}}\cos\theta - \hat{\ }\sin\theta\right)\right) = -\frac{d}{dt}\frac{r\left(\mu_o\right)^2 m^2}{\left(4\pi R^3\right)^2}\sin\theta\left(\hat{\ }\cos\theta + \hat{\ }\sin\theta\right) \tag{22.4.22}$$

So $$\int_{surface}\left[\frac{\mathbf{E}\times\mathbf{B}}{\mu_o}\right]\cdot\hat{\mathbf{r}}\,da = -2\pi\frac{\mu_o}{\left(4\pi\right)^2 R^3}\frac{dm^2}{dt}\int_{-1}^{1} d(\cos\theta)\sin^2\theta = -\frac{d}{dt}\frac{\mu_o m^2}{6\pi R^3}$$

The minus sign in this equation means the energy flow is inward, and it is exactly the amount we need to account for the rate at which magnetic energy is building up in the interior (see (22.4.17)).

So all of this makes sense, we are creating energy at the shell where we are doing work, and it is flowing out from where we create it at exactly the rate that we need for the build up of magnetic energy inside and outside of the sphere.

Unfortunately, to solve the problem in this relatively simple form, we have had to assume we are doing everything really slowly compared to the speed of light transit time across the sphere, so in none of our terms above do we explicitly see fields propagating at the speed of light. Our solutions just change instantaneously in time everywhere in space, and that is because we have essentially assumed that *c* is infinite to get the tractable expressions above. We can however solve the full problem without making any assumptions as to how *T* compares to *R/c*, but the solutions are much more complicated (see ://web.mit.edu/viz/spin/) When we do this, we obtain solutions in which the energy created at $r = R$ propagates inward and outward at the speed of light, reaching the center of the sphere at a time *R/c* after it was created at the cylinder walls. We show one frame of a visualization of this process, in which you can just see the inward and outward propagation of fields at the speed of light, in Figure 22-5. The full movie for Figure 22-5 can be found at

://web.mit.edu/viz/spin/visualizations/l=1/slow/slow.htm

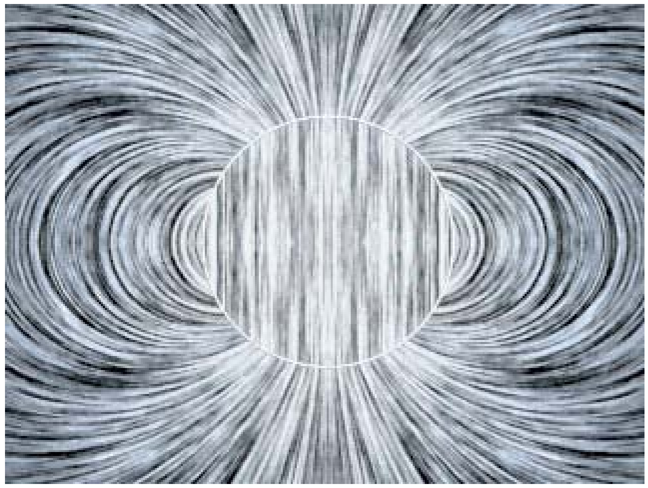


Figure 22-5: One frame of a movie showing the creation of magnetic energy

22.4.5 The electromagnetic angular momentum of the spinning sphere

Let us now look at the conservation of electromagnetic angular momentum, which we could not treat in the cylindrical case because it was stored in the fringing fields at infinity. The form that this conservation law takes is (4.5.7), which we reproduce below

$$\frac{d}{dt}\int_V \mathbf{r}\times\left[\varepsilon_o\mathbf{E}\times\mathbf{B}\right]d^3x+\int_S\left(-\mathbf{r}\times\overset{\leftrightarrow}{\mathbf{T}}\cdot\hat{\mathbf{n}}\right)da=-\int_V \mathbf{r}\times\left[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}\right]d^3x \qquad (22.4.23)$$

If we compute the total electromagnetic angular momentum of this spinning charge configuration, it is

$$\int_{\text{all space}} \mathbf{r}\times\left[\varepsilon_o\mathbf{E}\times\mathbf{B}\right]d^3x \qquad (22.4.24)$$

and we have

$$\mathbf{r}\times\left(\mathbf{E}\times\mathbf{B}\right)=\begin{cases}\mathbf{r}\boldsymbol{\theta}\times\left[-\dfrac{d}{dt}\dfrac{r\left(\mu_o\right)^2 m^2}{\left(4\pi R^3\right)^2}\sin\theta\left(\hat{\ }\cos\theta+\hat{\ }\sin\theta\right)\right] & r<R\\[2ex] \dfrac{\mu_o m(t)}{4\pi\, r^3}\mathbf{r}\times\left[\dfrac{Q}{4\pi\varepsilon_o r^2}\sin\theta\,\hat{\boldsymbol{\phi}}-\dfrac{\mu_o 2\sin\theta\cos\theta}{4\pi r^2}\dfrac{dm}{dt}\hat{\boldsymbol{\theta}}+\dfrac{\mu_o\sin^2\theta}{4\pi r^2}\dfrac{dm}{dt}\hat{\mathbf{r}}\right] & r>R\end{cases}$$

(22.4.25)

$$\mathbf{r}\times(\mathbf{E}\times\mathbf{B})=\begin{cases}\left[-\dfrac{d}{dt}\dfrac{r^2(\mu_o)^2m^2}{(4\pi R^3)^2}\sin\theta\left(\hat{\boldsymbol{\phi}}\cos\theta\right)\right] & r<R\\ -\dfrac{\mu_o m(t)}{4\pi\, r^3}\left[\dfrac{Q}{4\pi\varepsilon_o r}\sin\theta\,\hat{\boldsymbol{\theta}}-\dfrac{\mu_o 2\sin\theta\cos\theta}{4\pi r}\dfrac{dm}{dt}\hat{\boldsymbol{\phi}}\right] & r>R\end{cases} \tag{22.4.26}$$

We can ignore the *dm/dt* terms in (22.4.26) for two reasons: (1) these terms integrate to zero because of the $\hat{\boldsymbol{\phi}}$ dependence; and (b) these terms are small compared to the others. So we have

$$\mathbf{r}\times(\varepsilon_o\mathbf{E}\times\mathbf{B})=\begin{cases}0 & r<R\\ -\dfrac{m(t)}{4\pi\, r^3c^2}\dfrac{Q}{4\pi\varepsilon_o r}\sin\theta\,\hat{\boldsymbol{\theta}}= & r>R\end{cases} \tag{22.4.27}$$

Using $\hat{\boldsymbol{\theta}}=-\hat{\mathbf{z}}\sin\theta+\hat{\mathbf{x}}\cos\theta\cos\phi+\hat{\mathbf{y}}\cos\theta\sin\phi$, and realizing in advance that the *x* and *y* components will average to zero because of the $\cos\phi$ and $\sin\phi$ terms, we have

$$\mathbf{r}\times(\varepsilon_o\mathbf{E}\times\mathbf{B})=\begin{cases}0 & r<R\\ \dfrac{m(t)}{4\pi\, r^3c^2}\dfrac{Q\sin^2\theta}{4\pi\varepsilon_o r}\hat{\mathbf{z}} & r>R\end{cases} \tag{22.4.28}$$

The total electromagnetic angular momentum is located entirely outside the sphere and is given by

$$\begin{aligned}\int\mathbf{r}\times(\varepsilon_o\mathbf{E}\times\mathbf{B})\,d^3x &= \hat{\mathbf{z}}\frac{m(t)Q}{(8\pi\varepsilon_o)\,c^2}\int_R^\infty\frac{dr}{r^2}\int_{-1}^{1}d(\cos\theta)\sin^2\theta\\ &= \hat{\mathbf{z}}\frac{mQ}{(6\pi\varepsilon_o)\,R\,c^2}=\hat{\mathbf{z}}\frac{4\pi R^3\sigma\,\omega RQ}{3(6\pi\varepsilon_o)\,R\,c^2}\end{aligned} \tag{22.4.29}$$

$$\int\mathbf{r}\times(\varepsilon_o\mathbf{E}\times\mathbf{B})\,d^3x=\hat{\mathbf{z}}\frac{\omega RQ^2}{18\pi\varepsilon_o\,c^2} \tag{22.4.30}$$

22.4.6 The creation rate of electromagnetic angular momentum

The total rate at which electromagnetic angular momentum is being created as the sphere is being spun up is according to (22.4.23)

$$\int_{\text{all space}}-\mathbf{r}\times[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}]\,d^3x=\int_{surface}-\mathbf{r}\times\sigma\mathbf{E}\,da=\int_{surface}\left(\mathbf{r}\times\hat{\boldsymbol{\phi}}\right)\sigma\left(\frac{\mu_o\sin\theta}{4\pi R^2}\frac{dm}{dt}\right)da \tag{22.4.31}$$

so that

$$\int_{\text{all space}} -\mathbf{r}\times\left[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}\right]d^3x = -\int_{surface}\hat{\ }\left(\frac{\sigma\mu_o\sin\theta}{4\pi R}\frac{dm}{dt}\right)da$$
$$=\hat{\mathbf{z}}\left(\frac{\sigma\mu_o}{4\pi R}\frac{dm}{dt}\right)\int_{surface}\sin^2\theta\, da = \hat{\mathbf{z}}\left(\frac{\sigma\mu_o}{4\pi R}\frac{dm}{dt}\right)2\pi R^2\frac{4}{3} \tag{22.4.32}$$

$$\int_{\text{all space}} -\mathbf{r}\times\left[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}\right]d^3x = \hat{\mathbf{z}}\frac{d}{dt}\left(\frac{2\sigma\mu_o R\left(4\pi R^3\kappa\right)}{9}\right)$$
$$=\hat{\mathbf{z}}\frac{d}{dt}\left(\frac{2\sigma\mu_o 4\pi R^4\omega R\sigma}{9}\right)=\hat{\mathbf{z}}\frac{d}{dt}\left(\left(\frac{Q}{4\pi R^2}\right)^2\frac{8\mu_o\pi R^5\omega}{9}\right) \tag{22.4.33}$$

$$\int_{\text{all space}} -\mathbf{r}\times\left[\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}\right]d^3x = \hat{\mathbf{z}}\frac{d}{dt}\left(\frac{Q^2\omega R}{18\pi\varepsilon_o c^2}\right) \tag{22.4.34}$$

In the first of the equations above we have used (22.4.11), in the second we have ignored a term that looks like $\hat{\mathbf{r}}\sin\theta\cos\theta$ because it will integrate to zero when we do the theta integration, and in (22.4.34) we have used the definitions of *m* and *κ*. This is the rate we want to see, because when integrated over time it gives the total electromagnetic angular momentum created, as given in (22.4.30). If you are spinning up the sphere, it is you who are creating this angular momentum by the additional torque you must impose to overcome "back" force due to the induction electric field.

22.4.7 The flux of electromagnetic angular momentum

Finally, let us calculate the flux of electromagnetic angular momentum, $\int_{surface}\left[-\mathbf{r}\times\overleftrightarrow{\mathbf{T}}\right]\cdot\hat{\mathbf{r}}\,da$, through a sphere of radius r.

$$\int_{surface}\left[-\mathbf{r}\times\overleftrightarrow{\mathbf{T}}\right]\cdot\hat{\mathbf{r}}\,da = \int_{surface}\left[-\mathbf{r}\times\left(\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{r}}\right)\right]da = \int_{surface}\left[-\mathbf{r}\times\left(T_{rr}\hat{\mathbf{r}}+T_{r\theta}\hat{\ }+T_{r\phi}\hat{\boldsymbol{\phi}}\right)\right]da \tag{22.4.35}$$

For a sphere with *r* a little less than *R*, this flow will be zero, as we expect, because there is no electromagnetic angular momentum in the interior of the sphere when we are finishing spinning up the sphere. For *r* a little greater than *R*, we have

$$\int_{surface} \left[-\mathbf{r}\hat{\boldsymbol{\theta}} \times T_{r\phi}\hat{\boldsymbol{\phi}} \right] da = \hat{\ } \int_{surface} \boldsymbol{\theta} R T_{r\phi}\, da = -\hat{\ }\varepsilon_o \int_{surface} R \frac{Q}{4\pi\varepsilon_o R^2} \frac{\mu_o \sin\theta}{4\pi R^2} \frac{dm}{dt} da$$
$$= -\hat{\boldsymbol{\theta}}\varepsilon_o \frac{d}{dt} \int_{surface} R \frac{Q}{4\pi\varepsilon_o R^2} \frac{\mu_o \sin\theta}{4\pi R^2} \frac{4\pi R^3 \sigma\omega R}{3} da = +\hat{\mathbf{z}} \frac{d}{dt} \frac{Q^2 \omega R}{18\pi c^2 \varepsilon_o} \tag{22.4.36}$$

This is exactly the rate of flow that we want, because when integrated over time it gives us the electromagnetic angular momentum stored outside the sphere when the sphere is fully spun up.

23 The classical model of the electron

23.1 Learning Objectives

We extend the results we have obtained above for the magnetic energy and angular momentum of a spinning shell of charge to the case of a spherical shell of charge with a linear velocity. We compute the linear electromagnetic momentum involved in that motion. Finally, put this all in the context of the attempt in the early 1900's to make a purely electromagnetic model of the electron.

23.2 The momentum and angular momentum of a shell of charge

23.2.1 Angular momentum

In the previous section, we derived the total electromagnetic angular momentum of a spinning uniformly charged shell of radius *R* and total charge *Q*, spinning at an angular speed $\omega\hat{\mathbf{z}}$. The total angular momentum (all of which is contained in the region outside of the shell) is from (22.4.30)

$$\mathbf{L}_E = \int \mathbf{r} \times \left(\varepsilon_o \mathbf{E} \times \mathbf{B} \right) d^3x = \hat{\mathbf{z}} \frac{\omega R Q^2}{18\pi\varepsilon_o\, c^2} \tag{23.2.1}$$

The amount of energy that it took to assemble the charges in this shell, U_E, is given by

$$U_E = \frac{1}{2}\int \rho\phi d^3x = \frac{1}{2}\int \sigma\, \phi da = \frac{Q^2}{8\pi\varepsilon_o R} \tag{23.2.2}$$

Let us define the "electromagnetic inertial mass", m_E, of this charge configuration by the equation

$$m_E c^2 = U_E \Rightarrow m_E = \frac{Q^2}{8\pi\varepsilon_o R c^2} \tag{23.2.3}$$

With this definition, we see that the angular momentu in (23.2.1) can be written as

$$\mathbf{L}_E = \omega\hat{\mathbf{z}}\frac{\omega R Q^2 R}{18\pi\varepsilon_o\, c^2 R} = \hat{\mathbf{z}}\left(\frac{4}{9}m_E R^2\right)\omega = I_E \tag{23.2.4}$$

where the “electromagnetic moment of intertia” is defined by

$$I_E = \frac{4}{9}m_E R^2 \tag{23.2.5}$$

What about the magnetic energy of the spinning spherical shell? In (22.4.18) found that this magnetic energy is given by

$$U_B = \frac{\mu_o m^2}{4\pi R^3} = \frac{2}{9}m_E\omega^2 R^2 = \frac{2}{9}\frac{\omega^2 R^2}{c^2}U_E \tag{23.2.6}$$

so that we have $U_B << U_E$ if $\omega R << c$. Moreover, we can also write (23.2.6) as

$$U_B = \frac{2}{9}R^2 m_E\omega^2 = \frac{1}{2}I_E\omega^2 \tag{23.2.7}$$

At this point, bells start going off in our head’s, and the same bells went off in the heads of physicists in the early 1900’s. From mechanics we remember that a thin shell of mass m has a moment of inertia given by $I = \frac{2}{5}mR^2$. When it is rotating it has rotational kinetic energy of $\frac{1}{2}I\omega^2$ and angular momentum $I\omega$. That looks an awful like our expressions in (23.2.4) and (23.2.7). Maybe mass is just charge?

23.2.2 Linear momentum

Let see if we can get some kind of similar result for linear momentum. We move the spherical shell at velocity $V << c$ in the z direction. Assuming for the moment that the velocity V is constant, we know that in the rest frame of the shell, there is only an electric field. How can we use that fact to find the magnetic field in the laboratory frame, neglecting terms of order $(V/c)^2$? If we return to equation (14.3.15), the fact that the magnetic field in the barred frame is zero means that

$$\overline{B}_x = B_x = 0 \qquad \overline{B}_y = \gamma\left(B_y + \frac{\mathrm{v}}{c^2}E_z\right) = 0 \qquad \overline{B}_z = \gamma(B_z - \frac{\mathrm{v}}{c^2}E_y) = 0 \tag{23.2.8}$$

This set of equations can be written as

$$\mathbf{B} = \mathbf{V} \times \mathbf{E} / c^2 \tag{23.2.9}$$

We also have from (14.3.8) that

$$E_x = \bar{E}_x \qquad E_y = \gamma\left(\bar{E}_y + \mathrm{V}\,\bar{B}_z\right) \qquad E_z = \gamma\left(\bar{E}_z \text{ -} V\bar{B}_y\right) \tag{23.2.10}$$

Since the magnetic field in the barred frame is zero, we thus see from (23.2.10) that the electric field in the laboratory frame is just the Coubomb field in the rest frame of the sphere, neglecting terms of order $\left(V/c\right)^2$.

Now, let us consider a time in the lab frame where the spherical shell is just at the origin. Then the total amound of linear momentum in the fields of the sphere at that time is given by

$$\mathbf{P}_E = \int \varepsilon_o \mathbf{E} \times \mathbf{B}\, d^3x = \int \varepsilon_o \mathbf{E} \times \left[\mathbf{V} \times \mathbf{E}\right] / c^2 d^3x \tag{23.2.11}$$

We only need to do this integral over the exterior of the sphere since the electric field vanishes in the interior, so we have

$$\mathbf{P}_E = \frac{V}{c^2}\int \varepsilon_o \left(\frac{Q}{4\pi\varepsilon_o r^2}\right)^2 \hat{\mathbf{r}} \times \left[\hat{\mathbf{z}} \times \hat{\mathbf{r}}\right] r^2 da = \frac{V}{c^2}\int \varepsilon_o \left(\frac{Q}{4\pi\varepsilon_o r^2}\right)^2 \hat{\mathbf{r}} \times \left[\sin\theta\,\hat{\boldsymbol{\phi}}\right] r^2 da \tag{23.2.12}$$

$$\mathbf{P}_E\boldsymbol{\theta} = -\frac{V}{c^2}\int \varepsilon_o \left(\frac{\mathbf{z}\;Q}{4\pi\varepsilon_o r^2}\right)^2 \hat{\ } \sin\theta\, r^2 da = \hat{\ } \frac{2\pi V\varepsilon_o}{c^2}\left(\frac{Q}{4\pi\varepsilon_o}\right)^2 \int_{-1}^{1} \sin^2\theta d\left(\cos\theta\right)\int_R^\infty \frac{dr}{r^2} \tag{23.2.13}$$

$$\mathbf{P}_E = \hat{\mathbf{z}}\frac{4}{3}V\left(\frac{Q^2}{8\pi\varepsilon_o c^2 R}\right) = \frac{4}{3}m_E\mathbf{V} \tag{23.2.14}$$

So we see that the moving spherical shell has a net linear electromagnetic momentum associated with it, which looks something like its “electromagnetic mass” times its velocity, except for the funny fact of $\frac{4}{3}$.

23.3 The self force on an accelerating set of charges

Where did this linear electromagnetic momentum stored in the fields of the moving shell come from? We are going to make the following statement, and then give an example. Suppose you have an electrostatic problem where you have assembled a set of charges and in doing this you had to expend energy U_E. We assume that the charges are held in place by some rigid framework, and that if we try to move them they will move as a unit. Now suppose you come along and "kick" this distribution of charges, applying a force to it of $\mathbf{F}_{me}$. Then to total force on the charge will be given in part[6] by

$$\mathbf{F} = \mathbf{F}_{me} - \xi \frac{U_E}{c^2}\frac{d\mathbf{V}}{dt} = \mathbf{F}_{me} - \xi m_E \frac{d\mathbf{V}}{dt} \tag{23.3.1}$$

where ξ is a dimensionless factor of order unity and $m_E = U_E / c^2$. What this means is that a set of charges such as this resists being moved. That is, it has an inertia, and the order of magnitude of the "back reaction" to being moved is given in (23.3.1).

23.3.1 The electromagnetic inertia of a capacitor

The statement above in (23.3.1) is a generally true. Here we want to give a specific example with very idealized geometry to give you some feel for why this inertial "back reaction" exists for a charged object. Consider a capacitor oriented as shown in Figure 0-1, and moving upward. The plates of the capacitor have area *A*, and the distance between the plates is *d*. The right plate carries a charge per unit area of $+\sigma$, and the left plate carries a charge per unit area $-\sigma$. When the capacitor is at rest, the electric field between the plates is $\mathbf{E}_o = -\frac{\sigma}{\varepsilon_o}\hat{\mathbf{x}}$. The total electrostatic energy in the capacitor is thus

$$U_E = Ad\left(\frac{1}{2}\varepsilon_o E_o^2\right) = Ad\left(\frac{1}{2}\varepsilon_o\left(\frac{Q}{\varepsilon_o A}\right)^2\right) = \frac{d\,Q^2}{2\varepsilon_o A} \tag{23.3.2}$$

Now suppose we have gotten the capacitor up to speed $V\,\hat{\mathbf{y}}$. Then the upward motion of the positive sheet of charge on the right will correspond to a current sheet with current per unit length $\sigma V\,\hat{\mathbf{y}}$, and the upward motion of the negative sheet of charge on the left will correspond to a current sheet with current per unit length $-\sigma V\,\hat{\mathbf{y}}$. As a result of these current sheets, there will be a magnetic field between the sheets of current given by

$$\mathbf{B} = \mu_o \sigma V\,\hat{\mathbf{z}} \tag{23.3.3}$$

Now suppose we try to increase the speed of the sheet. This will result in an increase in the magnetic field given in (23.3.3), and therefore to the presence of an induced electric field. If we take a loop in the *xy* plane centered on the x axis with width *2x* in the *x*-direction and length *L* in the *y* direction, and integrate around that loop, the Faraday's Law gives, assuming $x < d/2$

---

[6] We say in part because we are neglecting the "radiation reaction" force, we considered above in Section 17.8, and which is proportional to the time derivative of the acceleration.

$$\oint \mathbf{E}\cdot d\mathbf{l} = -\frac{d}{dt}\left[L2xB\right] = -\frac{d}{dt}\left[2xL\mu_o\sigma V\right] = -2xL\mu_o\sigma\frac{dV}{dt} \qquad (23.3.4)$$

A little thought will convince you that this induced electric field is in the *y*-direction and an antisymmetic function of *y*, with the induced field for $|x| < d/2$ given by

$$\mathbf{E}_{induced} = -x\mu_o\sigma\frac{dV}{dt}\hat{\mathbf{y}} = -x\mu_o\frac{Q}{A}\frac{dV}{dt}\hat{\mathbf{y}} \qquad (23.3.5)$$

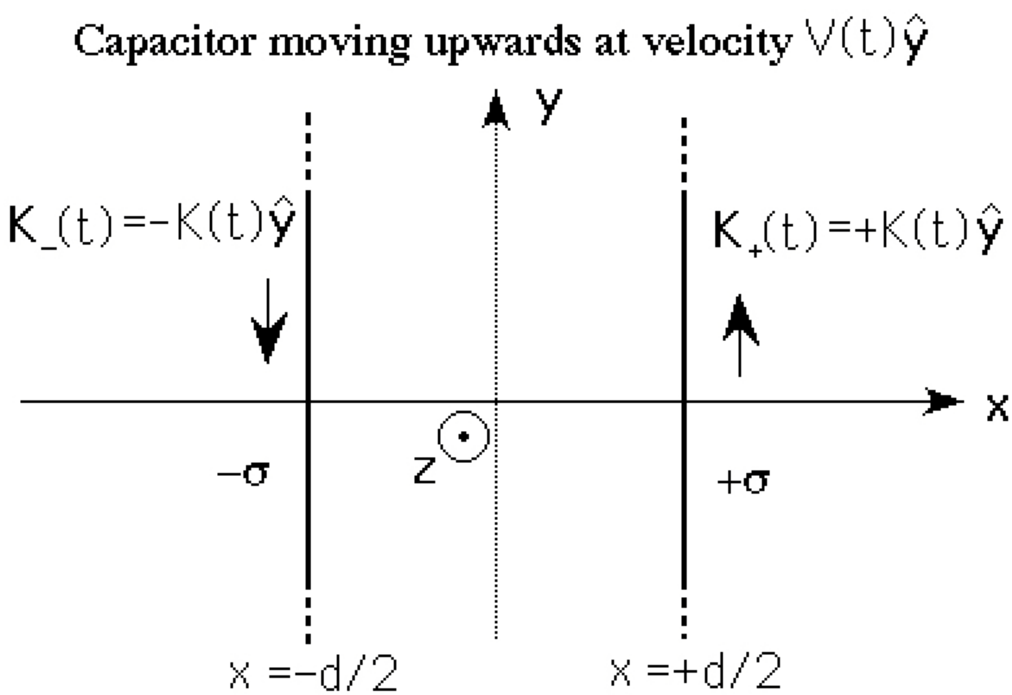


Figure 0-1: A Capacitor Moving Upward

This induced field will exext a downward force on the positive sheet at $x = d/2$ and a downward force on the negative sheet at $x = -d/2$ (because the induced electric field has switched diretion there), and therefore if we try to accelerate the capacitor the induced electric field resists that acceleration with a force given by

$$\mathbf{F}_{back} = -\hat{\mathbf{y}}2Q\frac{d}{2}\mu_o\frac{Q}{A}\frac{dV}{dt} = -\frac{dQ^2}{\varepsilon_o Ac^2}\frac{dV}{dt}\hat{\mathbf{y}} = -\frac{2U_E}{c^2}\frac{dV}{dt}\hat{\mathbf{y}} = -2m_E\frac{d\mathbf{V}}{dt} \qquad (23.3.6)$$

where we have used (23.3.2).

Equation (23.3.6) is the form given in (23.3.1), as we asserted that it must be. In this highly idealized geometry we see why this back reaction force arises. If we try to move a set of charges, we will create currents which will generate magnetic fields proportional to velocity. If we try to accelerate the set of charges, we will thus have a time changing magnetic field which will be associated with induced electric fields which are proportion to the time rate of change of the magnetic fields, or to the acceleration. These induced electric fields will exert self forces in a direction such as to resist any attempt to move the charges. That is why there is an additional electromagnetic momentum (23.2.14) in our moving spherical shell, and where it comes from. It as usually comes from me, the agent who got the shell to move, who had to exert an addition force because the shell was charged, and that force is the source of the additional momentum in the shell.

It is important to note that this process is *reversible*. That is, if an object is charged we have to put extra energy into getting it moving, but when we stop it we get all of that energy back. This is very different from the radiation reaction force, which is always a drain of energy. This is clear from out moving capacitor example above. When we try to stop this moving capacitor, we will be reducing the amount of current flowing and thus the magnetic field those currents produce, and thus there will be an associated induction electric field which will now try to kept the capacitor moving. That additional inertia due the charging means as the capacitor slows down we can use the induced electric field to do work on us, and the work that is done on us is exactly the additional work we had to put in to get the capacitor up to speed because of the charge.

23.3.2 The classical model of the electron

What do we mean when we talk about the inertial mass of an uncharged object? What we mean is that if we appy a known force $\mathbf{F}_{me}$ to the object, the resultant acceleration will be in the same direction as the force, and if we measure the acceleration *a* of the object due to this force, then the inertial mass of the object is given by

$$m_o = \frac{F_{me}}{a} \qquad (23.3.7)$$

where we are using m_o to denote the mass of the uncharged object. We will always get this same ratio, regardless of the magnitude of the force applied. That is, if we double F_{me}, we will observe double the acceleration *a*, giving the same interial mass as before.

If we now charge up this object, and measure its inertial mass after we have charged it up, then comparing (23.3.7) to (23.3.1), we see that the inertial mass after we have charged it up will increase, because the back reaction will have decreased the acceleration we observe for the same external force applied, $\mathbf{F}_{me}$. That is, we always have that

$$\mathbf{F} = m_o\mathbf{a} = \mathbf{F}_{me} - \xi\frac{U_E}{c^2}\mathbf{a} \qquad (23.3.8)$$

where the last expression comes from (23.3.1). If we solve (23.3.8) for **a** in terms of the applied force, we have

$$\left(m_o + \xi \frac{U_E}{c^2} \right) \mathbf{a} = \mathbf{F}_{me} \tag{23.3.9}$$

Or an inertial mass *m* of

$$m = \frac{F_{me}}{a} = m_o + \xi \frac{U_E}{c^2} \tag{23.3.10}$$

Thus when we charge up a neutral object and then measure its inertial mass, we see an increase in that mass.

The great excitement in the early 1900's was that when physicists realized that a charged object had properties that looked exactly like "ordinary" matter, e.g. momentum, angular momentum, and inertia, they suddenly realized that perhaps there was no "ordinary" matter at all, but that everything was simply electromagnetic in character. We know the mass of the electron m_e and its charge e, and if it were a ball of charge of radius r_c given by

$$r_e = \frac{e^2}{4\pi\varepsilon_o m_e c^2} \tag{23.3.11}$$

then the expression in (23.3.10) would account for all of its inertial mass, and we would not need the "neutral" inertial mass m_o.

The reason this does not work is that of course you need something besides electromagnetic fields to hold together an electron. If we only had electric fields, for example, the charge making up an electron would fly off to inifinty due to mutual repulsion. So there must be something else that is holding the electron together, and those stresses would also contribute to the inertia. The strange factors we get above (like the factor of 4/3 in (23.2.14) for one thing means that we can not construct a four momentum for a purely electrostatic electron that transforms correctly under Lorentz transformations, another indication that we are leaving out something significant.

In any case, with the advent of quantum mechanics, all of these classical models went by the boards. But it is worth pointing out that the length scale defined by (23.3.11) appears repeatedly in classical electromagnetic radiation calculations. For example, as you will show in Assignment 10, the square of this distance is proportional to the Thompson cross-section for an electron, that is its cross-section for scattering energy out of an incident electromagnetic wave.

24 Interactions with matter I

24.1 Learning Objectives

We begin our discussion of the interactions of electromagnetic fields with matter. First we look at insulators, that is dielectrics and magnetic materials.

24.2 The Displacement Current

Before we begin actually discussing matter, we first consider the origins of the displacement current term introduced by Maxwell. Before Maxwell, Ampere's Law was well known, which has the differential form

$$\nabla \times \mathbf{B} = \mu_o \mathbf{J} \tag{24.2.1}$$

Maxwell noted that this equation must be missing a term because if we take the divergence of both sides of (24.2.1), since the divergence of any curl is zero, we have

$$\nabla \cdot \mathbf{J} = 0 \tag{24.2.2}$$

But this cannot be because charge conservation in differential form is

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} \tag{24.2.3}$$

Maxwell suggested that (24.2.1) should instead be

$$\nabla \times \mathbf{B} = \mu_o \left(\mathbf{J} + \varepsilon_o \frac{\partial \mathbf{E}}{\partial t} \right) \tag{24.2.4}$$

With this term and $\nabla \cdot \mathbf{E} = \rho / \varepsilon_o$, we no longer have a problem with the divergence of (24.2.4) being zero on both sides. The term $\varepsilon_o \frac{\partial}{\partial t}\mathbf{E}$ in equation (24.2.4) has units of current density and is called the displacement current, ***even though it has nothing to do with displacement and nothing to do with current.*** The reason it is called that is because Maxwell knew about the polarization current, which we discuss below in (24.3.4), and he could ascribe this current to the real motion of charges producing a real current. He hypothesized that the aether was made up of positive and negative changes and that an electric field would produce a similar "displacement current" by moving those charges around. There is no such thing, and the concept is totally wrong. Nonetheless we continue to call $\varepsilon_o \frac{\partial}{\partial t}\mathbf{E}$ the displacement current.

24.3 The average dipole moment per unit volume

We begin our consideration of the interaction of material media with electromagnetic fields by realizing that for the most part matter is neutral, and that the interaction with fields is not through any net charge but through the effects of electric and magnetic dipoles. We are motivated therefore to define the two vectors **P** and **M** as the electric dipole moment per unit volume and the magnetic dipole moment per unit volume. That is if we have a small volume ΔV of material located at **r** at time *t*, with *N* electric dipoles $\{\mathbf{p}_i\}_{i=1}^N$ and *N* magnetic dipoles $\{\mathbf{m}_i\}_{i=1}^N$, then we define **P** and **M** at (**r**,t) as

$$\mathbf{P}(\mathbf{r},t) = \frac{1}{\Delta V}\sum_{i=1}^{N}\mathbf{p}_i \qquad\qquad \mathbf{M}(\mathbf{r},t) = \frac{1}{\Delta V}\sum_{i=1}^{N}\mathbf{m}_i \tag{24.3.1}$$

Here we assume the averaging volume ΔV is large enough to contain many particles, but small enough that any macroscopic variation in the material medium is on scales much larger than $\Delta V^{1/3}$. Since the dimensions of electric dipole moment are charge time length, **P** has units of charge per area. Similary, since the dimensions of magnetic dipole moment are current time area, **M** has units of current per length.

A variation of $\mathbf{P}(\mathbf{r},t)$ with space can be an accompanied by a "polarization" charge density. The easiest way to see this is to think of a long cylinder of uniformly polarized material with cross sectional area *A* and length *l* (see the left image of Figure 24-0-1). The total dipole moment of this cylinder is lAP, and we expect therefore from far away it should look like a charge $+Q$ on the top end accompanied by a charge $-Q$ on the bottom end such that the dipole moment of this arrangement, Ql, is equal to lAP. This means that there must be a dipole surface charge density on the ends of the cylinder of $\sigma_{pol} = Q/A = P$.

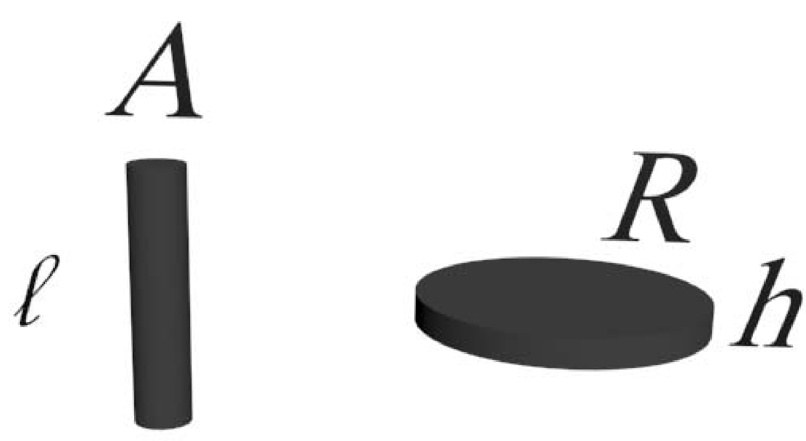


Figure 24-0-1: A uniform polarized "needle" and "disk"

Similary, a variation of $\mathbf{M}(\mathbf{r},t)$ with space can be accompanied by a "magnetization" current. . The easiest way to see this is to think of a disk of uniformly magnetized material with cross sectional radius *R* and height *l* (see the right image of Figure 24-0-1). The total magnetic dipole moment of this cylinder is $h\pi R^2 M$, and we expect therefore from far away it should look like a disk with a current *I* running around its circumference such that the dipole moment $I\pi R^2$ of this arrangement is equal to $h\pi R^2 M$. This means that there must be a current per unit height of the disk given by $\kappa_{mag} = I/h = M$.

When we go over to a non-uniform **P** amd **M**, then we expect there to be corresponding polarization charge densities and current densities given by

$$\rho_{pol} = -\nabla \cdot P \qquad\qquad \mathbf{J}_{mag} = \nabla \times \mathbf{M} \tag{24.3.2}$$

At the boundary of a dielectric or magnetic material, we will have polarization surface charges and magnetization surface currents given by

$$\sigma_{pol} = \hat{\mathbf{n}} \cdot \mathbf{P} \qquad\qquad \boldsymbol{\kappa}_{mag} = \mathbf{M} \times \hat{\mathbf{n}} \tag{24.3.3}$$

where $\hat{\mathbf{n}}$ is the normal pointing out of the material medium.

The polarization charge density in (24.3.2) will be accompanied by a polarization current to conserve polarization charge, that is there is a $\mathbf{J}_{pol}$ such that

$$\frac{\partial \rho_{pol}}{\partial t} + \nabla \cdot \mathbf{J}_{pol} = 0 \qquad\qquad \Rightarrow \mathbf{J}_{pol} = \frac{\partial \mathbf{P}}{\partial t} \tag{24.3.4}$$

Since $\nabla \cdot \mathbf{J}_{mag} = \nabla \cdot \left(\nabla \times \mathbf{M} \right) = 0$, there is no corresponding "magnetization" charge density, nor would we expect one.

24.4 Unformly polarized spheres

24.4.1 A dielectric sphere

Suppose we have a uniformly polarized dielectric sphere of radius *R*, with the direction of polarization in the z-direction. The polarization surface charge on the surface of the sphere will be

$$\sigma_{pol} = \hat{\mathbf{n}} \cdot \mathbf{P} = \hat{\mathbf{r}} \cdot P\hat{\mathbf{z}} = P\cos\theta \tag{24.4.1}$$

We have seen that a surface charge on a sphere with this angular dependence will produce a sphere which has an electric field given by

$$\mathbf{E} = \begin{cases} -E_o \hat{\mathbf{z}} & r < R \\ \left[\dfrac{2 p_o \cos\theta}{4\pi\varepsilon_o r^3} \hat{\mathbf{r}} + \dfrac{p_o \sin\theta}{4\pi\varepsilon_o r^3} \hat{\boldsymbol{\theta}} \right] & r > R \end{cases} \tag{24.4.2}$$

where

$$E_o = \frac{P}{3\varepsilon_o} \qquad\qquad p_o = \frac{4\pi R^3}{3} P \tag{24.4.3}$$

24.4.2 A magnetized sphere

Suppose we have a uniformly polarized magnetic sphere of radius *R*, with the direction of polarization in the z-direction. The polarization surface current on the surface of the sphere will be

$$\kappa_{mag} = \mathbf{M} \times \hat{\mathbf{n}} = M\, \hat{\mathbf{z}} \times \hat{\mathbf{r}} = M \sin\theta\, \hat{\boldsymbol{\phi}} \tag{24.4.4}$$

We have seen that a surface current on a sphere with this angular dependence will produce a sphere which has a magnetic field given by

$$\mathbf{B} = \begin{cases} B_o \hat{\mathbf{z}} & r < R \\ \left[\dfrac{\mu_o}{4\pi} \dfrac{2 m_o \cos\theta}{r^3} \hat{\mathbf{r}} + \dfrac{\mu_o}{4\pi} \dfrac{m_o \sin\theta}{r^3} \hat{\boldsymbol{\theta}} \right] & r > R \end{cases} \tag{24.4.5}$$

where

$$B_o = \frac{2}{3} \mu_o M \qquad\qquad m_o = \frac{4\pi R^3}{3} M \tag{24.4.6}$$

24.5 Maxwell's Equations in the presence of matter

So the presence of spatial and temporal variations in **P** and **M** mean that we have to modify Maxwell's equations to include the assoicated currents and charge densities, since **E** and **B** are produced by all currents and charges, regardless of from which they arise. That is if ρ_{free} and $\mathbf{J}_{free}$ are the "free" charge density and current density (that is, the ones we control) we have

$$\rho = \rho_{free} + \rho_{pol} = \rho_{free} - \nabla \cdot \mathbf{P} \tag{24.5.1}$$

$$\mathbf{J} = \mathbf{J}_{free} + \mathbf{J}_{mag} + \mathbf{J}_{mag} = \mathbf{J}_{free} + \nabla \times \mathbf{M} + \frac{\partial \mathbf{P}}{\partial t} \tag{24.5.2}$$

Maxwell's equations with these additional terms now become become (Faraday's Law and the divergence of B is zero are unchanged)

$$\nabla \cdot \mathbf{E} = \frac{\rho_{free}}{\varepsilon_o} - \frac{\nabla \cdot \mathbf{P}}{\varepsilon_o} \tag{24.5.3}$$

$$\nabla \times \mathbf{B} = \mu_o \mathbf{J}_{free} + \mu_o \nabla \times \mathbf{M} + \mu_o \frac{\partial \mathbf{P}}{\partial t} + \mu_o \varepsilon_o \frac{\partial \mathbf{E}}{\partial t} \tag{24.5.4}$$

or

$$\nabla \cdot \left(\varepsilon_o \mathbf{E} + \mathbf{P} \right) = \rho_{free} \tag{24.5.5}$$

$$\nabla \times \left(\frac{\mathbf{B}}{\mu_o} - \mathbf{M} \right) = \mathbf{J}_{free} + \frac{\partial}{\partial t} \left(\mathbf{P} + \varepsilon_o \mathbf{E} \right) \tag{24.5.6}$$

We define

$$\mathbf{D} = \varepsilon_o \mathbf{E} + \mathbf{P} \tag{24.5.7}$$

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_o} - \mathbf{M} \qquad \text{or} \qquad \mathbf{B} = \mu_o \left(\mathbf{H} + \mathbf{M} \right) \tag{24.5.8}$$

So then our equations are

$$\nabla \cdot \mathbf{D} = \rho_{free} \tag{24.5.9}$$

$$\nabla \times \mathbf{H} = \mathbf{J}_{free} + \frac{\partial}{\partial t} \mathbf{D} \tag{24.5.10}$$

24.6 Linear media

We assume that we have media where the only polarization arises *because of external applied fields.* Note that this is *very* different from the situation we discussed in Section 24.4 above, where somehow the dipoles were aligned on their own (e.g. as in a permanent magnet) and we then looked to see what fields this polarization produced. Not only are we going to suppose that the polarization arises because of external fields, but we are going to assume there is a linear relationship between the external field and the resultant polarization, that is we assume that

$$\mathbf{P} = \varepsilon_o \chi_e \mathbf{E} \tag{24.6.1}$$

$$\mathbf{M} = \chi_m H \tag{24.6.2}$$

You might argue that if we are going to make **P** proportional to **E** we should corrrenspondingly make **M** proportional to **B**, not **H**, but this is the custom and we do not deviate from it. The constant of proportionality χ_e is called the electric susceptibility and the constant of proportionality χ_m is called the magnetic susceptibility. These are dimensionless quantities that vary from one substance to another. If we go back to (24.5.7) and (24.5.8), we see that for linear media, we have

$$\mathbf{D} = \varepsilon_o \mathbf{E} + \mathbf{P} = \varepsilon_o \mathbf{E} + \varepsilon_o \chi_e \mathbf{E} = \varepsilon_o \left(1 + \chi_e\right) = \varepsilon \mathbf{E} \tag{24.6.3}$$

$$\mathbf{B} = \mu_o \left(\mathbf{H} + \mathbf{M}\right) = \mu_o \left(1 + \chi_m\right) \mathbf{H} = \mu \mathbf{H} \tag{24.6.4}$$

where

$$\mu = \mu_o \left(1 + \chi_m\right) \tag{24.6.5}$$

$$\varepsilon = \varepsilon_o \left(1 + \chi_e\right) \tag{24.6.6}$$

We call ε the permittivity of the material, and μ the permeability of the material. We also call $\left(1 + \chi_e\right)$ the dielectric constant of the material, defined as

$$K_e = 1 + \chi_e \tag{24.6.7}$$

In the case of linear materials, we can therefore write Maxwell's equations as

$$\nabla \cdot \varepsilon \mathbf{E} = \rho_{free} \tag{24.6.8}$$

$$\nabla \times \mathbf{B} = \mu \mathbf{J}_{free} + \mu \varepsilon \frac{\partial}{\partial t} \mathbf{E} \tag{24.6.9}$$

24.7 Boundary conditions of E, B, H, and D and spheres in uniform fields

24.7.1 Boundary conditions

If you look at our boundary conditions on E and B in Section Sections 19.6 and 20.7 above, it is clear that with material media, the boundary conditions on D, H, B, and E are

$$\mathbf{E}_{2t} = \mathbf{E}_{1t} \tag{24.7.1}$$

$$B_{2n} = B_{1n} \tag{24.7.2}$$

$$\mathbf{H}_{2t} - \mathbf{H}_{1t} = \mathbf{K}_{free} \times \hat{\mathbf{n}} \tag{24.7.3}$$

$$D_{2n} - D_{1n} = \hat{\mathbf{n}} \cdot \left(\mathbf{D}_2 - \mathbf{D}_1\right) = \sigma_{free} / \varepsilon_o \tag{24.7.4}$$

24.7.2 A linear dielectric sphere in a uniform field

Suppose we have sphere made out of linear dielectric material with dielectric constant K_e sitting in a constant field in the z-direction, $\mathbf{E} = E_o\hat{\mathbf{z}}$. We are going to guess our solution is of the form

$$\mathbf{E} = \begin{cases} E_1\hat{\mathbf{z}} & r < R \\ \begin{bmatrix} E_o\hat{\mathbf{z}} + \dfrac{2p_1\cos\theta}{4\pi\varepsilon_o r^3}\hat{\mathbf{r}} \\ +\dfrac{p_1\sin\theta}{4\pi\varepsilon_o r^3}\hat{\boldsymbol{\theta}} \end{bmatrix} & r > R \end{cases} \tag{24.7.5}$$

where the unknowns here are E_1 and p_1, and we will determine these constants using the boundary conditions in the problem. At the poles we have from (24.7.4) that

$$\varepsilon_o K_e E_1 = \varepsilon_o\left(E_o + \frac{2p_1}{4\pi\varepsilon_o R^3}\right) \tag{24.7.6}$$

and at the equator we have from (24.7.1) that

$$E_1 = E_o - \frac{p_1}{4\pi\varepsilon_o R^3} \tag{24.7.7}$$

Solving these equations gives

$$p_1 = 4\pi\varepsilon_o R^3 E_o \frac{\left(K_e - 1\right)}{\left(K_e + 2\right)} \tag{24.7.8}$$

$$E_1 = E_o \frac{3}{\left(2 + K_e\right)} \tag{24.7.9}$$

In the limit that $K_e = 1$, we recover the field we expect. In the limit that $K_e = \infty$, we have zero field inside the sphere. This limit is the same as if the sphere were made out of conducting material. We can find the polarization surface charge now that we have solved the problem by computing

$$\sigma_{pol} = \varepsilon_o\left(E_{2n} - E_{1n}\right) = \left(\varepsilon_o E_o\cos\theta + \frac{2p_1\cos\theta}{4\pi\varepsilon_o r^3}\right) - \varepsilon_o E_1\cos\theta \tag{24.7.10}$$

$$\sigma_{pol} = \varepsilon_o \cos\theta\left(E_o - E_1 + \frac{2p_1}{4\pi\varepsilon_o R^3}\right) = \varepsilon_o E_o \cos\theta\left(1 - \frac{3}{(2+K_e)} + \frac{2(K_e - 1)}{(K_e + 2)}\right)$$
$$\sigma_{pol} = \varepsilon_o E_o \cos\theta \frac{3(K_e - 1)}{(2+K_e)} \tag{24.7.11}$$

Again we get the corrent limits when $K_e = 1$ and when $K_e = \infty$. When $K_e = 1$ the polarization charge vanishes, as we expect, and when $K_e = \infty$ the polarization charge is $\sigma_{pol} = 3\varepsilon_o E_o \cos\theta$, which is what we need to make the electric field just outside the sphere at the pole to drop from $3E_o$ to zero just inside the pole. The field in this case ($K_e = \infty$) is

$$\mathbf{E} = E_o \begin{cases} 0 & r < R \\ \left[\begin{array}{l} \hat{\mathbf{z}} + 2\cos\theta\left(\dfrac{R}{r}\right)^3 \hat{\mathbf{r}} \\ +\sin\theta\left(\dfrac{R}{r}\right)^3 \hat{\boldsymbol{\theta}} \end{array}\right] & r > R \end{cases} \tag{24.7.12}$$

We see that the presence of the linear dielectric decreased the electric field inside the sphere below what it would have otherwise be with out the dielectric, as we expect. In the corresponding magnetic material case (a sphere with permeability μ sitting in a constant magnetic field $\mathbf{B} = B_o\hat{\mathbf{z}}$), the magnetic field inside the sphere will be enhanced instead of diminished, as we expect.

24.8 Conservation of energy in linear media

We can carry out the following formal manipulations, assuming μ and ε are not functions of space, and using (24.5.10)

$$\begin{aligned} &\nabla\cdot(\mathbf{E}\times\mathbf{H}) = \mathbf{H}\cdot(\nabla\times\mathbf{E}) - \mathbf{E}\cdot(\nabla\times\mathbf{H}) \\ &= -\mathbf{H}\cdot\frac{\partial\mathbf{B}}{\partial t} - \mathbf{E}\cdot\left(\mathbf{J}_{free} + \frac{\partial}{\partial t}\mathbf{D}\right) = -\mathbf{H}\cdot\frac{\partial\mathbf{B}}{\partial t} - \mathbf{E}\cdot\frac{\partial}{\partial t}\mathbf{D} - \mathbf{E}\cdot\mathbf{J}_{free} \end{aligned} \tag{24.8.1}$$

Or

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\mathbf{E}\cdot\mathbf{D} + \frac{1}{2}\mathbf{B}\cdot\mathbf{H}\right) + \nabla\cdot(\mathbf{E}\times\mathbf{H}) = -\mathbf{E}\cdot\mathbf{J}_{free} \tag{24.8.2}$$

This suggests that we take the energy density in linear media to be

$$u = \frac{1}{2}\mathbf{E}\cdot\mathbf{D} + \frac{1}{2}\mathbf{B}\cdot\mathbf{H} \tag{24.8.3}$$

and the energy flux density to be

$$\mathbf{S} = \mathbf{E}\times\mathbf{H} \tag{24.8.4}$$

We could write down similar formal equations for the conservation of momentum, but this is a complicated subject and we refer the reader to e.g. Jackson's treatment.

24.9 Propagation speed of electromagnetic waves

Suppose we have no free currents or charges and are looking at the propagation of waves in the media. The taking the curl of (24.6.9), we have

$$\nabla\times\left(\nabla\times\mathbf{B}\right) = \mu\varepsilon\frac{\partial}{\partial t}\nabla\times\mathbf{E} = -\mu\varepsilon\frac{\partial^2\mathbf{B}}{\partial t^2} = -\nabla^2\mathbf{B} \tag{24.9.1}$$

where we have used Faraday's Law and the fact that $\nabla\cdot\mathbf{B} = 0$ in (24.9.1). We thus have

$$\nabla^2\mathbf{B} - \mu\varepsilon\frac{\partial^2\mathbf{B}}{\partial t^2} = 0 \tag{24.9.2}$$

which means that a plane wave will no longer propagate at speed speed of *c*, but rather at the speed

$$\frac{\omega}{k} = \frac{1}{\sqrt{\mu\varepsilon}} = \frac{c}{\sqrt{\left(1+\chi_e\right)\left(1+\chi_m\right)}} \tag{24.9.3}$$

We define the index of refraction *n* of a medium by

$$n = \frac{c}{\omega/k} = \sqrt{\left(1+\chi_e\right)\left(1+\chi_m\right)} \tag{24.9.4}$$

25 Why is the speed of light not *c* in a dielectric?

25.1 Learning Objectives

We want to discuss the propagation of light when there is matter present. We look at the case where we only have a dielectric present, which leads to electromagnetic plane waves propagating at speeds different from the speed of light, as we saw in (24.9.3). We want to see what this means and how it is possible. To this end, we will make a "model" of a dielectric medium where we deduce a physical mechanism producing the dipole moment per unit volume **P**, and thus offer up a physical basis for the origin of (24.6.1).

25.2 Review of plane wave generated by an oscillating sheet of current

We know that a current sheet oscillating in the *y*-direction at the origin, $\mathbf{J}(x,t) = \hat{\mathbf{y}}\,\kappa(t)\,\delta(x)$, will generate electromagnetic plane waves propagating in the $+x$ and $-x$ direction, with **E** and **B** fields given by

$$\mathbf{E}(x,t) = -\hat{\mathbf{y}}\frac{c}{2}\mu_o\,\kappa\left(t-\frac{|x|}{c}\right) \tag{25.2.1}$$

$$\mathbf{B}(x,t) = -\hat{\mathbf{z}}\frac{1}{2}\mu_o\,\kappa\left(t-\frac{|x|}{c}\right) sign(x) \tag{25.2.2}$$

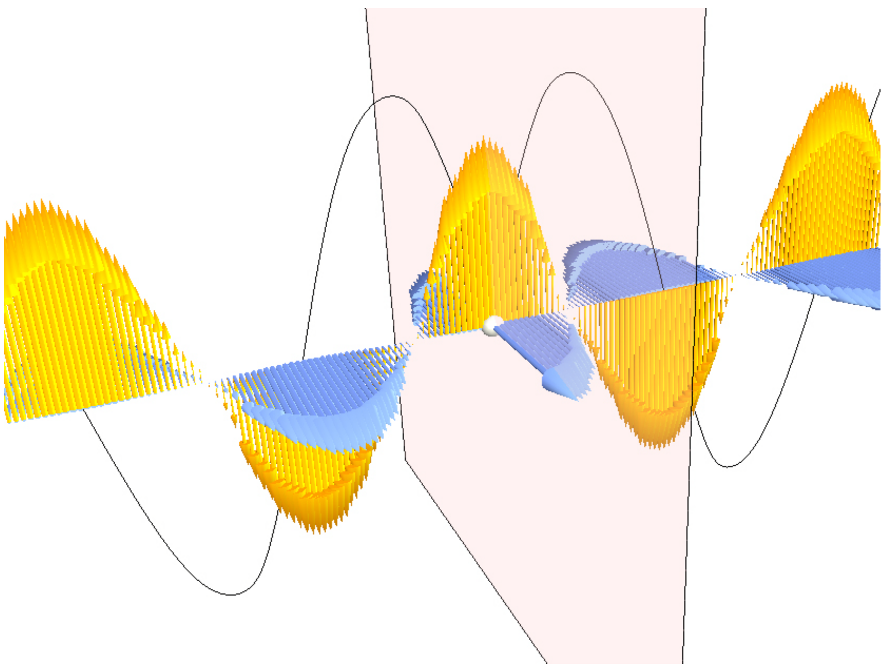


Figure 25-1: The E and B fields of an oscillating current sheet at *x = 0*

We will use these expressions below to calculate the electromagnetic wave generated by a thin dielectric sheet oscillating up and down.

25.3 The slow-down of electromagnetic waves traversing a thin dielectric sheet

25.3.1 The "polarization" current density

Suppose we add matter consisting of many point dipoles. For example, we could have a number of massive ions which we regard as fixed in space, with number density *n*, and with each of the ions having a movable electron attached to it by a spring with spring constant *k*. Suppose the vector separation between the electron and the ion for the i^{th} ion/electron pair is $\Delta\mathbf{r}_e$, and the electron sees an electric field **E** and magnetic field **B** at its location. Then the non-relativistic equation of motion for the electron is

$$m_e \frac{d}{dt}\mathbf{u}_e = -e\left(\mathbf{E} + \mathbf{u}_e \times \mathbf{B}\right) - k\Delta\mathbf{r}_e \tag{25.3.1}$$

We will assume that that $cB \le E$ and that the speed of the electron is much less than c, in which case we can neglect the $\mathbf{u}_e \times \mathbf{B}$ compared to the **E** term in (25.3.1), so that we have

$$m_e \frac{d}{dt}\mathbf{u}_e = -e\mathbf{E} - k\Delta\mathbf{r}_e \tag{25.3.2}$$

or

$$\frac{d}{dt}\mathbf{u}_e = -\frac{e}{m_e}\mathbf{E} - \omega_o^2\Delta\mathbf{r}_i \qquad \omega_o^2 = \frac{k}{m_e} \tag{25.3.3}$$

If we now assume that the electric field varies as $\cos(\omega t)$ and that $\omega \ll \omega_o = \sqrt{k/m_e}$, we can neglect the $\frac{d}{dt}\mathbf{u}_e$ term in (25.3.3) compared to the $\omega_o^2\Delta\mathbf{r}_i$ term, giving us

$$0 = -\frac{e}{m_e}\mathbf{E} - \omega_o^2\Delta\mathbf{r}_e \qquad \Rightarrow \quad \Delta\mathbf{r}_e\left(t\right) = -\frac{e}{\omega_o^2 m_e}\mathbf{E}\left(t\right) \tag{25.3.4}$$

What (25.3.4) tells us is that if the electron sees an electric field varying with time at its position, and if the frequency of the electric field variation is low compared to the natural oscillation frequency of the electron, ω_o, then the position of the electron will vary with the electric field **E**, and its velocity will vary as

$$\mathrm{u}_e(\mathrm{t}) = \frac{d}{dt}\Delta\mathbf{r}_e\left(t\right) = -\frac{e}{\omega_o^2 m_e}\frac{d}{dt}\mathbf{E}\left(t\right) \tag{25.3.5}$$

This means that a time varying electric field in a dielectric will induce a current due to the motion of the "bound" charges, and this is called the "polarization current". We can compute this current using (25.3.5) and the definition of current density to get

$$\mathbf{J}_{polarization}\left(t\right) = -ne\mathrm{u}_e(t) = +\frac{ne^2}{\omega_o^2 m_e}\frac{d}{dt}\mathbf{E}\left(t\right) \tag{25.3.6}$$

It is clear that if we have an electric field that varies in space and time, then (25.3.6) will generalize to

$$\mathbf{J}_{polarization}\left(\mathbf{r},t\right) = -ne\mathrm{u}_e(t) = +\frac{ne^2}{\omega_o^2 m_e}\frac{\partial}{\partial t}\mathbf{E}\left(\mathbf{r},t\right) \tag{25.3.7}$$

We define

$$\chi_e = \frac{n e^2}{\varepsilon_o \omega_o^2 m_e} \tag{25.3.8}$$

and with this definition (25.3.7) becomes

$$\mathbf{J}_{polarization}(\mathbf{r},t) = \varepsilon_o \chi_e \frac{\partial}{\partial t}\mathbf{E}(\mathbf{r},t) \tag{25.3.9}$$

Thus we have come up with a physical model which explains why we would have a relation like (24.6.1).

25.3.2 Change in the speed in the dielectric due to the polarization current

We can easily see that the polarization current given in (25.3.9) will change the speed of electromagnetic waves in the dielectric. We saw above that if we only have a dielectric, and $\mu = \mu_o$, the velocity of propagation of a plane electromagnetic wave is ***NOT*** the speed of light *c*, but

$$\frac{\omega}{k} = \frac{1}{\sqrt{\mu_o \varepsilon_o (1+\chi_e)}} = \frac{c}{\sqrt{(1+\chi_e)}} = \frac{c}{\sqrt{K_e}} \tag{25.3.10}$$

where we have defined the dielectric constant K_e such that

$$K_e = 1 + \chi_e \tag{25.3.11}$$

So the addition of our bound charges and their associated polarization currents in the presence of time changing electric fields leads to a slowing of the propagation speed of electromagnet waves. How can this be?

That is, how does the presence of a dielectric slow down an electromagnetic wave propagating through it? What happens is that the time-varying electric field of the incoming wave drives an oscillating current in the dielectric. These oscillating current sheets, of necessity, must generate electromagnetic waves. The new waves are out of phase with the incident wave, and as a result of the interference between these two waves, the phase of the combination differs from the phase of the incident light. We observe this change in phase as a change in speed. Let us make this qualitative description quantitative[7].

---

[7] This treatment is suggested by a similar approach found in Chapter 31 of Feynman, Leighton, and Sands, *The Feynman Lectures on Physics, Vol 1*, Addison-Wesley, 1963.

25.3.3 The effects of a thin dielectric slab

Suppose we have totally empty space except for a thin dielectric slab lying in the *y-z* plane at *x = 0*. A plane electromagnetic wave whose electric field is polarized in the *y*-direction propagates in the +*x*-direction and encounters the thin dielectric slab of width *D* with dielectric constant K_e . We assume that this is very close to one, that so that the speed of light in the slab is only slightly less than *c*. If there we no dielectric present, the wave would propagate through the slab in a time *D/c*. We now show that the additional time Δt it takes to get through the slab because the thin dielectric slab "slows" down the light can be written as $\Delta t = D\chi_e / 2c$, which corresponds to the slow-down we expect because the speed of the wave in the slab is no longer *c* but $c/\sqrt{1+\chi_e} \approx c - c\chi_e/2$, so that the time to "get through" the slab is no longer D/c but $\dfrac{D}{c(1-\chi_e/2)} \approx \dfrac{D}{c} + \dfrac{D\chi_e}{2c}$.

The electric field of our incident plane wave is given by

$$\delta \mathbf{E}_K = \hat{\mathbf{y}}\,\delta E_o \cos\omega\left(t - \frac{x}{c}\right) \tag{25.3.12}$$

as shown in the top wave form in Figure 25-2. We put at the origin a dielectric sheet with dielectric constant $K_e = 1 + \chi_e$ and width *D*, which we assume is small compared to a wavelength of our incoming wave, that is $\lambda = 2\pi c/\omega >> D$. Because of the presence of the bound charges in the slab, which will oscillate up and down under the influence of the electric field of the incident wave, we will see a polarization current density given by (see (25.3.9) and (25.3.12))

$$\mathbf{J}_{bound} = \varepsilon_o \chi_e \frac{\partial}{\partial t}\mathbf{E} = -\varepsilon_o \chi_e \delta E_o \omega \sin(\omega t)\,\hat{\mathbf{y}} \tag{25.3.13}$$

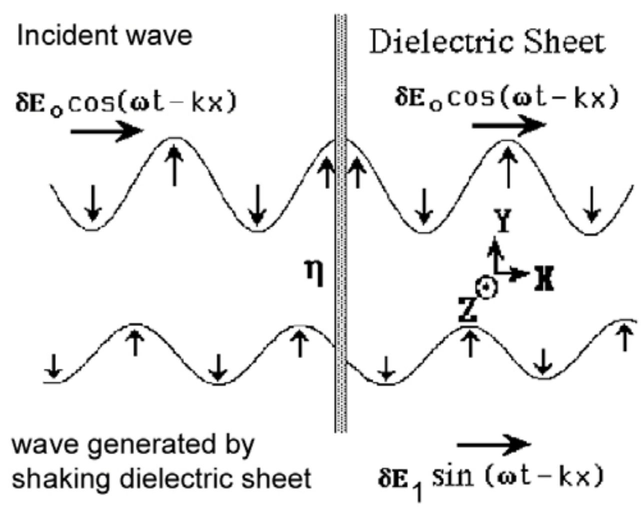


Figure 25-2: An electromagnetic wave encoungers a thin dielectric sheet

Thus the electric field of our incoming wave sets up an oscillating sheet of current with current per unit length in the *y*-direction of

$$\kappa(t) = -\varepsilon_o \chi_e \delta E_o D\omega \sin\left(\omega t\right) \tag{25.3.14}$$

Using equation (25.2.1) above, for $x > 0$, we see that the electric field of the wave *generated by this current sheet* is given by

$$\delta\mathbf{E}_1 = -\hat{\mathbf{y}}\frac{c}{2}\mu_o\,\kappa(t-\frac{x}{c}) = \hat{\mathbf{y}}\left[\frac{\omega D\chi_e}{2c}\right]\delta E_o \sin\omega\left(t-\frac{x}{c}\right) \quad \textit{for } x > 0 \tag{25.3.15}$$

This wave is shown by the lower wave form in Figure 25-2 as a function of *x* for a given instant of time $t = 0$. The total electric field for $x > 0$ is thus

$$\delta\mathbf{E}_{total} = \delta\mathbf{E}_0 + \delta\mathbf{E}_1 = \hat{\mathbf{y}}\,\delta E_o\left\{\cos\omega\left(t-\frac{x}{c}\right) + \frac{\omega D\chi_e}{2c}\sin\omega\left(t-\frac{x}{c}\right)\right\} \quad \textit{for } x > 0 \tag{25.3.16}$$

To see that (25.3.16) implies that the wave is delayed in going through the dielectric, we assume that

$$\frac{\omega D\chi_e}{2c} = \frac{\pi D\chi_e}{\lambda} << 1 \tag{25.3.17}$$

where λ is the wavelength of the wave. We now use the fact that if $\beta << 1$ then

$$\cos\left(\psi-\beta\right) = \cos\psi\cos\beta + \sin\psi\sin\beta \approx \cos\psi + \beta\sin\psi \tag{25.3.18}$$

And comparing (25.3.18) to (25.3.16) assuming (25.3.17) allows us to write (25.3.16) as

$$\delta\mathbf{E}_{total}(x,t) = \hat{\mathbf{y}}\,\delta E_o \cos\omega\left(t-\Delta t-\frac{x}{c}\right) \quad \textit{for } x > 0 \tag{25.3.19}$$

where $\Delta t = D\chi_e / 2c$. Equation (25.3.19) shows that the peaks in amplitude for x > 0, are delayed by a time Δt from the time we expect if there were no dielectric, and this delay in the peak corresponds the a longer time for the wave to “get through” the dielectric sheet, as we noted above.

Note that we have gotten this delay by adding up two waves in (25.3.16), ***both of which are traveling at the speed of light.***

26 Interactions with matter II

26.1 Learning Objectives

We turn now from insulators to conductors. We first look at the "microscopic" relation between **J** and **E** in conductors, or Ohm's Law, and then move on to consider the consequences of this relationship.

26.2 The microscopic form of Ohm's Law

In constrast to our treatment of the relationship between **J** and **E** for dielectrics in Section 25.3.1 above, we consider here a different model for currents in a material. Again, we suppose we have a number of massive ions which we regard as fixed in space, with number density *n*, with an equal number of electrons, but now with the electrons free to move and not bound to the ions. However, the electrons do "collide" with the ions, and this results in a frictional drag. Suppose the vector displacement of an electron from its rest position is $\Delta\mathbf{r}_e$, and that the electron sees an electric field **E** and magnetic field **B** at its location, with $\frac{d}{dt}\Delta\mathbf{r}_e = \mathbf{u}_e$. Then the non-relativistic equation of motion for the electron is

$$m_e \frac{d}{dt}\mathbf{u}_e = -e\left(\mathbf{E} + \mathbf{u}_e \times \mathbf{B}\right) - m_e\gamma_c\mathbf{u}_e \tag{26.2.1}$$

where γ_c is a collision frequency and the $-m_e\gamma_c\mathbf{u}_e$ term represents the frictional drag. Again, we assume that that $cB \le E$ and that the speed of the electron is much less than c, in which case we can neglect the $\mathbf{u}_e \times \mathbf{B}$ compared to the **E** term in (26.2.1), so that we have

$$m_e \frac{d}{dt}\mathbf{u}_e = -e\mathbf{E} - m_e\gamma_c\mathbf{u}_e \tag{26.2.2}$$

or

$$\frac{d}{dt}\mathbf{u}_e = -\frac{e}{m_e}\mathbf{E} - \gamma_c\mathbf{u}_e \tag{26.2.3}$$

If we now assume that the electric field varies as $\cos(\omega t)$ and that $\omega \ll \gamma_c$, we can neglect the $\frac{d}{dt}\mathbf{u}_e$ term in (26.2.3) compared to the $-\gamma_c\mathbf{u}_e$ term, giving us

$$0 = -\frac{e}{m_e}\mathbf{E} - \gamma_c\mathbf{u}_e \quad \Rightarrow \quad \mathbf{u}_e(t) = -\frac{e}{\gamma_c m_e}\mathbf{E}(t) \tag{26.2.4}$$

What (26.2.4) tells us is that if the electron sees an electric field varying with time at its position, and if the frequency of the electric field variation is low compared to the collision frequency, then the velocity of the electron will vary with the electric field **E** in the manner given in (26.2.4). This means that a time varying electric field in a conductor will induce a current due to the motion of the "free" electrons. We can compute this current using (25.3.5) and the definition of current density to get

$$\mathbf{J}(t) = -ne\mathrm{u}_e(t) = +\frac{n e^2}{\gamma_e m_e}\mathbf{E}(t) \tag{26.2.5}$$

It is clear that if we have an electric field that varies in space and time, then (26.2.5) will generalize to

$$\mathbf{J}(\mathbf{r},t) = -ne\mathrm{u}_e(t) = +\frac{n e^2}{\gamma_e m_e}\mathbf{E}(\mathbf{r},t) \tag{26.2.6}$$

We define the conductivity σ_c to be

$$\sigma_c = \frac{n e^2}{\gamma_e m_e} \tag{26.2.7}$$

and with this definition (26.2.6) becomes

$$\mathbf{J}(\mathbf{r},t) = \sigma_c\,\mathbf{E}(\mathbf{r},t) \tag{26.2.8}$$

26.3 Reflection of an electromagnetic wave by a conducting sheet

26.3.1 The conceptual basis

How does a very good conductor reflect an electromagnetic wave falling on it? In words, what happens is the following. The time-varying electric field of the incoming wave drives an oscillating current on the surface of the conductor, following Ohm's Law. That oscillating current sheet, of necessity, must generate waves propagating in both directions from the sheet. One of these waves is the reflected wave. The other wave cancels out the incoming wave inside the conductor. Let us make this qualitative description quantitative.

Suppose we have an infinite plane wave propagating to the right, generated by currents far to the left and not shown. Suppose that the electric and magnetic fields of this wave are given by

$$\mathbf{E}_{incident}(x,t) = \hat{\mathbf{y}}\,\delta E_o \cos\omega\left(t-\frac{x}{c}\right) \qquad \mathbf{B}_{incident}(x,t) = \hat{\mathbf{z}}\,\delta B_o \cos\omega\left(t-\frac{x}{c}\right) \tag{26.3.1}$$

as shown in the top wave form in the Figure 26-1. We put at the origin ($x = 0$) a conducting sheet with width D, which we assume is small compared to a wavelength of our incoming wave. This conducting sheet will *reflect* our incoming wave. How?

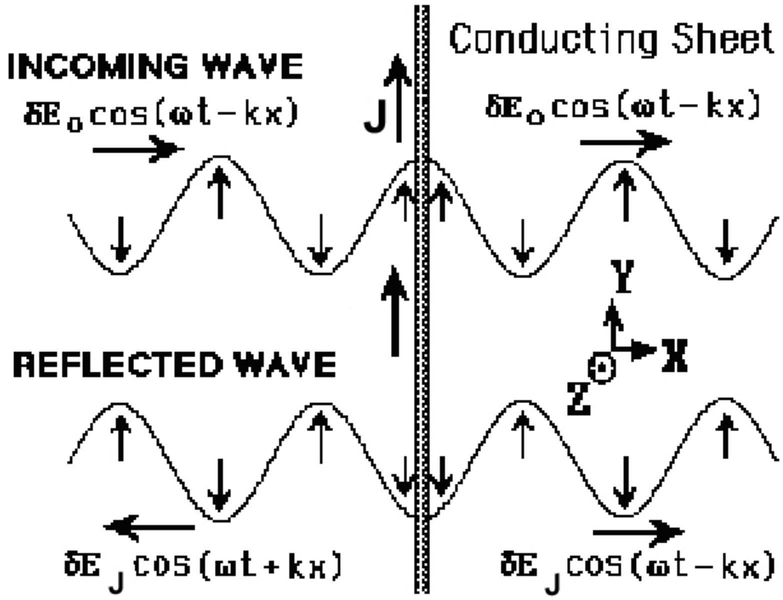


Figure 26-1: An incoming electromagnetic wave reflected by a conducing sheet.

The electric field of the incoming wave will cause a current $\mathbf{J} = \sigma_c \mathbf{E}$ to flow in the sheet, where σ_c is the conductivity. Moreover, the direction of $\mathbf{J}$ *will be in the same direction as the electric field of the incoming wave*, as shown in Figure 26-1. Thus our incoming wave sets up an oscillating sheet of current with current per unit length $\boldsymbol{\kappa} = D\mathbf{J}$. As in our discussion in Section 25.3.3, this current sheet *will generate electromagnetic waves*, moving both to the right and to the left (see Figure 26-1, lower wave form) away from the oscillating sheet of charge. What is the amplitude of these waves?

Using equation (25.2.1) above, for $x > 0$ the electric field of the wave generated by the current **J**, which we denote by $\mathbf{E}_\mathbf{J}(x,t)$ will be

$$\mathbf{E}_\mathbf{J}(x,t) = -\frac{c\mu_o}{2} D\mathbf{J}\left(t - \frac{|x|}{c}\right) \tag{26.3.2}$$

and this represents a wave propagating both to the left and to the right at the speed of light. The sign of this electric field at $x = 0$; it is *down* when the sheet of current is *up*, and vice-versa. Thus, for $x > 0$, the electric field $\mathbf{E}_\mathbf{J}(x,t)$ generated by the current in the sheet will always be *opposite* the direction of the electric field of the incoming wave, and *it will tend to cancel out the incoming wave for $x > 0$*. For a very good conductor, in fact, we show below that DJ will be equal to $2\delta E_o / c\mu_o$, so that for $x > 0$ we will have $\mathbf{E}_\mathbf{J}(x,t) = -\hat{\mathbf{y}}\delta E_o \cos(\omega t - kx)$. That is, for a very good conductor, the electric field of the wave generated by the current *will exactly cancel* the electric field of the incoming wave *for x > 0!*

And that's what a very good conductor does. It supports exactly the amount of current per unit length needed to cancel out the incoming wave for *x > 0* ($2\delta E_o / c\mu_o$, or equivalently $2\delta B_o / \mu_o$). For *x < 0, this same current* generates a "reflected" wave propagating back in the direction from which the original incoming wave came, *with the same amplitude as the original incoming wave*. This is how a very good conductor *totally* reflects electromagnetic waves. Below we show that our current density will in fact approach the value needed to accomplish this in the limit that the conductivity σ_C approaches infinity. We also have you obtain this result in the more standard manner in Problem Set 10.

26.3.2 The quantitative result in the limit of infinite conductivity

We show here that a perfect conductor will perfectly reflect an incident wave. To approach the limit of a perfect conductor, we first consider the finite resistivity case, and then let the conductivity go to infinity. As we pointed out above, the electric field of the incoming wave will, by Ohm's Law, cause a current $\mathbf{J} = \sigma_c \mathbf{E}$ to flow in the sheet, where σ_c is the conductivity. Since the sheet is assumed thin compared to a wavelength, we can assume that the entire sheet sees essentially the same electric field, so that $\mathbf{J}$ will be uniform across the thickness of the sheet, and outside of the sheet we will see fields appropriate to a equivalent surface current $\boldsymbol{\kappa} = D\mathbf{J}$. This current sheet will generate electromagnetic waves, moving both to the right and to the left, away from the oscillating current sheet. The total electric field, $\mathbf{E}_{total}(x,t)$, will be the sum of the incident electric field and the electric field generated by the current sheet. Using equations (26.3.1) and (26.3.2) above, we thus have for the total electric field the following expression:

$$\mathbf{E}_{total}(x,t) = \mathbf{E}_{incident}(x,t) + \mathbf{E}_{\mathbf{J}}(x,t) = \hat{\mathbf{y}}\delta E_o \cos\omega\left(t - \frac{x}{c}\right) - \frac{c\mu_o}{2} D\mathbf{J}\left(t - \frac{|x|}{c}\right) \qquad (26.3.3)$$

We also have a relation between the current density $\mathbf{J}$ and $\mathbf{E}_{total}$ from Ohm's Law, which is

$$\mathbf{J}(t) = \sigma_c \mathbf{E}_{total}(0,t) \qquad (26.3.4)$$

Where $\mathbf{E}_{total}(0,t)$ is the total electric field at the position of the conducting sheet (which remember is very thin compared to a wavelength of the wave). Note that is appropriate to use the ***total*** electric field in Ohm's Law--the currents arise from the total electric field, irrespective of the origin of that field. So we have

$$\boldsymbol{\kappa}(t) = D\mathbf{J}(t) = D\sigma_c \mathbf{E}_{total}(0,t) \qquad (26.3.5)$$

If we look at (26.3.3) at $x = 0$, we have

$$\mathbf{E}_{total}(0,t) = \hat{\mathbf{y}}\delta E_o \cos\omega\left(t\right) - \frac{c\mu_o}{2} D\mathbf{J}\left(t\right) \qquad (26.3.6)$$

or using (26.3.5)

$$\mathbf{E}_{total}(0,t) = \hat{\mathbf{y}}\delta E_o \cos\omega\left(t\right) - \frac{c\mu_o}{2} D\sigma_c \mathbf{E}_{total}(0,t) \qquad (26.3.7)$$

We can now solve equation (26.3.7) for $\mathbf{E}_{total}(0,t)$, with the result that

$$\mathbf{E}_{total}(0,t) = \left[\frac{1}{1+\frac{c\mu_o D\sigma_c}{2}}\right]\hat{\mathbf{y}}\delta E_o \cos\omega\left(t\right) \qquad (26.3.8)$$

and therefore, using equation (26.3.8) and (26.3.5)

$$\boldsymbol{\kappa}(t) = D\sigma_c \mathbf{E}_{total}(0,t) = \left[\frac{D\sigma_c}{1+\frac{c\mu_o D\sigma_c}{2}}\right]\hat{\mathbf{y}}\delta E_o \cos\omega\left(t\right) \qquad (26.3.9)$$

If we take the limit that σ_c approaches infinity (no resistance, that is, a perfect conductor), then we can easily see using equation (26.3.8) that $\mathbf{E}_{total}(0,t)$ goes to zero, and that using equation (26.3.9)

$$\boldsymbol{\kappa}(t) = \hat{\mathbf{y}}\frac{2\delta E_o}{c\mu_o}\cos\omega\left(t\right) = \hat{\mathbf{y}}\frac{2\delta B_o}{\mu_o}\cos\omega\left(t\right) \qquad (26.3.10)$$

In this same limit equation (26.3.3) becomes

$$\mathbf{E}_{total}(x,t) = \hat{\mathbf{y}}\delta E_o\left[\cos\omega\left(t-\frac{x}{c}\right) - \cos\omega\left(t-\frac{|x|}{c}\right)\right] \qquad (26.3.11)$$

or

$$\mathbf{E}_{total}(x,t) = \begin{cases} 0 \quad \text{for } x > 0 \\ \hat{\mathbf{y}}\delta E_o\left[\cos\omega\left(t-\frac{x}{c}\right) - \cos\omega\left(t+\frac{x}{c}\right)\right] \quad \text{for } x < 0 \end{cases} \qquad (26.3.12)$$

Again in the same limit of infinite conductivity, our total magnetic fields become

$$\mathbf{B}_{total}(x,t) = \begin{cases} 0 \quad \text{for } x > 0 \\ \hat{\mathbf{z}}\delta B_o\left[\cos\omega\left(t - \dfrac{x}{c}\right) + \cos\omega\left(t + \dfrac{x}{c}\right)\right] \quad \text{for } x < 0 \end{cases} \tag{26.3.13}$$

Thus, we see that we get ***no electromagnetic wave for x > 0***, and standing electromagnetic waves for $x < 0$. Note that right at $x = 0$, the total electric field vanishes. The current per unit length $\boldsymbol{\kappa}(t) = \hat{\mathbf{y}}\dfrac{2\delta B_o}{\mu_o}\cos\omega(t)$ is just the current per length we need to bring the magnetic field down from its value at $x < 0$ to zero for $x > 0$

You may be perturbed by the fact that in the limit of a perfect conductor, the electric field vanishes at $x = 0$, since it is the electric field at $x = 0$ that is driving the current there! But this is ok in this limit, since the conductivity is going to infinity in the same limit. In the limit of very small resistance, the electric field required to drive any current you want can go to zero, because the product of the infinite conductivity and the zero electric field can assume any value you need. That is, even a very small value of the electric field can generate a perfectly finite value of the current.

26.4 Radiation pressure on a perfectly conducting sheet

In the process of the reflection, there is a force per unit area exerted on the perfect conductor. This is just the $\mathbf{J}\times\mathbf{B}$ force due to the current flowing in the presence of the magnetic field of the incoming wave. If we calculate the total force $d\mathbf{F}$ acting on a cylindrical volume with area dA and length D of the conductor, using (26.3.10) and (26.3.13) we find that it is in the $+x$ direction, as follows

$$\begin{aligned} d\mathbf{F}(t) &= dA\,D\mathbf{J}\times\mathbf{B}_{incident}(0,t) \\ &= dA\boldsymbol{\kappa}(t)\times\mathbf{B}_{incident}(0,t) \\ &= \hat{\mathbf{x}}\,dA\left[\frac{2\delta B_o^2}{\mu_o}\right]\cos^2(\omega t) \\ &= \hat{\mathbf{x}}dA\left(\frac{2\delta E_o\delta B_o}{c\mu_o}\right)\cos^2(\omega t) \end{aligned} \tag{26.4.1}$$

so that the force per unit area, $d\mathbf{F}/dA$, or the radiation pressure, is just twice the Poynting flux divided by the speed of light c. Note that in (26.4.1) we use on the magnetic field due to the incident wave. Including the magnetic field due to the generated wave is unnecessary, since that wave reverses sign across $x = 0$ and thus generates net force on the sheet. The factor of two is appropriate for total reflection, as is the case here. For total absorption, the factor of two becomes unity. We can obtain this same expression by integrating the Maxwell stress tensor over surface of the same cylindrical volume.

27 Moving Magnets, Einstein, and Faraday

27.1 Learning Objectives

We first review the example that Einstein gave in the first paragraph of his 1905 paper on special relativity, and try to explain what motivated him to focus on this phenomena in one of his most famous papers. This involves thinking about the magnetic field of a moving magnetic and what electric field is associated with it, and conversely the force that will be seen by a conductor moving in the magnetic field of a stationary magnet.

27.2 What did Einstein mean?

On the Electrodynamics of Moving Bodies
by A. Einstein
(Translated from "Zur Elektrodynamik bewegter Korper, " Annalen der Physik, 17, 1905)

It is known that Maxwell's electrodynamics- as usually understood at the present time- when applied to moving bodies, leads to asymmetries which do not appear to be inherent in the phenomena. Take, for example, the reciprocal electrodynamic action of a magnet and a conductor. The observable phenomena here depends only on the relative motion of the conductor and the magnet, whereas the customary view draws a sharp distinction between the two cases in which either the one or the other of these bodies is in motion. For if the magnet is in motion and the conductor at rest, there arises in the neighborhood of the magnet an electric field with a certain definite energy, producing a current at the places where parts of the conductor are situated. But, if the magnet is stationary and the conductor in motion, no electric field arises in the neighborhood of the magnet. In the conductor, however, we find an electromotive force, to which in itself there is no corresponding energy, but which gives rise- assuming equality of relative motion in the two cases discussed- to electric currents of the same path and intensity as those produced by the electric forces in the former case.

Examples of this sort, together with the unsuccessful attempts to discover any motion of the earth relatively to the "light medium," suggest that the phenomena of electrodynamics as well as of mechanics possess no properties corresponding to the idea of absolute rest....We will raise this conjecture (the purport of which will hereafter be called the "Principle of Relativity") to the status of a postulate, and also introduce another postulate, which is only apparently irreconcilable with the former, namely, that light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body.....

Let us try to understand in detail what Einstein meant by the example he gives above about moving magnets versus stationary magnets. First, he talks about the electric field of magnets in motion. Let us investigate the electric fields of moving magnets.

27.2.1 The electric field of a magnet moving at constant velocity

Suppose we have a magnet whose dipole moment is along the *x*-direction moving in the *x*-direction at constant velocity $\mathbf{V} = V\,\hat{\mathbf{x}}$ as seen in the laboratory frame. Suppose **E** and **B** are the fields of the magnet as seen in the laboratory frame. If we look at (14.3.8) for the transformation of fields between co-moving frames, we have

$$\overline{E}_x = E_x \qquad \overline{E}_y = \gamma\left(E_y - V\,B_z\right) \qquad \overline{E}_z = \gamma\left(E_z + VB_y\right) \tag{27.2.1}$$

where the "barred" frame is the rest frame of the magnet. If we neglect terms of order $\left(\mathrm{V}/c\right)^2$ equation (27.2.1) can be written as

$$\overline{\mathbf{E}} = \mathbf{E} + \mathbf{V}\times\mathbf{B} \tag{27.2.2}$$

But in the rest frame of the magnet, the electric field $\overline{\mathbf{E}}$ is zero, so that we must have the following relationship between the electric and magnetic fields of the magnet as seen in the laboratory frame, where the magnet is moving:

$$\mathbf{E} = -\mathbf{V}\times\mathbf{B} \tag{27.2.3}$$

Now how does the magnetic field in the laboratory frame relate to the dipolar magnetic field we have in the rest frame of the magnet? If we look at the transformations (14.3.15) and consider them when we are going from the rest frame of the magnet to the laboratory frame (this reverses the sign of the velocity and changes a bar to an unbar and vice versa), we have

$$B_x = \overline{B}_x \qquad B_y = \gamma\left(\overline{B}_y - \frac{\mathrm{V}}{c^2}\overline{E}_z\right) \qquad B_z = \gamma(\overline{B}_z + \frac{\mathrm{V}}{c^2}\overline{E}_y) \tag{27.2.4}$$

Since we know that the electric field $\overline{\mathbf{E}}$ in the rest frame of the magnet is zero, this means that to first order in $\left(\mathrm{V}/c\right)$, we have that the magnetic field in the laboratory frame is the same as the magnetic field in the rest frame of the magnet, which is dipolar. Therefore we have from this fact and (27.2.3)

$$\mathbf{E} = -\mathbf{V}\times\mathbf{B}_{dipole} \tag{27.2.5}$$

In Figure 27-1 we show this "motional electric field" for the case where the magnet moves to the right. In Figure 27-2 we show the electric field where the magnet moves to the left. In both cases the electric field moves in circles about the direction of motion of the magnet, with the sense of the circulation of the electric field reversing from in front of the magnet to behind the magnet.

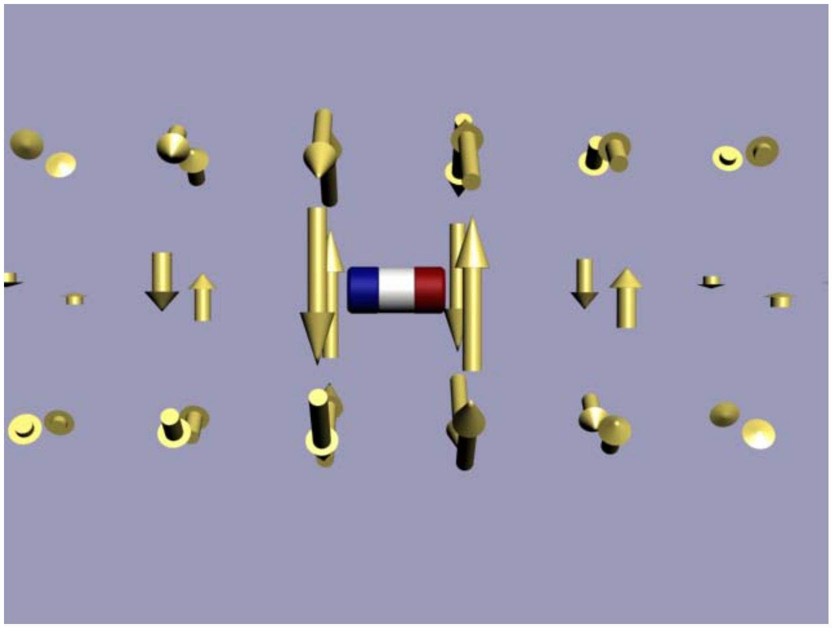


Figure 27-1: The E field of a magnet moving to the right (red is the north pole)

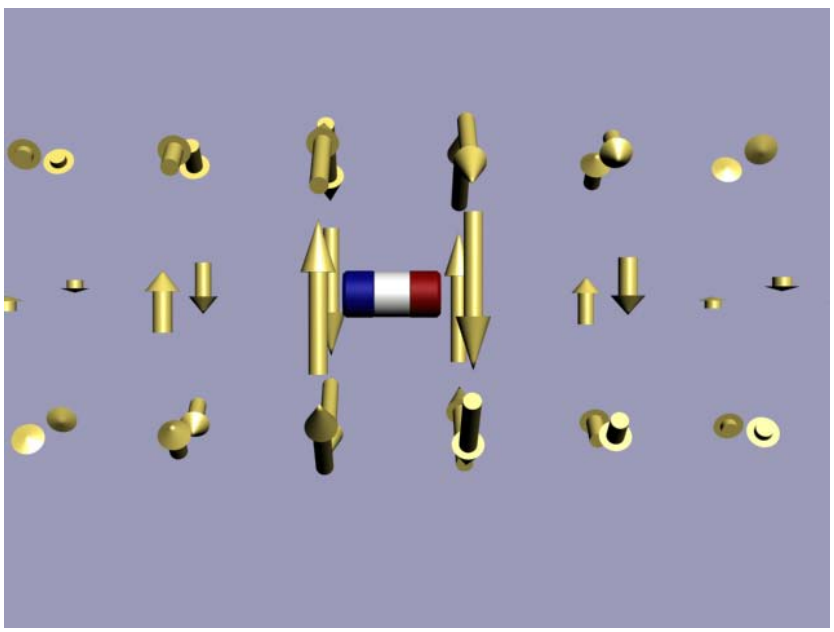


Figure 27-2: The E field of a magnetic moving to the left (red is the north pole)

27.2.2 "… the magnet is in motion and the conductor at rest…"

Now let us return to Einstein's example above. He first considers the situation where the magnet is in motion and the conductor is at rest. If we actually do this experiment, as in Figure 27-3, which shows one frame of a movie of the experiment[8], we find that the current in the coil is left-handed with respect to the magnetic dipole vector (pointing to the right in the figure) when the magnet is moving toward the coil, and left-handed with respect to the magnetic dipole vector when the magnet is being pulled away from the coil. This is in keeping with the direction of the electric field shown in Figure 27-1 and in Figure 27-2, and this is what Einstein means when he says that "…there arises in the neighborhood of the [moving] magnet an electric field with a certain definite energy, producing a current at the places where parts of the conductor are situated…". That is, there is a current in the loop because of the electric field of the moving magnet.

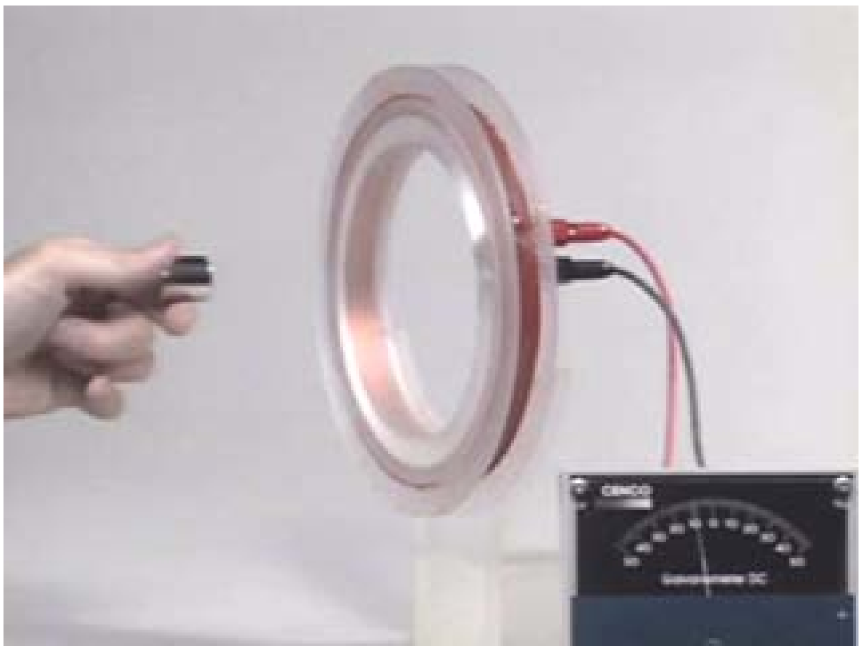


Figure 27-3: A magnet moving toward and away from a stationary loop of wire.

27.2.3 "…the magnet is stationary and the conductor in motion…"

On the other hand, suppose the magnet is at rest and the conductor is moving. Then there is no electric field as seen in the conductor. But there is a $\mathbf{V}\times\mathbf{B}$ force on the charges in the coil, because they are now moving along with the coil, and a little thought shows that this force is in the same direction and has the same magnitude as the electric field given in (27.2.5) (remember we had the magnet moving toward the coil at velocity $\mathbf{V} = V\,\hat{\mathbf{x}}$, so this means that the coil is moving toward the magnet at velocity $-\mathbf{V}$). Thus in either case we get the same current in the coil, but in one case the observer would say that the charges in the coil feel a force producing a current because of the electric field they see, and in the other an observer would say that the charges feel a force producing a current because of the $\mathbf{V}\times\mathbf{B}$ force they see. Regardless of what the source of the force

[8] http://web.mit.edu/viz/EM/visualizations/faraday/faradaysLaw/

is, we see a current as a result. Let us look at the mathematical form of Faraday's Law and put these qualitative ideas into quantitative form.

27.3 Differential and Integral Forms of Faraday's Law

27.3.1 Faraday's Law in Differential Form

Faraday's Law in differential form is

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{27.3.1}$$

We can begin the process of writing Faraday's Law in integral form by integrating both sides of (27.3.1) over any open surface *S(t)* and converting the integral of the curl of **E** to a line integral of **E** over the bounding contour *C(t)* using Stokes Theorem, giving

$$\int_{S(t)} \left(\nabla \times \mathbf{E}\right) \cdot \hat{\mathbf{n}} da = \oint_{C(t)} \mathbf{E} \cdot d\mathbf{l} - \int_{S(t)} \frac{\partial \mathbf{B}}{\partial t} \cdot \hat{\mathbf{n}} da \tag{27.3.2}$$

We would like to move the $\frac{\partial}{\partial t}$ term out from under the integral sign to become a $\frac{d}{dt}$ in front of the integral sign, but we have to be careful about doing this because we frequently apply Faraday's Law in integral form to moving circuits. We have to pause for a bit to prove the following mathematical theorem.

27.3.2 A Mathematical Theorem

The theorem we now prove has nothing specifically to do with electromagnetism, it is a general theorem about the flux through moving open surfaces. Consider the following problem. You are given a vector field $\mathbf{F}(\mathbf{r}, t)$ which is a function of space and time. You are also given an open surface *S* with associated bounding contour *C*, and this surface is moving in space, with each element of the surface moving at some specified velocity $\mathbf{v}(\mathbf{r}, t)$ (the **r** here of course must lie on the open surface). All of these things are given. Suppose you now compute the flux of **F** through *S* at any given instant of time *t*, in the usual way, that is

$$\Phi_{\mathrm{F}}(t) = \int_{\mathrm{S(t)}} \mathbf{F}(\mathbf{r}, t) \cdot \hat{\mathbf{n}}\, da \tag{27.3.3}$$

where we have used the notation *S(t)* to indicate that the surface is changing in time, as well as **F**. Note that $\Phi_{\mathrm{F}}(t)$ depends only on time--we have integrated over space. Now, here is the question we want to answer. What is the time derivative of $\Phi_{\mathrm{F}}(t)$? Let's start out by giving the answer, and then we will show how it comes about.

$$\frac{d\Phi_{\mathrm{F}}(t)}{dt} = \int_{\mathrm{S(t)}} \frac{\partial \mathbf{F}(\mathbf{r},t)}{\partial t} \cdot \hat{\mathbf{n}}\, da - \oint_{C(t)} \left[\mathbf{v} \times \mathbf{F}(\mathbf{r},t)\right] \cdot \mathbf{dl} + \int_{S(t)} \left[\nabla \cdot \mathbf{F}(\mathbf{r},t)\right] \mathbf{v} \cdot \hat{\mathbf{n}}\, da \qquad (27.3.4)$$

The first term on the right of this equation is due to the intrinsic time variability of **F**. The other two terms on the right, a line integral around the bounding contour *C(t)*, and another surface integral over *S(t)*, arise purely because of the motion of *S* (that is, they disappear when $\mathbf{v}(\mathbf{r},t)$ is zero everywhere).

To prove (27.3.4), let us start out with the definition of the derivative of a function:

$$\frac{d\Phi_{\mathrm{F}}(t)}{dt} = \lim_{\Delta t \to 0} \frac{\Phi_{\mathrm{F}}(t+\Delta t) - \Phi_{\mathrm{F}}(t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[\int_{\mathrm{S}(t+\Delta t)} \mathbf{F}(\mathbf{r},t+\Delta t) \cdot \hat{\mathbf{n}}\, da - \int_{\mathrm{S(t)}} \mathbf{F}(\mathbf{r},t) \cdot \hat{\mathbf{n}}\, da\right] \qquad (27.3.5)$$

We use a Taylor series expansion of **F** about *t*,

$$\mathbf{F}(\mathbf{r},t+\Delta t) = \mathbf{F}(\mathbf{r},t) + \Delta t \frac{\partial \mathbf{F}(\mathbf{r},t)}{\partial t} + ... \qquad (27.3.6)$$

to write the first term on the right of equation (27.3.5) as two terms, namely

$$\frac{d\Phi_{\mathrm{F}}(t)}{dt} = \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[\Delta t \int_{\mathrm{S}(t+\Delta t)} \frac{\partial \mathbf{F}(\mathbf{r},t)}{\partial t} \cdot \hat{\mathbf{n}}\, da + \int_{\mathrm{S}(t+\Delta t)} \mathbf{F}(\mathbf{r},t) \cdot \hat{\mathbf{n}}\, da - \int_{\mathrm{S(t)}} \mathbf{F}(\mathbf{r},t) \cdot \hat{\mathbf{n}}\, da\right] \qquad (27.3.7)$$

Now the next part is tricky. The divergence theorem is just as good for time varying functions as for functions which do not vary in time, and it says the following. Pick any closed volume *V′* bounded by the closed surface *S′*. ***This is different from the surface S above*** (for one thing, *S′* is closed, and *S* is open). Now, the divergence theorem states that at any time *t*, we have

$$\int_{V'} \left[\nabla \cdot \mathbf{F}(\mathbf{r}',t)\right] d^3x' = \oint_{S'} \mathbf{F}(\mathbf{r},t) \cdot \hat{\mathbf{n}}'\, da' \qquad (27.3.8)$$

This equation is true for any volume *V′* and any given instant of time, and we are going to apply it at time *t* to the following volume (brace yourself). The volume *V′* at time *t* which we are going to use in (27.3.8) is *the volume swept out by our original open surface S, as it moves through space between time t and time t+Δt.*

This at first seems like a peculiar volume to use at time *t*, since it depends on things that happen in the future, but it is a perfectly well defined volume at time *t*, we just need to know what is going to happen with *S(t)* in the future to define this volume of space at *t*, but that is given. Figure 27-4 shows this closed surface and the associated volume.

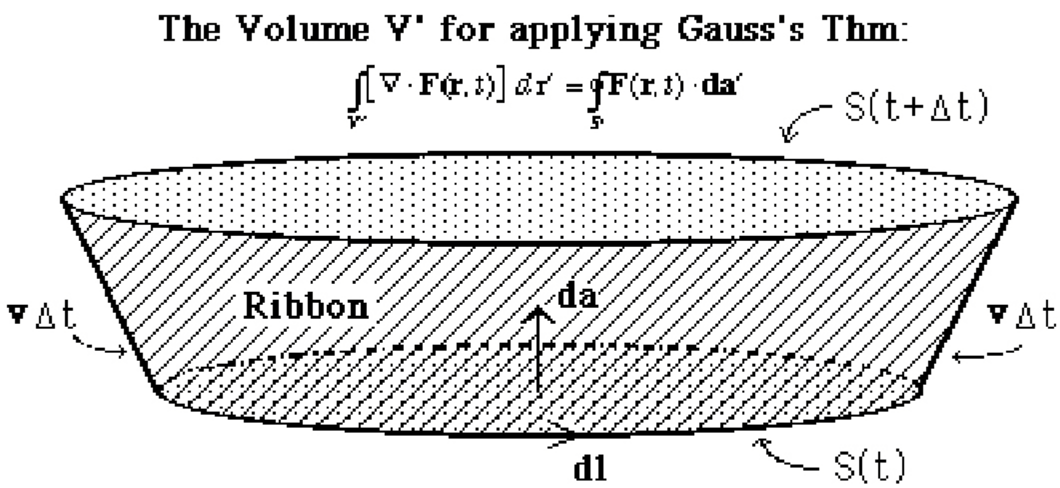


Figure 27-4: The closed surface at time *t* we use for applying the divergence theorem

The infinitesimal line element **dl** and the infinitesimal area element $\mathbf{da} = \hat{\mathbf{n}}\, da$ shown in Figure 27-4 are associated with the open surface *S*. They must be right-handed with respect to one another, which is why **da** must be up if **dl** is counterclockwise. The vector $\mathbf{da} = \hat{\mathbf{n}}\, da$ is *not* the infinitesimal area element **da'** associated with the closed surface *S'*--that vector must always point away from the volume of interest, namely *V'*. So **da'** is anti-parallel to **da** on the bottom of the volume shown above, and parallel to **da** on the top of the volume shown above.

Applying Gauss's Theorem to the volume *V'* gives

$$\int_{V'} [\nabla \cdot \mathbf{F}(\mathbf{r},t)]\, d^3x' = \oint_{S'} \mathbf{F}(\mathbf{r},t) \cdot \mathbf{da}' = \oint_{S(t+\Delta t)} \mathbf{F}(\mathbf{r},t) \cdot \mathbf{da}' + \oint_{S(t)} \mathbf{F}(\mathbf{r},t) \cdot \mathbf{da}' + \oint_{\text{ribbon}} \mathbf{F}(\mathbf{r},t) \cdot \mathbf{da}' \tag{27.3.9}$$

where we have gone all around the closed surface *S'* containing *V'*, including the "ribbon" of area that is swept out by the moving contour *C(t)* between times *t* and *t+Δt.* If we use the fact that $\mathbf{da}' = \mathbf{da}$ on the surface *S(t+Δt)* and $\mathbf{da}' = -\mathbf{da}$ on the surface *S(t)*, we have

$$\int_{V'} [\nabla \cdot \mathbf{F}(\mathbf{r},t)]\, d^3x' = \oint_{S'} \mathbf{F}(\mathbf{r},t) \cdot \mathbf{da}' = \oint_{S(t+\Delta t)} \mathbf{F}(\mathbf{r},t) \cdot \mathbf{da} - \oint_{S(t)} \mathbf{F}(\mathbf{r},t) \cdot \mathbf{da} + \oint_{\text{ribbon}} \mathbf{F}(\mathbf{r},t) \cdot \mathbf{da}' \tag{27.3.10}$$

Now what is d^3x' for this volume *V'*? It is pretty easy to see that $d^3x' = \mathbf{v} \cdot \mathbf{da}\, \Delta t$. This is dimensionally correct, and has the right behavior (if **v** is perpendicular to **da** at some point on *S* there is no increase in differential volume d^3x' at that point between time *t* and $t+\Delta t$). In addition, it is also pretty obvious that the infinitesimal area element on the surface of the "ribbon" is given by $\mathbf{da}'_{\text{ribbon}} = \mathbf{dl} \times \mathbf{v}\,\Delta t$. To see this clearly, consider Figure 27-5**Error! Reference source not found.**, which is a blowup of an area of detail of Figure 27-4 . This sketch shows the geometry of the situation. This form $\mathbf{da}'_{\text{ribbon}} = \mathbf{dl} \times \mathbf{v}\,\Delta t$ has the right dependence--if **v** is parallel to **dl**, at some point, the contour *C* is moving parallel to itself at that point, and there is no area swept out by *C* at that point between *t* and $t+\Delta t$.

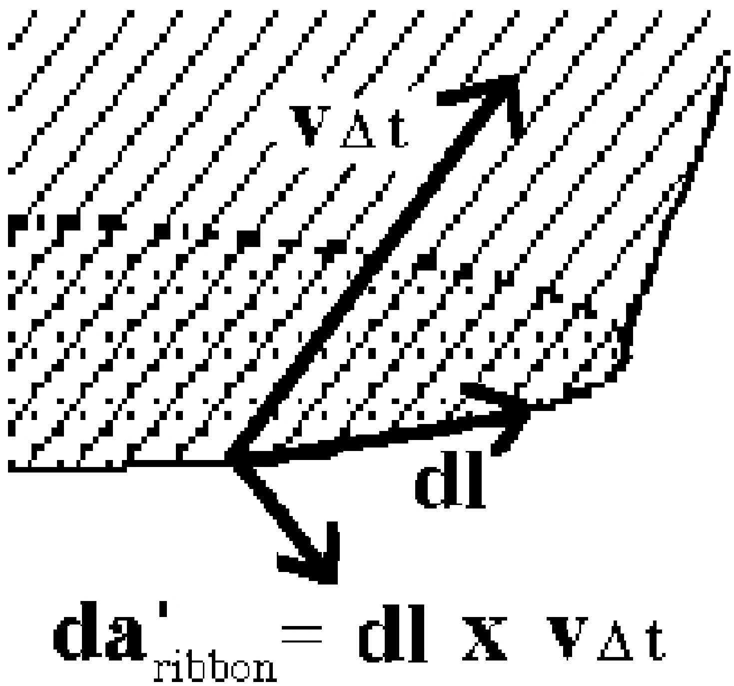


Figure 27-5: The infinitesimal area element for the ribbon

So, using these two forms for d^3x' and $\mathbf{da}'_{\text{ribbon}}$ in equation (27.3.10), we can convert the volume integral on the left hand side into a surface integral over *S(t)*, and the area integral over the ribbon on the right hand side to a line integral over *C(t)*. We certainly make some error in doing this, but the corrections will be of order Δt, and we already have a Δt in these terms, so that the corrections will vanish as we go to the limit of $\Delta t = 0$. Thus, we have that

$$\int_{S(t)} \left[\nabla\cdot\mathbf{F}(\mathbf{r},t)\right]\,\mathbf{v}\cdot\mathbf{da}\;\Delta t = \oint_{S(t+\Delta t)} \mathbf{F}(\mathbf{r},t)\cdot\mathbf{da} - \oint_{S(t)} \mathbf{F}(\mathbf{r},t)\cdot\mathbf{da} + \oint_{C(t)} \mathbf{F}(\mathbf{r},t)\cdot\left(\mathbf{dl}\ \mathbf{x}\ \mathbf{v}\right)\Delta t \tag{27.3.11}$$

which we can rewrite as

$$\oint_{S(t+\Delta t)} \mathbf{F}(\mathbf{r},t)\cdot\mathbf{da} - \oint_{S(t)} \mathbf{F}(\mathbf{r},t)\cdot\mathbf{da} = -\Delta t \oint_{C(t)} \mathbf{F}(\mathbf{r},t)\cdot\left(\mathbf{dl}\times\mathbf{v}\right) + \Delta t \int_{S(t)} \left[\nabla\cdot\mathbf{F}(\mathbf{r},t)\right]\mathbf{v}\cdot\mathbf{da} \tag{27.3.12}$$

Thus we can write (27.3.7) as

$$\frac{d\Phi_{\mathrm{F}}(t)}{dt} = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left[\Delta t \int_{S(t+\Delta t)} \frac{\partial \mathbf{F}(\mathbf{r},t)}{\partial t}\cdot\mathbf{da} - \Delta t \oint_{C(t)} \mathbf{F}(\mathbf{r},t)\cdot\left(\mathbf{dl}\times\mathbf{v}\right) + \Delta t \int_{S(t)} \left[\nabla\cdot\mathbf{F}(\mathbf{r},t)\right]\mathbf{v}\cdot\mathbf{da}\right] \tag{27.3.13}$$

If we take the limit, we have

$$\frac{d\Phi_{\mathrm{F}}(t)}{dt} = \int_{S(t)} \frac{\partial \mathbf{F}(\mathbf{r},t)}{\partial t}\cdot\mathbf{da} - \oint_{C(t)} \mathbf{F}(\mathbf{r},t)\cdot\left(\mathbf{dl}\ \mathbf{x}\ \mathbf{v}\right) + \int_{S(t)} \left[\nabla\cdot\mathbf{F}(\mathbf{r},t)\right]\mathbf{v}\cdot\mathbf{da} \tag{27.3.14}$$

If we use the vector identity $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = (\mathbf{C}\times\mathbf{A})\cdot\mathbf{B}$ in the second term on the right, we can also write this as

$$\frac{d\Phi_{\mathrm{F}}(t)}{dt} = \int_{S(t)} \frac{\partial \mathbf{F}(\mathbf{r},t)}{\partial t}\cdot\mathbf{da} - \oint_{C(t)} \left[\mathbf{v}\times\mathbf{F}(\mathbf{r},t)\right]\cdot\mathbf{dl} + \int_{S(t)} \left[\nabla\cdot\mathbf{F}(\mathbf{r},t)\right]\mathbf{v}\cdot\mathbf{da} \tag{27.3.15}$$

which is the result we were after. What this equation says is that the flux of **F** through the moving open surface can change in three ways. First, there can be changes due to the innate time dependence of **F** (first term on the right in (27.3.15)). But also the flux can change because flux is lost out of the boundary of the surface as it moves along (second term on the right). And finally, the flux can change because the surface sweeps across sources of **F**, that is regions where the divergence of F is non-zero.

27.3.3 Faraday's Law in Integral Form

We are now ready to write Faraday's Law in integral form. We apply (27.3.15) to the magnetic field, using the fact that we always have $\nabla\cdot\mathbf{B} = 0$

$$\frac{d\Phi_B(t)}{dt} = \int_{S(t)} \frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t}\cdot\hat{\mathbf{n}}\,da - \oint_{C(t)} \left[\mathbf{v}\times\mathbf{B}(\mathbf{r},t)\right]\cdot d\mathbf{l} \tag{27.3.16}$$

Using this equation, (27.3.2) becomes

$$\oint_{C(t)} \left[\mathbf{E} + \mathbf{v}\times\mathbf{B}(\mathbf{r},t)\right]\cdot d\mathbf{l} = -\frac{d}{dt}\int_{S(t)} \mathbf{B}\cdot\hat{\mathbf{n}}\,da \tag{27.3.17}$$

But we have just seen this term on the left before. Non-relativistically it is the electric field *as seen in the rest frame of d***l**. Equation (27.3.17) is the correct form for the integral version of Faraday's Law. If we define $\overline{\mathbf{E}} = \mathbf{E} + \mathbf{v} \times \mathbf{B}(\mathbf{r},t)$, then we have

$$\oint_{C(t)} \overline{\mathbf{E}} \cdot d\mathbf{l} = -\frac{d\Phi_B(t)}{dt} \tag{27.3.18}$$

27.4 Faraday's Law Applied to Circuits with Moving Conductors

We consider several examples that demonstrate how the form of Faraday's Law given in (27.3.18) applies to moving conductors.

27.4.1 The Falling Magnet/Falling Loop

To see how our final form of Faraday's Law removes the distinction between whether the magnet is moving or the conductor is moving, consider the following problem. We have a magnet with dipole moment $\mathbf{m} = m\hat{\mathbf{z}}$ and mass *M*, and associated magnetic field $\mathbf{B}_{dipole}$. The dipole moment of the magnet is always up, and the magnet is constrained to move only on the z-axis, but it is allowed to move freely up and down on that axis (see Figure 27-6). Let $Z(t)$ be the location of the magnet at time *t*. The z-axis is

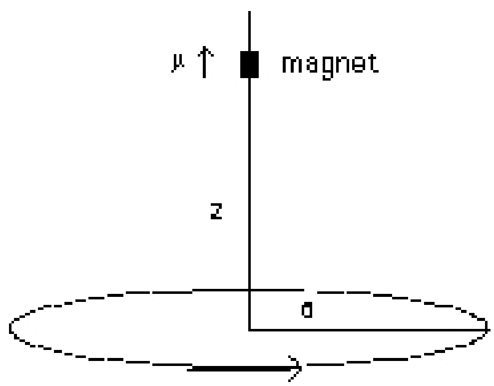


Figure 27-6: A magnet falling on the axis of a conducting loop

also the axis of a circular stationary loop of radius *a*, resistance *R*, and inductance *L*, fixed in place at *z = 0*. The magnet moves downward under the influence of gravity due to the force $-Mg\hat{\mathbf{z}}$. There will be a current *I* induced in the loop as the magnet falls, because of the changing magnetic flux through the loop, and that current will produce a magnetic field $\mathbf{B}_{loop}(\mathbf{r},t)$. The falling magnet will feel a force due to the current it induces in the loop due to the magnetic field $\mathbf{B}_{loop}$, given by $\mathbf{m}\cdot\nabla\mathbf{B}_{loop}$ (see Griffiths equation (6.3) page 258). The equation of motion is thus

$$M\frac{d^2Z}{dt^2} = -Mg + m\frac{dB_z^{loop}}{dz} \qquad (27.4.1)$$

We need another equation to solve the problem, which we get from Faraday's Law. Faraday's Law (27.3.18) applied to the loop is

$$\oint \overline{\mathbf{E}}\cdot\mathbf{dl} = -\frac{d}{dt}\int \mathbf{B}\cdot\mathbf{dA} = -\frac{d}{dt}\int\left[\mathbf{B}_{loop} + \mathbf{B}_{dipole}\right]\cdot\hat{\mathbf{n}}\,da \qquad (27.4.2)$$

From the definition of the self-inductance of the loop L, we have $LI = \int \mathbf{B}_{loop}\cdot\hat{\mathbf{n}}\,da$, so

$$\oint \overline{\mathbf{E}}\cdot\mathbf{dl} = -L\frac{dI}{dt} - \frac{d}{dt}\int\left[\mathbf{B}_{dipole}\right]\cdot\hat{\mathbf{n}}da \qquad (27.4.3)$$

If ρ is the resistivity of the loop material, Ohm's Law in microscopic form is $\overline{\mathbf{E}} = \rho\,\mathbf{J}$, and if $A_{\text{cross section}}$ is the area of the cross section of the wire, then $J = I / A_{\text{cross section}}$ and

$$\oint \overline{\mathbf{E}}\cdot\mathbf{dl} = \oint \rho\,\mathbf{J}\cdot\mathbf{dl} = \frac{I}{A_{\text{cross section}}}\oint \rho\,dl = I\left[\frac{2\pi a\rho}{A_{\text{cross section}}}\right] = IR \qquad (27.4.4)$$

so that Faraday's Law can be written as

$$IR = -L\frac{dI}{dt} - \frac{d}{dt}\int \mathbf{B}_{dipole}\cdot\mathbf{dA} \qquad (27.4.5)$$

Here is the crucial point. If we were to apply Faraday's Law to the situation where the magnet is at rest and the ring is falling, (27.4.5) ***does not change.*** It does not change because in getting to (27.4.4), we used Ohm's Law in microscopic form (see Section 26.2, (26.2.8), and the proper electric field to use in Ohm's Law is always the electric field in the rest frame of the conductor. That is what causes charges to move, and that is where we want to evaluate the electric field and not in any other frame. It is clear then that our equations of motion (27.4.5) and (27.4.1) depend only on the relative position of the magnet and the loop and the rate at which that is changing, and not on whether the loop is moving or the magnet is moving, or whether we are in a frame where both move. The resultant relative motion of the two is the same, regardless of the inertial frame in which we describe it.

27.4.2 A Circuit with a Sliding Bar

We consider another problem with a moving conductor. In Figure 27-7 we have a highly conducting cylindrical bar has mass *M*. It moves in the +*x*-direction along two frictionless horizontal rails separated by a distance *W*, as shown in the sketch. The rails are connected on the far right by a resistor of resistance *R*, as shown. The resistance of the bar, the rails, and the external resistor is $R_{\textbf{total}}$. For $t < 0$, the bar is in the region

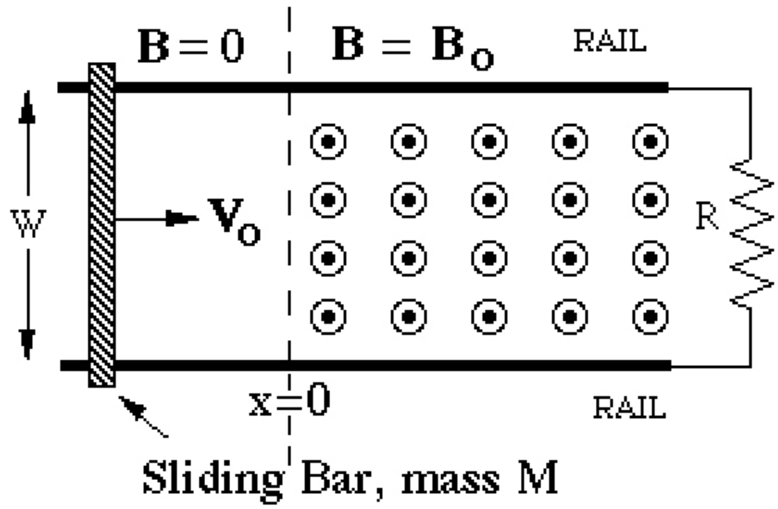


Figure 27-7: The sliding bar circuit

$x < 0$, sliding at a constant speed V_0 through a region with no magnetic field (see sketch). At time $t = 0$, when the bar is at location $x = 0$, the bar enters a region containing magnetic field $\boldsymbol{B_o}$, which is directed out of paper. After this time, suppose the bar has speed $V(t)$.

Faraday's Law tells us that $\oint \overline{\mathbf{E}} \cdot \mathbf{dl}$ around the circuit is equal to the negative of the time rate of change of the magnetic flux. Once the bar enters the field the flux is

$$\Phi(t) = \int_{surface} \mathbf{B} \cdot \hat{\mathbf{n}}\, da = B_o (L - X(t))W \qquad \text{where} \quad \frac{dX(t)}{dt} = V(t) \tag{27.4.6}$$

In the surface integral, we have taken $\hat{\mathbf{n}}$ to be out of the page, which means that the direction of the contour integral $d\mathbf{l}$ is positive counterclockwise. So we have

$$\oint \overline{\mathbf{E}} \cdot \mathbf{dl} = IR_{\text{total}} = -\frac{d\Phi}{dt} = +B_o V(t) W$$
$$\Rightarrow I = \frac{B_o V(t) W}{R_{\text{total}}} \tag{27.4.7}$$

where the plus sign in this equation means that the induced electric field and the resultant current is counterclockwise around the circuit. This means that the current in the bar is down in the drawing above.

Again, the crucial point here is that when we do the evaluation of $\oint \overline{\mathbf{E}} \cdot \mathbf{dl}$, we can always use Ohm's Law in the form $\overline{\mathbf{E}} = \rho\,\mathbf{J}$, and we will always end up with the total resistance of the resistors in the circuit, even though some are moving and some are not. This is because in Faraday's Law we are always evaluating the electric field in the rest frame of the circuit element.

If we ask about the force $\boldsymbol{F}$ on the bar at time $t > 0$, we see that the total force on the bar is the force per unit length $\mathbf{I} \times \mathbf{B}_o$ times the length W, and the direction is in the $-\hat{\mathbf{x}}$ direction, so we have

$$\mathbf{F} = -\hat{\mathbf{x}} W I B_o = -\hat{\mathbf{x}} V(t) \frac{\left(B_o W\right)^2}{R_{\text{total}}} \qquad (27.4.8)$$

Given this force $\boldsymbol{F}$ on the bar, the differential equation for $V(t)$ for $t > 0$, is

$$\begin{aligned} & M \frac{dV}{dt} = -V(t) \frac{\left(B_o W\right)^2}{R_{\text{total}}} \\ & \Rightarrow V(t) = V_o e^{-t/\tau} \\ & \text{where} \quad \tau = \frac{M R_{\text{total}}}{B_o^2 W^2} \text{ is the e-folding time} \end{aligned} \qquad (27.4.9)$$

The time dependence of I is given by (see (27.4.7))

$$I(t) = \frac{B_o W}{R_{\text{total}}} V_o e^{-t/\tau} \qquad (27.4.10)$$

Thus we easily have

$$\begin{aligned} & \frac{d}{dt}\left(\frac{1}{2} M V^2\right) = M V \frac{dV}{dt} = \frac{M}{\tau} V_o^2 e^{-2t/\tau} \\ & = \frac{M}{\tau} I^2 \left(\frac{R_{\text{total}}}{B_o W}\right)^2 = I^2 \frac{M B_o^2 W^2}{M R_{\text{total}}} \left(\frac{R_{\text{total}}}{B_o W}\right)^2 = I^2 R_{\text{total}} \end{aligned} \qquad (27.4.11)$$

Thus the Joule heating rate at any time is equal to the rate at which the moving bar is losing kinetic energy, as we would expect from the conservation of energy.

27.4.3 A Loop Falling Out of a Magnetic Field

We consider one last example of Faraday's Law applied to moving conductors.

A loop of mass *M*, resistance *R*, inductance *L*, height *H*, and width *W* sits in a magnetic field given by $\mathbf{B} = \hat{\mathbf{x}}\begin{cases} B_o & z \geq 0 \\ 0 & z < 0 \end{cases}$. At *t = 0* the loop is at rest and its mid-point is at *z = 0*, as shown in Figure 27-8, and the current in the loop is zero at *t = 0*. The acceleration of gravity is downward at *g*.

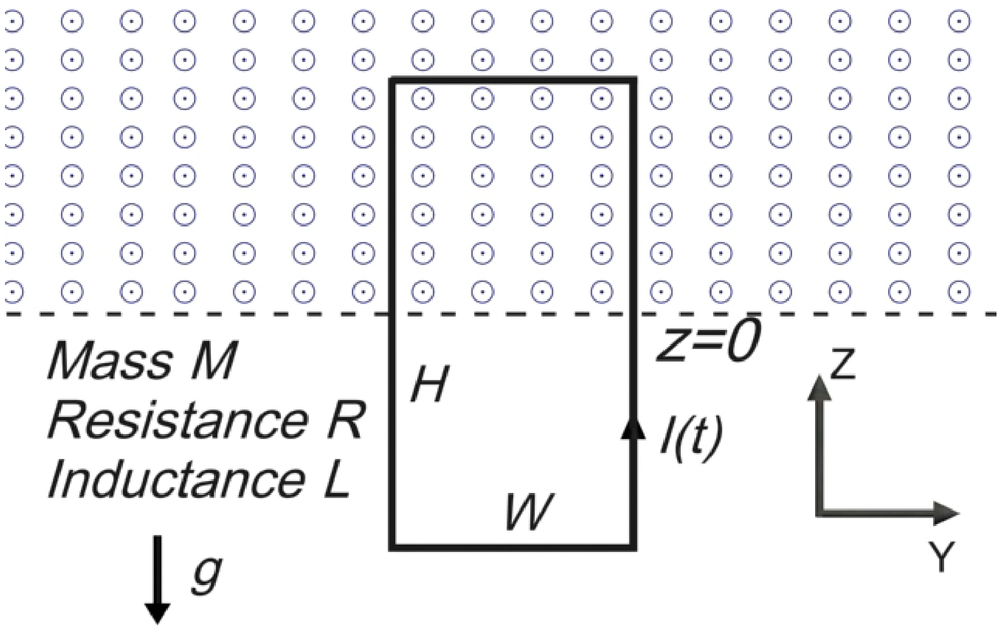


Figure 27-8: A Loop Falling Out of a Magnetic Field

The two differential equations that determine the subsequent behavior of the loop are as follows, where we take the direction of positive current to be counterclockwise, as shown in the figure. From Faraday's Law we have

$$IR = -L\frac{dI}{dt} - \frac{d}{dt}W\left[B_o(H/2 + z(t))\right]$$
$$\Rightarrow IR = -L\frac{dI}{dt} - WB_o\mathrm{v}(t) \tag{27.4.12}$$

and the equation of motion is simply

$$M\frac{d}{dt}\mathrm{v}(t) = -Mg + IWB_o \tag{27.4.13}$$

If we multiply the first equation by *I* and the second equation by v, we have

$$I^2R = -LI\frac{dI}{dt} - WIB_o\mathrm{v}(t) = -\frac{d}{dt}\frac{1}{2}LI^2 - WIB_o\mathrm{v}(t) \qquad (27.4.14)$$

and

$$M\mathrm{v}\frac{d}{dt}\mathrm{v}(t) = -Mg\mathrm{v} + IWB_o\mathrm{v} = \frac{d}{dt}\frac{1}{2}M\mathrm{v}^2 \qquad (27.4.15)$$

Putting these two together gives conservation of energy

$$\frac{d}{dt}\left[\frac{1}{2}M\mathrm{v}^2 + Mgz + \frac{1}{2}LI^2\right] = -I^2R \qquad (27.4.16)$$

Let solve the case where the resistance of the loop is zero. Assuming that the loop never falls out of the magnetic field, (27.4.13) is

$$\begin{aligned} &M\frac{d}{dt}\mathrm{v}(t) = -Mg + IWB_o \\ &\Rightarrow M\frac{d^2}{dt^2}\mathrm{v}(t) = WB_o\frac{dI}{dt} \end{aligned} \qquad (27.4.17)$$

and (27.4.14) is

$$\begin{aligned} &0 = -L\frac{dI}{dt} - WB_o\mathrm{v}(t) \\ &\Rightarrow \frac{dI}{dt} = -\frac{WB_o\mathrm{v}(t)}{L} \end{aligned} \qquad (27.4.18)$$

Thus we have

$$\begin{aligned} &M\frac{d^2}{dt^2}\mathrm{v}(t) = WB\frac{dI}{dt} = -\frac{W^2B_o^2}{L}\mathrm{v}(t) \\ &\Rightarrow \frac{d^2}{dt^2}\mathrm{v}(t) + \frac{W^2B_o^2}{ML}\mathrm{v}(t) = 0 \end{aligned} \qquad (27.4.19)$$

And therefore our solution for the velocity is

$$\mathrm{v}(t) = -\frac{g}{\omega}\sin(\omega t) \quad \text{where} \quad \omega^2 = \frac{W^2B_o^2}{ML} \qquad (27.4.20)$$

where we have picked the sin so that v is 0 at $t = 0$. This implies that

$$z(t) = \frac{g}{\omega^2}\left(\cos\left(\omega t\right) - 1\right) \tag{27.4.21}$$

where we have picked the integration constant so that z is 0 at t =0
Finally, we determine the constant A from the fact that we have to satisfy at $t = 0$ the equation

$$M\frac{d}{dt}\mathrm{v}(t) = -Mg + IWB_o \tag{27.4.22}$$

This means that we must have $M\frac{d}{dt}\mathrm{v}(0) = MA\omega\cos\left(0\right) = -Mg \Rightarrow A = -\frac{g}{\omega}$

so

$$\mathrm{v}(t) = -\frac{g}{\omega}\sin\left(\omega t\right) \quad \text{where} \quad \omega^2 = \frac{W^2 B_o^2}{ML} \quad \text{and} \quad z(t) = \frac{g}{\omega^2}\left(\cos\left(\omega t\right) - 1\right) \tag{27.4.23}$$

28 EMF's and Faraday's Law in Circuits

28.1 Learning Objectives

We now consider the concept of electromotive force in a circut. We have seen above that a given observer may think that the motion of charges is driven either by a $q\mathbf{E}$ or a $q\mathbf{v}\times\mathbf{B}$ force in the observer's rest frame, or some combination of both, but that this always reduces to the electric field in the rest frame of the circuit element. There may also be or some other force entirely (see our example of the battery below), and that leads us to consider the general concept of an EMF. We then turn our attention to the typical circuit elements: batteries, resistors, capacitors, and inductors. We pay special attention to inductors, since there are a huge number of misconceptions about the "voltage drop" across an inductor. We will only consider single loop circuits here. In the beginning we will only consider circuits with batteries with resistors. Then we will add inductors, and finally capacitors. ***Unlike in the situation above, in this Section we assume that all parts of the circuit are at rest.***

28.2 The electromotive force

Suppose we have a current flowing in a closed circuit. To have a flow of current in our single loop circuit, there must be at every point in the circuit some force per unit charge **f** on the charge carriers which causes them to move. For the moment we consider a circuit containing only batteries and resistors. Ohm's Law states the relation between the force per unit charge **f** at any point in the circuit and the current density **J** at that point is $\mathbf{J} = \sigma_c\mathbf{f}$. This is the same relation we have derived in Section 26.2, except that there we considered only the force per unit charge due to an electric field **E,** *whereas here we* *consider the force per unit charge due to any force.* The constant σ_C is called the conductivity (we use the symbol σ_C to distinguish this quantity from surface charge σ). Why do we have this particular relationship? The classical model is given in Section 26.2, and we do not repeat those arguments here, other than that we point out that those arguments easily generalize from the electric field to *any* force per unit charge **f.**

Now, the electromotive force (or emf) of our single loop circuit is denoted by the symbol $\mathcal{E}$, and is defined by the equation

$$\mathcal{E} = \oint \mathbf{f} \cdot d\mathbf{l} \tag{28.2.1}$$

where the integral is around the complete circuit, and at every point in the circuit **f** is the force per unit charge that is felt by the charge carriers located at that point (the same **f** that we were dealing with above). The terminology here is poor, since an "electromotive force" is not a force at all, but instead is a closed line integral of a force per unit charge. Note that the units of emf are Joules/coulomb, or Volts.

In any case, it is the emf $\mathcal{E}$ defined in (28.2.1) that determines how much current will flow in a circuit, by the following argument. The crucial step in this argument is understanding the following point.

For a single loop circuit, the current *I* is to an good approximation the same in all parts of the circuit

Why is this so? Basically, although the current will start out at *t=0* being unequal in different parts of the circuit, those inequalities mean that charge is piling up somewhere. The accumulating charge at the pile up will quickly produce an electric field, and *this electric field is always in the sense so as to even out the inequalities in the current*. We give a semi-qualitative example of this evening-out below in Section 28.8, but for the moment we simply accept it.

The upshot is that in a very short time electric fields will be set up around the circuit along with various pockets of accumulated charge, all arranged so as to made the current the same in every part of the circuit. This will be true all around the circuit as long as we long as we are considering time changes in the current that are long compared to the speed of light transit time across the circuit.

Given this, let's figure out how the resultant current *I* is related to the emf $\mathcal{E}$. We have

$$\mathcal{E} = \oint \mathbf{f} \cdot d\mathbf{l} = \oint \frac{\mathbf{J}}{\sigma_c} \cdot d\mathbf{l} = \oint \frac{I}{A\sigma_c} dl = I \oint \frac{dl}{A\sigma_c} \tag{28.2.2}$$

where A is the cross-sectional area of the circuit at any point, and σ_C its conductivity, and both of these quantities can vary at different points in the circuit. However, the current I

does not vary around the circuit, which allows us to take it out of the integral in equation (28.2.2). We have also assumed that the current at any point is uniformly distributed across the cross-sectional area *A*. This is not a crucial assumption, and we can relax that easily. Thus if we define the total resistance R_T of our one loop circuit as

$$R_T = \oint \frac{dl}{A\sigma_c} \tag{28.2.3}$$

then we have the relationship between the emf, the current *I*, and the total resistance of the circuit R_T :

$$I = \frac{\mathcal{E}}{R_T} \tag{28.2.4}$$

From equation (28.2.4), the units of the resistance are clearly *volt/amp* which we define as an *ohm*. Thus we see that the conductivity σ_C must have units of $ohm^{-1}\ m^{-1}$. Sometimes we will refer to the resistivity of a material. The resistivity is just the inverse of the conductivity, and has units of *ohm meters*.

28.3 An example of an EMF: batteries

Let's apply these ideas to a concrete example--a simple circuit consisting of a battery connected by highly conducting wires to a resistor (a conductor with low conductivity σ_C). Batteries provide one of the most familiar examples of a source of electromagnetic energy. Batteries are in many ways like capacitors, with one fundamental difference. There are chemical people inside batteries that do work, and provide an "electromotive force", as follows. The positive terminal of a battery carries positive charge, just like the positive plate of a (charged) capacitor, and similarly for the negative terminal (see sketch next page). There is therefore an internal electric field $\mathbf{E}_B$ in the battery, going from the positive terminal to the negative terminal, and the positive terminal of the battery is at a higher electric potential than the negative terminal. When the battery is placed in a circuit, say with a resistor, there is then a path for positive charge to flow from higher to lower potential (through the resistor--see below), and charge will do just that when the circuit is established. Up to this point, we have a situation that looks very much like a capacitor that is discharging through a resistor.

But the terminals of the battery do *not* discharge. The cartoon essence of a battery is the following. Suppose a charge $+dq$ leaves the positive terminal, flows through the resistor, and then arrives at the negative terminal. When that charge arrives at the negative terminal, a chemical person picks it up and, and applying a force per unit charge $\mathbf{f}_S$, moves it *against* the internal electric field of the battery, and deposits it on the positive terminal again. No matter how rapidly charge flows off the positive terminal through the external circuit and arrives at the bottom plate, the chemical people manage to keep up, transferring the incoming charge on the negative terminal to the positive terminal, as fast as it arrives. Clearly our "chemical people" are doing work in this process, just as we do work in charging a capacitor, because they are moving positive charge against the electric field of the battery by applying the force $\mathbf{f}_S$.

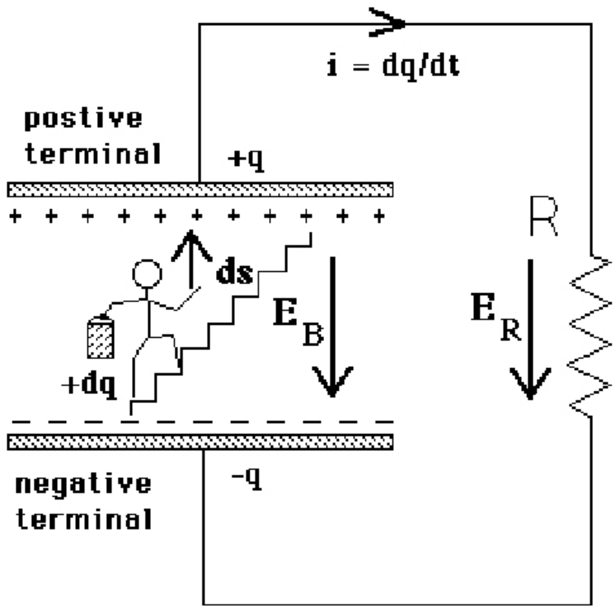


Figure 28-1: A cartoon view of a battery in a circuit

However, if we look at the complete circuit, there are really two forces per unit charge involved in driving current around the circuit: the "source" of the emf, $\mathbf{f}_S$, which is ordinarily confined to one portion of the loop (inside the battery here), and the electrostatic force per unit charge **E** whose function is to smooth out the flow and communicate the influence of the source to distant parts of the circuit (*Griffiths*, page 292). The total force per unit charge, **f**, that is the **f** that appears in equation (28.2.1) above, is therefore given by the sum

$$\mathbf{f} = \mathbf{f}_S + \mathbf{E} \tag{28.3.1}$$

In the case of the circuit here, a battery and a resistor, $\oint \mathbf{E} \cdot \mathbf{dl} = 0$ because this is an electrostatic field. Furthermore, $\mathbf{f}_S$ vanishes outside of the battery, so that we have simply that $\mathcal{E} = \int_{\text{bottom battery}}^{\text{top battery}} \mathbf{f}_s \cdot \mathbf{dl}$. This is just the work done per unit charge by the chemical people in moving the charge against the electric field of the battery from the bottom plate to the top plate. The *rate* at which they are doing work is *dq/dt* times this work per unit charge, or

$$P_{\text{rate at which work done by battery}} = I\mathcal{E} \tag{28.3.2}$$

and this is the rate that the battery is providing energy to the circuit.

The origin of the electromotive force in a battery is the internal mechanism (the "chemical people") that transports charge carriers in a direction opposite to that in which the electric field would move them. In ordinary batteries, it is chemical energy that makes the charge carriers move against the internal field of the battery. That is, a positive charge will move towards higher electric potential if in so doing it can engage in a chemical reaction that will yield more energy than it costs to move against the electric field. The electromotive force in a battery depends on atomic properties. The values of potential differences between the battery terminals lie in the range of volts because the binding energies of the outer electrons of atoms are in the range of several electron volts, and it is essentially the differences in these binding energies that determine the voltage of the battery.

The actual details of all this are complicated, so we will not go beyond the cartoon level. Purcell (*Electricity and Magnetism*, *Berkeley Physics Course*, *Volume 2*, McGraw-Hill, 1965) has a excellent discussion of batteries in Chapter 4 of that volume.

28.4 The resistance of a resistor

Finally, let's finish up our discussion of the simple battery and resistor circuit shown in Figure 28-1. Suppose the resistor in our circuit consists of a conducting cylinder of length *L* and cross-sectional area A, with conductivity . Since the only force per unit charge in the resistor is the electric field in the resistor, $\mathbf{E}_R$ (the battery "source" $\mathbf{f}_S$ is zero there), the current density in the resistor, $\mathbf{J}_R$, is $\sigma_c \mathbf{E}_R$ by the microscopic form of Ohm's Law. Let $\Delta V_R = -\int_{\text{bottom R}}^{\text{top R}} \mathbf{E}_R \cdot d\mathbf{l}$ be the potential difference from the bottom to the top of the resistor (see Figure 28-1). Then

$$I / A = J_R = \sigma_c E_R = \sigma_c \Delta V_R / L \tag{28.4.1}$$

Solving equation (28.4.1) for *I* in terms of ΔV_R, we obtain

$$I = \Delta V_R / R \quad \text{with} \quad R = \frac{L}{\sigma_c A} \tag{28.4.2}$$

This is the macroscopic form of Ohm's Law with which we are most familiar. The quantity *R* is the resistance of the resistor, in ohms, and is a function of both the fundamental properties of the material, via σ_c and of the shape of the material, via *A* and *L*.

Suppose now that the conductivity in the connecting wires and in the battery is so much larger than σ_c that we can take them to be infinite for all practical purpose. Then equation (28.2.3) becomes $R_T = R$ That is, the total resistance of our circuit is to a good approximation just the resistance of our (low conductivity) resistor. If we consider equation (28.2.4) (with $R_T = R$), we obtain $I = \mathcal{E}/R = \Delta V_R/R$, and therefore $\mathcal{E} = \Delta V_R$.

Since $\oint \mathbf{E}\cdot d\mathbf{l} = 0$, this only can be true if inside the battery we have approximately that $\mathbf{f}_s = -\mathbf{E}_B$, where $\mathbf{E}_B$ is the electric field in the battery. Note that since **J** and **E** are anti-parallel in the battery, the creation rate of electromagnetic energy, $-\mathbf{E}\cdot\mathbf{J}$, is positive there, and therefore electromagnetic energy is being created in the battery.

28.5 Joule heating

Since $\mathbf{J}_R$ and $\mathbf{E}_R$ are parallel in the resistor, the creation rate of electromagnetic energy, $-\mathbf{E}\cdot\mathbf{J}$, is negative, and therefore electromagnetic energy must be being destroyed in the resistor. Where is this energy going? Well, the electric field is doing work on the charges at the rate of $qn\mathbf{E}_\mathbf{R}\cdot\mathbf{v}_e$, or $+\mathbf{E}_\mathrm{R}\cdot\mathbf{J}_\mathrm{R}$, per unit volume, just as we would expect (if electromagnetic energy is disappearing, it must be going some place). That work done by the field in the steady state is transmitted from the charges to the lattice via collisions, i.e., to random thermal energy. The total rate at which this heating takes place is $+\mathbf{E}_\mathrm{R}\cdot\mathbf{J}_\mathrm{R}$ times the volume *AL*, or *EJAL*, or *VI*. Thus the heating rate P_{heating}, in joules per second, is
(28.6.1)

$$P_{\text{heating}} = \Delta V_R I = I^2 R \tag{28.5.1}$$

This is the familiar form of the Joule Heating Law. If we furthermore use the fact that $\mathcal{E} = \Delta V_R$, we have that the rate at which the battery is doing work is equal to the rate at which energy is appearing as heat in the resistor. Thus electromagnetic energy is being created in the battery at the same rate at which it is being destroyed in the resistor.

28.6 Self-inductance and simple circuits with "one-loop" inductors:

The addition of time-changing magnetic fields to simple circuits means that the closed line integral of the electric field around a circuit is no longer zero. Instead, we have, for any open surface

$$\oint_C \mathbf{E}\cdot d\mathbf{l} = -\int_S \frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t}\cdot\hat{\mathbf{n}}\,da \tag{28.6.1}$$

Any circuit in which the current changes with time will have time-changing magnetic fields, and therefore associated “induced” electric fields, *which are due to the time changing currents, not to the time changing magnetic field (association is not causation).* How do we solve simple circuits taking such effects into account? We discuss here a consistent way to understand the consequences of introducing time-changing magnetic fields into circuit theory--that is, *self-inductance.*

As soon as we introduce time-changing currents, and thus time changing magnetic-fields, the electric potential difference between two points in our circuit is not longer well-defined. When the line integral of the electric field around a closed loop is no longer zero, the potential difference between points *a* and *b*, say, is no longer independent of the path used to get from *a* to *b*. That is, the electric field is no longer a conservative field, and the electric potential is no longer an appropriate concept (that is, **E** can no longer be written as the negative gradient of a scalar potential). However, we can still write down in a straightforward fashion the differential equation for *I(t)* that determines the time-behavior of the current in the circuit.

To show how to do this, consider the circuit shown in Figure 28-2. We have a battery, a resistor, a switch *S* that is closed at $t = 0$, and a "one-loop inductor". It will become clear what the consequences of this "inductance" are as we proceed. For $t > 0$, current will flow in the direction shown (from the positive terminal of the battery to the negative, as usual). What is the equation that governs the behavior of our current *I* for $t > 0$?

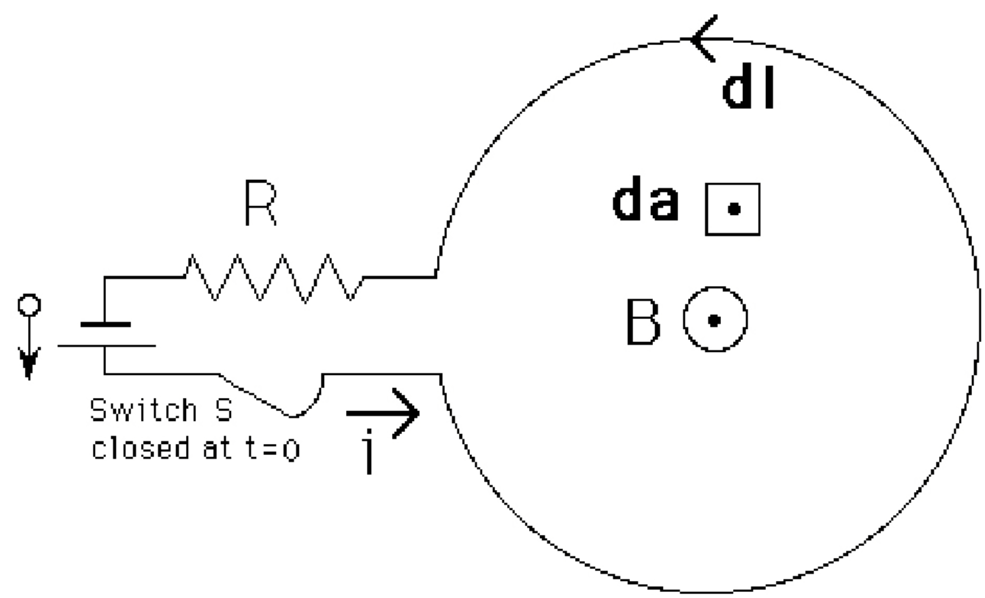


Figure 28-2: A simple circuit with battery, resistor, and a one-loop inductor

To investigate this, we apply Faraday's Law to the open surface bounded by our circuit, where we take $d\mathbf{a} = \hat{\mathbf{n}}\, da$ to be out of the page, and thus $d\mathbf{l}$ is counter-clockwise, as shown. First, what is the integral of the electric field around this circuit? That is, what is the left-hand side of (28.6.1)? Well, there is an electric field in the battery in the direction of $d\mathbf{l}$ that we have chosen, we are moving against that electric field, so that $\int \mathbf{E} \cdot d\mathbf{l}$ is negative. Thus the contribution of the battery to our integral is $-\mathcal{E}$ (see the discussion in Section 28.3 above). Then there is an electric field in the resistor, in the direction of the current, so when we move through the resistor in that direction, $\int \mathbf{E} \cdot d\mathbf{l}$ is positive, and that contribution to our integral is $+I\,R$. What about when we move through our "one-loop inductor"? There is no electric field in this loop if the resistance of the wire making up the loop is zero, so there is no contribution to $\int \mathbf{E}\cdot d\mathbf{l}$ from this part of the circuit. This may bother you, and we talk at length about it below. So, going totally around the closed loop, we have

$$\oint \mathbf{E}\cdot d\mathbf{l} = -\mathcal{E} + I\,R \tag{28.6.2}$$

Now what is the right hand side of (28.6.1). Since we have assumed in this Section that the circuit is not moving, we can take the partial with respect to time outside of the surface integral and then we simply have the time derivative the magnetic flux through the loop. What is the magnetic flux through the open surface? First of all, we arrange the geometry so that the part of the circuit which includes the battery, the switch, and the resistor, makes only a small contribution to the magnetic flux as compared to the (much larger area) of the open surface which consititutes our "one-loop inductor". Second, we know that the sign of the magnetic flux is positive in that part of the circuit because current flowing counter-clockwise will produce a **B** field out of the paper, which is the same direction of $\hat{\mathbf{n}}\,da$, so that $\mathbf{B}\cdot\hat{\mathbf{n}}\,da$ is positive. Note that our magnetic field here is the *self* magnetic field—that is the magnetic field produced by the current flowing in the circuit, and not by an currents external to this circuit.

We also know that at any point in space, **B** is proportional to the current *I*, since it can be computed from the Biot-Savart Law, that is,

$$\mathbf{B}(\mathbf{r},t) = \frac{\mu_o I(t)}{4\pi}\oint \frac{d\mathbf{l}'\times(\mathbf{r}-\mathbf{r}')}{\left|(\mathbf{r}-\mathbf{r}')\right|^3} \tag{28.6.3}$$

You may immediately object that the Biot-Savart Law is only good in time-independent situations, but in fact, as we have seen before when considering radiation, as long as the current is varying on time scales *T* long compared to the speed of light travel time across the circuit and we are within a distance *cT* of the currents, then (28.6.3) is an excellent approximation to the time dependent magnet field. If we look at (28.6.3), although for a general point in space it involves a very complicated integral over the circuit, it is clear that $\mathbf{B}(\mathbf{r},t)$ is everywhere propostional to *I(t)*. That is, if we double the current, **B** at any point in space will also double. It then follows that the magnetic flux itself must be proportional to *I*, because it is the surface integral of **B**, and **B** is everywhere proportional to *I*. That is,

$$\Phi(t) = \int_{S(t)} \mathbf{B}(\mathbf{r},t)\cdot\hat{\mathbf{n}}da = \int_S \left\{ \frac{\mu_o I(t)}{4\pi}\oint \frac{d\mathbf{l}'\times(\mathbf{r}-\mathbf{r}')}{\left|(\mathbf{r}-\mathbf{r}')\right|^3} \right\}\cdot\hat{\mathbf{n}}\,da = I(t)\int_S \left\{ \frac{\mu_o}{4\pi}\oint \frac{d\mathbf{l}'\times(\mathbf{r}-\mathbf{r}')}{\left|(\mathbf{r}-\mathbf{r}')\right|^3} \right\}\cdot\hat{\mathbf{n}}\,da \tag{28.6.4}$$

or

$$\Phi(t) = LI(t) \qquad\qquad L = \frac{\mu_o}{4\pi}\int_S \left\{ \oint \frac{d\mathbf{l}' \times (\mathbf{r}-\mathbf{r}')}{\left|(\mathbf{r}-\mathbf{r}')\right|^3} \right\} \cdot \hat{\mathbf{n}}\, da \tag{28.6.5}$$

So the magnetic flux is a constant *L* times the current. Note that *L* is a constant in the sense that it stays the same as long as we do not change the geometry of the circuit. If we change the geometry of the circuit (for example we halve the radius of the circle in our Figure above), we will change L, but for a given geometry, L does not change. Even though it may be terrifically difficult to do the integrals in (28.6.5), once we have done it for a given circuit geometry we know *L*, and *L* is a constant for that geometry. The quantity L is called the self-inductance of the circuit, or simply the inductance. From the definition in (28.6.5), you can show that the dimensions of *L* are μ_o times a length. In Assignment 10, you show that a lower limit to the inductance of a single loop of wire is proportional to μ_o times its radius.

Regardless of how hard or easy it is to compute *L*, it is a constant for a given circuit geometry and now we can write down the equation that governs the time evolution of *I*. If $\Phi(t) = LI(t)$, then $d\Phi(t)/dt = LdI(t)/dt$, and equation (28.6.1) becomes

$$\oint \mathbf{E}\cdot d\mathbf{l} = -\mathcal{E} + I\,R = -L\frac{dI}{dt} \tag{28.6.6}$$

If we divide (28.6.6) by *L* and rearrange terms, we find that the equation that determines the time dependence of *I* is

$$\frac{dI}{dt} + \frac{R}{L}I = \frac{\mathcal{E}}{L} \tag{28.6.7}$$

The solution to this equation given our initial conditions is

$$I(t) = \frac{\mathcal{E}}{R}(1 - e^{-tR/L}) \tag{28.6.8}$$

This solution reduces to what we expect for large times, that is $I = \frac{\mathcal{E}}{R}$, but it also shows a continuous rise of the current from 0 initially to this final value, with a characteristic time τ_L defined by

$$\tau_L = \frac{L}{R} \tag{28.6.9}$$

This time constant is known as the inductive time constant. This is the effect of having a non-zero inductance in a circuit, that is, of taking into account the "induced" electric fields which always appear when there are time changing **B** fields. And this is what we expect—the reaction of the system is to try to keep things the same, that is to delay the build-up of current (or its decay, if we already have current flowing in the circuit).

28.7 Kirchhoff's Second Law modified for inductors: a warning

We can write the governing equation for *I(t)* from above as

$$\mathcal{E} - I\,R - L\frac{dI}{dt} = \sum_{\mathrm{i}} \Delta V_i = 0 \tag{28.7.1}$$

where we have now cast it in a form that "looks like" a version of Kirchhoff's Second Law, a rule that is often quoted in elementary electromagnetism texts. Kirchhoff's Second Law states that the sum of the potential drops around a circuit is zero. In a circuit with no inductance, this is just a statement that the line integral of the electric field around the circuit is zero, which is certainly true if there is no time variation. However, in circuits with currents that vary in time, this "Law" is no longer true.

Unfortunately, many elementary texts choose to approach circuits with inductance by preserving "Kirchhoff's Second Law", or the loop theorem, by specifying the "potential drop" across an inductor to be *-LdI/dt* if the inductor is traversed in the direction of the current. Use of this formalism will give the correct equations. However, the continued use of Kirchhoff's Second Law with inductors is misleading at best, for the following reasons.

Kirchhoff's Second Law was originally based on the fact that the integral of **E** around a closed loop was zero. With time-changing currents and thus time-changing self-magnetic fields, *this is no longer true* (the **E** field is no longer conservative), and thus the sum of the "potential drops" around the circuit, if we take that to mean the *negative* of the closed loop integral of **E**, is ***no longer zero***--in fact it is LdI/dt (this is equation (28.6.6) with the sign reversed).

The continued use of Kirchhoff's Second Law in this way gives the right equations, but it confuses the physics. In particular, saying that there is a "potential drop" across the inductor of $-LdI/dt$ implies that there is an electric field in the inductor such that the integral of **E** through the inductor is equal to $-LdI/dt$. **This is not always, or even usually, true**. For example, suppose in our "one-loop" inductor above that the wire making up the loop has negligible resistance compared to the resistance *R*. The integral of **E** through our "one-loop" inductor above is then **very small**, **NOT** $-LdI/dt$. Why is it very small? Well, to repeat our assertion above

For a single loop circuit, the current *I* is to an good approximation the same in all parts of the circuit

This is just as valid in a circuit with inductance. Again, although the current may start out at $t=0$ being unequal in different parts of the circuit, those inequalities mean that charge is piling up somewhere. The accumulating charge at the pile-up will quickly produce an electric field, and this *electric field is always in the sense so as to even out the inequalities in the current*. In this particular case, if the conductivity of the wires making up our one-loop inductor is very large, then there will be a very small electric field in those wires, because it takes only a small electric field to drive any current you need. The amount of current needed is determined in part by the larger resistance in other parts of the circuit, and it is the charge accumulation at the ends of those low conductivity resistors that cancel out the field in the inductor and enhance it in the resistor, so as to maintain constant current in the circuit. We return to this point below in Section 28.8.

One final point, to confuse the issue further. If you have ever put the probes of a voltmeter across an the terminals of an inductor (with very small resistance) in a circuit, what you measured on the meter of the voltmeter was a "voltage drop" of $-LdI/dt$. But that is not because there is an electric field in the inductor! It is because putting the voltmeter in the circuit will result in a time changing magnetic flux through the voltmeter circuit, consisting of the inductor, the voltmeter leads, and the large internal resistor in the voltmeter. A current will flow in the voltmeter circuit because there will be an electric field in the large internal resistance of the voltmeter, with a potential drop across *that* resistor of $-LdI/dt$, by Faraday's Law applied to the voltmeter circuit, and that is what the voltmeter will read. The voltmeter as usual gives you a measure of the potential drop across its own internal resistance, but this is *not* a measure of the potential drop across the inductor. It is a measure of the time rate of change of magnetic flux in the voltmeter circuit! As before, there is only a very small electric field in the inductor if it has a very small resistance compared to other resistances in the circuit.

28.8 How can the electric field in an inductor be zero?

Students are always confused about the electric field in inductors, in part because of the kinds of problems they have seen. Quite often in simple problems with time varying magnetic fields, there is an “induced” electric field right where the time varying magnetic field was non-zero. What has changed in our circuit above to make the electric field zero in the wires of the (resistanceless) inductor zero, even though there is a time changing magnetic flux through it? This is a very suble point and a source of endless confusion, so let’s look at it very carefully.

Your intuition that there should be an electric field in the wires of an inductor is based on doing problems like that shown in Figure 28-3 below. We have a loop of wire of radius a and total resistance R immersed in an external magnetic field which is out of the page and increasing with time as shown. In considering this circuit, unlike in our “one-loop” circuit above, we neglect the the magnetic field due to the currents in the wire itself, assuming that the external field is much bigger than the self-field. The conclusions we arrive at here can be applied to the self-inductance case as well.

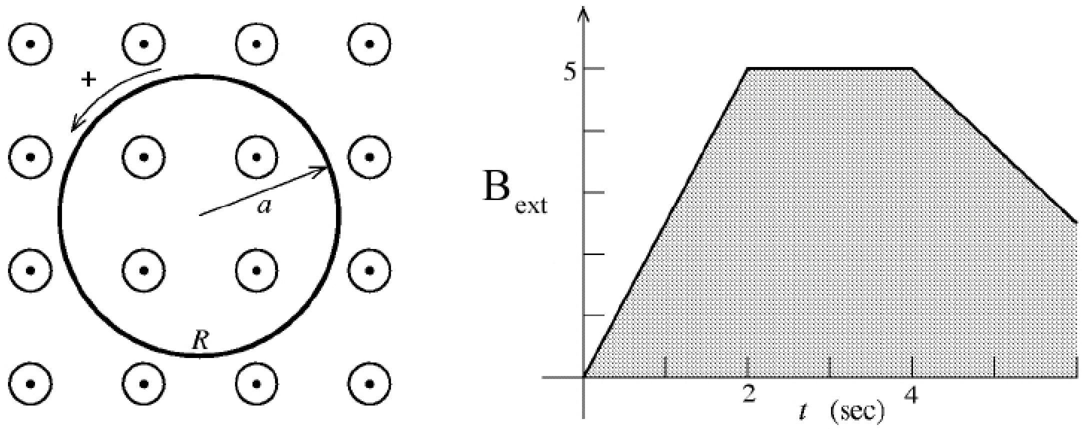


Figure 28-3: A loop of wire sitting in a time-changing external magnetic field

The changing extermal magnetic field will give rise to an "induced" electric field in the loop of the wire, with

$$\oint \mathbf{E}\cdot d\mathbf{l} = -\frac{d}{dt}\left(B_{ext}\pi a^2\right) \tag{28.8.1}$$

This "induced" electric field is azimuthal and uniformly distributed around the loop as long as the resistance in the loop is uniform, and in the loop itself we easily have from (28.8.1) that the electric field right at the loop is given by

$$\oint \mathbf{E}\cdot d\mathbf{l} = 2\pi a E_\phi = -\frac{d}{dt}\left(B_{ext}\pi a^2\right) \quad \Rightarrow \mathbf{E}\big|_{r=a} = -\hat{\boldsymbol{\phi}}\frac{a}{2}\frac{dB_{ext}}{dt} \tag{28.8.2}$$

Thus if the resistance is distribed uniformly around the wire loop, we get a uniform induced electric field in the loop, circulating clockwise for the external magnetic field increasing in time (see Figure 28-4). This electric field causes a current to flow, and the current will circulate clockwise in the same sense as the electric field. The total current in the loop will be the total "potential drop" around the loop divided by its resistance *R*, or

$$I = \frac{\pi a^2}{R}\frac{dB_{ext}}{dt} \tag{28.8.3}$$

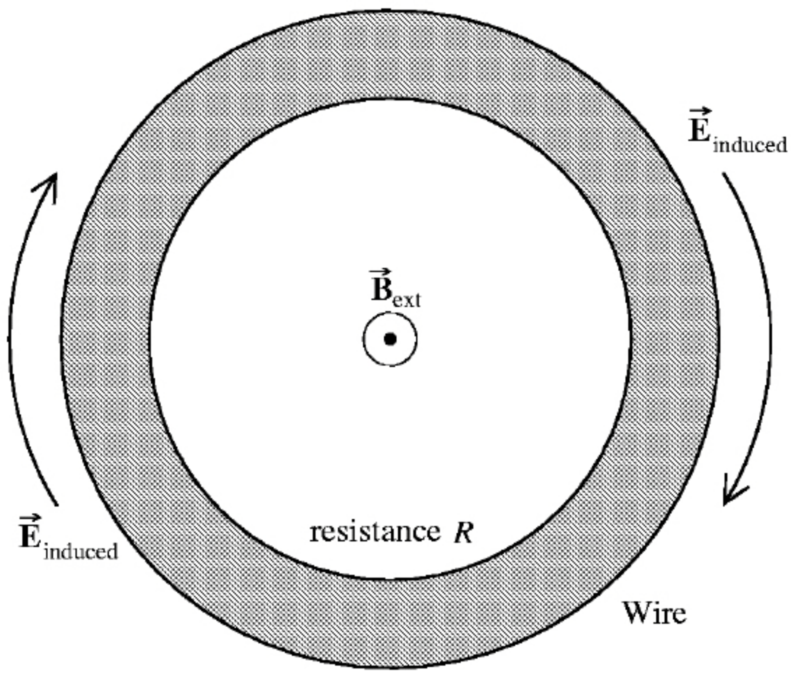


Figure 28-4: A loop of wire with resistance *R* in an external field out of the page

But what happens if we don't distribute the resistance uniformly around the wire loop? For example, let us make the left half of our loop out of wire with resistance R_1 and the right half of the loop out of wire with resistance R_2, with $R_1 + R_2 = R$, so that we have the same total resistance as before (see Figure 28-5). Let us futher assume that $R_1 < R_2$. How is the electric field distributed around the loop now?

First of all, the electromotive force around the loop (see (28.8.1)) is the same, as is the resistance, so that the current *I* has to be the same as in (28.8.3). Moreover it is the same on both sides of the loop by charge conservation. But the electric field in the left half of the loop $\mathbf{E}_1$ must now be different from the electric field in the right half of the loop $\mathbf{E}_2$. This is so because the line integral of the electric field on the left side is $\pi a E_1$, and from Ohm's Law in macroscopic form, this must be equal to IR_1. Similarly, $\pi a E_2 = IR_2$. Thus

$$\frac{E_1}{E_2} = \frac{R_1}{R_2} \Rightarrow E_1 < E_2 \text{ since } R_1 < R_2 \tag{28.8.4}$$

This makes sense. We get the same current on both sides, even though the resistances are different, and we do this by adjusting the electric field on the side with the smaller resistance to ***be*** smaller. Because the resistance is also smaller, we produce the same current as on the opposing side, with this smaller electric field.

But what happened to our uniform electric field. Well there are two ways to produce electric fields—one from time changing currents and their associated time changing magnet fields, and the other from electric charges. Nature accomplishes the reduction of $\mathbf{E}_1$ compared to $\mathbf{E}_2$ by charging at the junctions separating the two wire segments (see Figure 28-5), positive on top and negative on bottom.

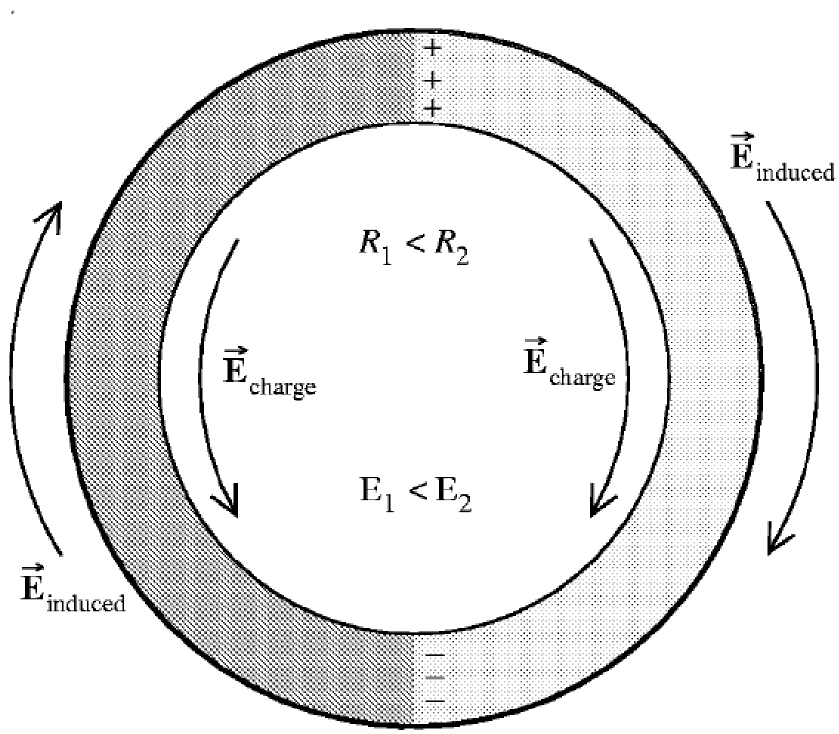


Figure 28-5: The electric field in the case of unequal resistances in the loop

The total electric field is the sum of the “induced” electric field and the electric field associated with the charges, as shown in the Figure above. It is clear that the addition of these two contributions to the electric field will reduce the total electric field on the left (side 1) and enhance it on the right (side 2). The field $\mathbf{E}_1$ will always be clockwise, but it can be made arbitrarily small by making $R_1 << R_2$.

Thus we see that we can make a non-uniform electric field in an inductor by using non-uniform resistance, even though our intuition tells us (correctly) that the “induced” electric field should be uniform at a given radius. All that Faraday’s Law tells us is that the line integral of the electric around a closed loop is equal to the negative of the time rate of change of the magnet flux through the enclosed surface. It does not tell us at what locations the electric field is non-zero around the loop, and it may be non-zero (or zero!) in unexpected places. The field in the wire making up the "one-loop" inductor we considered above is zero (or least very small) for exactly the kinds of reasons we have been discussing here.

29 Index

M

N

O

P

R

S

T

U

V

30 References

8

Cabral, B. and C. Leedom. Imaging Vector Fields Using Line Integral Convolution, Proc. SIGGRAPH '93, 263-270, 1993.. 21

Griffith David. J., *Introduction to Electrodynamics*, Prentice Hall, 1999 18

Hertz, Heinrich, *Electric waves; being researches on the propagation of electric action with finite velocity through space*, Macmillan and Co., 1893 18

Hunt, Bruce J., *The Maxwellians*, Cornell History of Science, 2005 18

Jackson, J. D., Examples of the zeroth theorem of the history of science, American Journal of Physics, **76** (8), 704 (2008).. 26

Jackson, John David, *Classical Electrodynamics*, 2nd Edition, John Wiley and Sons, 1975 18

Jefimenko, Oleg, *Electricity and Magnetism*, 2nd Edition Electret Scientific Company, 1989 18

Maxwell, James Clerk, *A Treatise on Electricity and Magnetism,* Oxford, Clarendon Press, 1881.. 14

Maxwell, James Clerk, On Faraday's Lines of Force, Transactions of the Cambridge Philosophical Society, Vol X, Part I, 1855, as quoted in *Maxwell on the Electromagnetic Field* by Thomas K. Simpson, Rutgers University Press, p. 55 (1997) 18

Newton, Isaac, in a letter dated February 25, 1693, to Dr. Richard Bentley, Bishop of Worcester, and reprinted in I. B. Cohen, *Isaac Newton's Papers & Letters on Natural Philosophy*, Harvard University Press, 1978.. 14

Panofsky, Wolfgang and Melba Phillips, *Classical Electricity and Magnetism*, 2nd Edition, Addison-Wesley 1962... 18

Weinberg, Steven, in Conceptual Foudations of Quantum Field Theory, edited by Tian Yu Cao, Cambridge, 1999 15