28

Electromagnetic Radiation

28–1 Electromagnetism

The most dramatic moments in the development of physics are those in which great syntheses take place, where phenomena which previously had appeared to be different are suddenly discovered to be but different aspects of the same thing. The history of physics is the history of such syntheses, and the basis of the success of physical science is mainly that we are *able* to synthesize.

Perhaps the most dramatic moment in the development of physics during the 19th century occurred to J. C. Maxwell one day in the 1860's, when he combined the laws of electricity and magnetism with the laws of the behavior of light. As a result, the properties of light were partly unravelled—that old and subtle stuff that is so important and mysterious that it was felt necessary to arrange a special creation for it when writing Genesis. Maxwell could say, when he was finished with his discovery, "Let there be electricity and magnetism, and there is light!"

For this culminating moment there was a long preparation in the gradual discovery and unfolding of the laws of electricity and magnetism. This story we shall reserve for detailed study next year. However, the story is, briefly, as follows. The gradually discovered properties of electricity and magnetism, of electric forces of attraction and repulsion, and of magnetic forces, showed that although these forces were rather complex, they all fell off inversely as the square of the distance. We know, for example, that the simple coulomb law for stationary charges is that the electric force field varies inversely as the square of the distance. As a consequence, for sufficiently great distances there is very little influence of one system of charges on another. Maxwell noted that the equations or the laws that had been discovered up to this time were mutually inconsistent when he tried to put them all together, and in order for the whole system to be consistent, he had to add another term to his equations. With this new term there came an amazing prediction, which was that a part of the electric and magnetic fields would fall off much more slowly with the distance than the inverse square, namely, inversely as the first power of the distance! And so he realized that electric currents in one place can affect other charges far away, and he predicted the basic effects with which we are familiar today—radio transmission, radar, and so on.

It seems a miracle that someone talking in Europe can, with mere electrical influences, be heard thousands of miles away in Los Angeles. How is it possible? It is because the fields do not vary as the inverse square, but only inversely as the first power of the distance. Finally, then, even light itself was recognized to be electric and magnetic influences extending over vast distances, generated by an almost incredibly rapid oscillation of the electrons in the atoms. All these phenomena we summarize by the word *radiation* or, more specifically, *electromagnetic radiation*, there being one or two other kinds of radiation also. Almost always, radiation means electromagnetic radiation.

And thus is the universe knit together. The atomic motions of a distant star still have sufficient influence at this great distance to set the electrons in our eye in motion, and so we know about the stars. If this law did not exist, we would all be literally in the dark about the exterior world! And the electric surgings in a galaxy five billion light years away—which is the farthest object we have found so far—can still influence in a significant and detectable way the currents in the great "dish" in front of a radio telescope. And so it is that we see the stars and the galaxies.

This remarkable phenomenon is what we shall discuss in the present chapter. At the beginning of this course in physics we outlined a broad picture of the world, but we are now better prepared to understand some aspects of it, and so we shall now go over some parts of it again in greater detail. We begin by describing the position of physics at the end of the 19th century. All that was then known about the fundamental laws can be summarized as follows.

First, there were laws of forces: one force was the law of gravitation, which we have written down several times; the force on an object of mass m, due to another of mass M, is given by

$$\mathbf{F} = GmM\mathbf{e}_r/r^2, \tag{28.1}$$

where $\mathbf{e}_r$ is a unit vector directed from m to M, and r is the distance between them.

Next, the laws of electricity and magnetism, as known at the end of the 19th century, are these: the electrical forces acting on a charge q can be described by two fields, called $\mathbf{E}$ and $\mathbf{B}$, and the velocity $\mathbf{v}$ of the charge q, by the equation

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{28.2}$$

To complete this law, we have to say what the formulas for **E** and **B** are in a given circumstance: if a number of charges are present, **E** and the **B** are each the sum of contributions, one from each individual charge. So if we can find the **E** and **B** produced by a single charge, we need only to add all the effects from all the charges in the universe to get the total **E** and **B**! This is the principle of superposition.

What is the formula for the electric and magnetic field produced by one individual charge? It turns out that this is very complicated, and it takes a great deal of study and sophistication to appreciate it. But that is not the point. We write down the law now only to impress the reader with the beauty of nature, so to speak, i.e., that it is *possible* to summarize all the fundamental knowledge on one page, with notations that he is now familiar with. This law for the fields of an individual charge is complete and accurate, so far as we know (except for quantum mechanics) but it looks rather complicated. We shall not study all the pieces now; we only write it down to give an impression, to show that it can be written, and so that we can see ahead of time roughly what it looks like. As a matter of fact, the most *useful* way to write the correct laws of electricity and magnetism is not the way we shall now write them, but involves what are called *field equations*, which we shall learn about next year. But the mathematical notations for these are different and new, and so we write the law in an inconvenient form for calculation, but in notations that we now know.

The electric field, **E**, is given by

$$\mathbf{E} = \frac{-q}{4\pi\epsilon_0}\left[\frac{\mathbf{e}_{r'}}{r'^2} + \frac{r'}{c}\frac{d}{dt}\left(\frac{\mathbf{e}_{r'}}{r'^2}\right) + \frac{1}{c^2}\frac{d^2}{dt^2}\,\mathbf{e}_{r'}\right]. \tag{28.3}$$

What do the various terms tell us? Take the first term, $\mathbf{E} = -q\mathbf{e}_{r'}/4\pi\epsilon_0 r'^2$. That, of course, is Coulomb's law, which we already know: q is the charge that is producing the field; $\mathbf{e}_{r'}$ is the unit vector in the direction from the point P where **E** is measured, r is the distance from P to q. But, Coulomb's law is wrong. The discoveries of the 19th century showed that influences cannot travel faster than a certain fundamental speed c, which we now call the speed of light. It is not correct that the first term is Coulomb's law, not only because it is not possible to know where the charge is *now* and at what distance it is *now*, but also because the only thing that can affect the field at a given place and time is the behavior of the charges in the *past*. How *far* in the past? The time delay, or *retarded time*, so-called, is the time it takes, at speed c, to get from the charge to the field point P. The delay is r'/c.

So to allow for this time delay, we put a little prime on r, meaning how far away it *was* when the information now arriving at P left q. Just for a moment suppose that the charge carried a light, and that the light could only come to P at the speed c. Then when we look at q, we would not see where it is now, of course, but where it *was* at some earlier time. What appears in our formula is the *apparent* direction $\theta_{r'}$ —the direction it used to be—the so-called *retarded* direction—and at the *retarded* distance r'. That would be easy enough to understand, too, but it is also wrong. The whole thing is much more complicated.

There are several more terms. The next term is as though nature were trying to allow for the fact that the effect is retarded, if we might put it very crudely. It suggests that we should calculate the delayed coulomb field and add a correction to it, which is its rate of change times the time delay that we use. Nature seems to be attempting to guess what the field at the present time is going to be, by taking the rate of change and multiplying by the time that is delayed. But we are not yet through. There is a third term—the second derivative, with respect to t, of the unit vector in the direction of the charge. Now the formula *is* finished, and that is all there is to the electric field from an arbitrarily moving charge.

The magnetic field is given by

$$\mathbf{B} = -\mathbf{e}_{r'} \times \mathbf{E}/c. \tag{28.4}$$

We have written these down only for the purpose of showing the beauty of nature or, in a way, the power of mathematics. We do not pretend to understand *why* it is possible to write so much in such a small space, but (28.3) and (28.4) contain the machinery by which electric generators work, how light operates, all the phenomena of electricity and magnetism. Of course, to complete the story we also need to know something about the behavior of the materials involved—the properties of matter—which are not described properly by (28.3).

To finish with our description of the world of the 19th century we must mention one other great synthesis which occurred in that century, one with which Maxwell had a great deal to do also, and that was the synthesis of the phenomena of heat and mechanics. We shall study that subject soon.

What had to be added in the 20th century was that the dynamical laws of Newton were found to be all wrong, and quantum mechanics had to be introduced to correct them. Newton's laws are approximately valid when the scale of things is sufficiently large. These quantum-mechanical laws, combined with the laws of electricity, have only recently been combined to form a set of laws called *quantum electrodynamics*. In addition, there were discovered a number of new phenomena, of which the first was radioactivity, discovered by Becquerel in 1898—he just sneaked it in under the 19th century. This phenomenon of radioactivity was followed up to produce our knowledge of nuclei and new kinds of forces that are not gravitational and not electrical, but new particles with different interactions, a subject which has still not been unravelled.

For those purists who know more (the professors who happen to be reading this), we should add that when we say that (28.3) is a complete expression of the knowledge of electrodynamics, we are not being entirely accurate. There was a problem that was not quite solved at the end of the 19th century. When we try to calculate the field from all the charges *including the charge itself that we want the field to act on*, we get into trouble trying to find the distance, for example, of a charge from itself, and dividing something by that distance, which is zero. The problem of how to handle the part of this field which is generated by the very charge on which we want the field to act is not yet solved today. So we leave it there; we do not have a complete solution to that puzzle yet, and so we shall avoid the puzzle for as long as we can.

28–2 Radiation

That, then, is a summary of the world picture. Now let us use it to discuss the phenomena called radiation. To discuss these phenomena, we must select from Eq. (28.3) only that piece which varies inversely as the distance and not as the square of the distance. It turns out that when we finally do find that piece, it is so simple in its form that it is legitimate to study optics and electrodynamics in an elementary way by taking it as "the law" of the electric field produced by a moving charge far away. We shall take it temporarily as a given law which we will learn about in detail next year.

Of the terms appearing in (28.3), the first one evidently goes inversely as the square of the distance, and the second is only a correction for delay, so it is easy to show that both of them vary inversely as the square of the distance. All of the effects we are interested in come from the third term, which is not very complicated, after all. What this term says is: look at the charge and note the direction of the unit vector (we can project the end of it onto the surface of a unit sphere). As the charge moves around, the unit vector wiggles, and *the acceleration of that unit vector is what we are looking for.* That is all. Thus

$$\mathbf{E} = \frac{-q}{4\pi\epsilon_0 c^2} \frac{d^2\mathbf{e}_{r'}}{dt^2} \tag{28.5}$$

is a statement of the laws of radiation, because that is the only important term when we get far enough away that the fields are varying inversely as the distance. (The parts that go as the square have fallen off so much that we are not interested in them.)

Now we can go a little bit further in studying (28.5) to see what it means. Suppose a charge is moving in any manner whatsoever, and we are observing it from a distance. We imagine for a moment that in a sense it is "lit up" (although it is light that we are trying to explain); we imagine it as a little white dot. Then we would see this white dot running around. But we don't see *exactly* how it is running around right *now*, because of the delay that we have been talking about. What counts is how it was moving *earlier*. The unit vector $\mathbf{e}_{r'}$ is pointed toward the apparent position of the charge. Of course, the end of $\mathbf{e}_{r'}$ goes on a slight curve, so that its acceleration has two components. One is the transverse piece, because the end of it goes up and down, and the other is a radial piece because it stays on a sphere. It is easy to demonstrate that the latter is much smaller and varies as the inverse square of r when r is very great. This is easy to see, for when we imagine that we move a given source farther and farther away, then the wigglings of $\mathbf{e}_{r'}$ look smaller and smaller, inversely as the distance, but the radial component of acceleration is varying much more rapidly than inversely as the distance. So for practical purposes all we have to do is project the motion on a plane at unit distance. Therefore we find the following rule: Imagine that we look at the moving charge and that everything we see is delayed—like a painter trying to paint a scene on a screen at a unit distance. A real painter, of course, does not take into account the fact that light is going at a certain speed, but paints the world as he sees it. We want to see what his picture would look like. So we see a dot, representing the charge, moving about in the picture. The acceleration of that dot is proportional to the electric field. That is all—all we need.

Thus Eq. (28.5) is the complete and correct formula for radiation; even relativity effects are all contained in it. However, we often want to apply it to a still simpler circumstance in which the charges are moving only a small distance at a relatively slow rate. Since they are moving slowly, they do not move an appreciable distance from where they start, so that the delay time is practically constant. Then the law is still simpler, because the delay time is fixed. Thus we imagine that the charge is executing a very tiny motion at an effectively constant distance. The delay at the distance r is r/c. Then our rule becomes the following: If the charged object is moving in a very small motion and it is laterally displaced by the distance $x(t)$, then the angle that the unit vector $\mathbf{e}_{r'}$ is displaced is x/r, and since r is practically constant, the x-component of $d^2\mathbf{e}_{r'}/dt^2$ is simply the acceleration of x itself at an earlier time, and so finally we get the law we want, which is

$$E_x(t) = \frac{-q}{4\pi\epsilon_0 c^2 r} a_x\left(t - \frac{r}{c}\right). \tag{28.6}$$

Only the component of a_x perpendicular to the line of sight is important. Let us see why that is. Evidently, if the charge is moving in and out straight at us, the unit vector in that direction does not wiggle at all, and it has no acceleration. So it is only the sidewise motion which is important, only the acceleration that we see projected on the screen.

28-3 The dipole radiator

As our fundamental "law" of electromagnetic radiation, we are going to assume that (28.6) is true, i.e., that the electric field produced by an accelerating charge which is moving nonrelativistically at a very large distance r approaches that form. The electric field varies inversely as r and is proportional to the acceleration of the charge, projected onto the "plane of sight," and this acceleration is not today's acceleration, but the acceleration that it had at an earlier time, the amount of delay being a time, r/c. In the remainder of this chapter we shall discuss this law so that we can understand it better physically, because we are going to use it to understand all of the phenomena of light and radio propagation, such as reflection, refraction, interference, diffraction, and scattering. It is the central law, and is all we need. All the rest of Eq. (28.3) was written down only to set the stage, so that we could appreciate where (28.6) fits and how it comes about.

We shall discuss (28.3) further next year. In the meantime, we shall accept it as true, but not just on a theoretical basis. We may devise a number of experiments which illustrate the character of the law. In order to do so, we need an accelerating charge. It should be a single charge, but if we can make a great many charges move together, all the same way, we know that the field will be the sum of the effects of each of the individual charges; we just add them together. As an example, consider two pieces of wire connected to a generator, as shown in Fig. 28–1. The idea is that the generator makes a potential difference, or a field, which pulls electrons away from piece A and pushes them into B at one moment, and then, an infinitesimal time later, it reverses the effect and pulls the electrons out of B and pumps them back into A! So in these two wires charges, let us say, are accelerating upward in wire A and upward in wire B for one moment, and a moment later they are accelerating downward in wire A and downward in wire B. The fact that we need two wires and a generator is merely that this is a way of doing it. The net result is that we merely have a charge accelerating up and down as though A and B were one single wire. A wire that is very short compared with the distance light travels in one oscillation period is called an *electric dipole oscillator*. Thus we have the circumstance that we need to apply our law, which tells us that this charge makes an electric field, and so we need an instrument to detect an electric field, and the instrument we use is the same thing—a pair of wires like A and B! If an electric field is applied to such a device, it will produce a force which will pull the electrons up on both wires or down on both wires. This signal is detected by means of a rectifier mounted between A and B, and a tiny, fine wire carries the information into an amplifier, where it is amplified so we can hear the audio-frequency tone with which the radiofrequency is modulated. When this probe feels an electric field, there will be a loud noise coming out of the loudspeaker, and when there is no electric field driving it, there will be no noise.

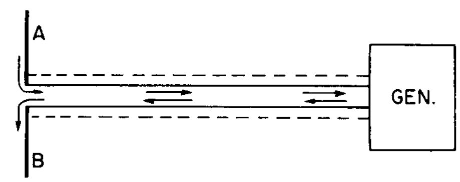

Fig. 28–1. A high-frequency signal generator drives charges up and down on two wires.

Because the room in which the waves we are measuring has other objects in it, our electric field will shake electrons in these other objects; the electric field makes these other charges go up and down, and in going up and down, these also produce an effect on our probe. Thus for a successful experiment we must hold things fairly close together, so that the influences from the walls and from ourselves—the reflected waves—are relatively small. So the phenomena will not turn out to appear to be precisely and perfectly in accord with Eq. (28.6), but will be close enough that we shall be able to appreciate the law.

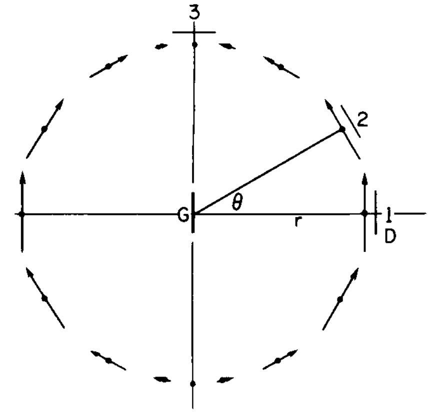

Fig. 28–2. The instantaneous electric field on a sphere centered at a localized, linearly oscillating charge.

Now we turn the generator on and hear the audio signal. We find a strong field when the detector D is parallel to the generator G at point 1 (Fig. 28–2). We find the same amount of field also at any other azimuth angle about the axis of G, because it has no directional effects. On the other hand, when the detector is at 3 the field is zero. That is all right, because our formula said that the field should be the acceleration of the charge *projected perpendicular* to the line of sight. Therefore when we look down on G, the charge is moving toward and away from D, and there is no effect. So that checks the first rule, that there is no effect when the charge is moving directly toward us. Secondly, the formula says that the electric field should be perpendicular to r and in the plane of G and r; so if we put D at 1 but rotate it 90°, we should get no signal. And this is just what we find, the electric field is indeed vertical, and not horizontal. When we move D to some intermediate angle, we see that the strongest signal occurs when it is oriented as shown, because although G is vertical, it does not produce a field that is simply parallel to itself—it is the *projection of the acceleration perpendicular to the line of sight* that counts. The signal is weaker at 2 than it is at 1, because of the projection effect.

28–4 Interference

Next, we may test what happens when we have two sources side by side a few centimeters apart (Fig. 28–3). The law is that the two sources should add their effects at point 1 when both of the sources are connected to the same generator and are both moving up and down the same way, so that the total electric field is the sum of the two and is twice as strong as it was before.

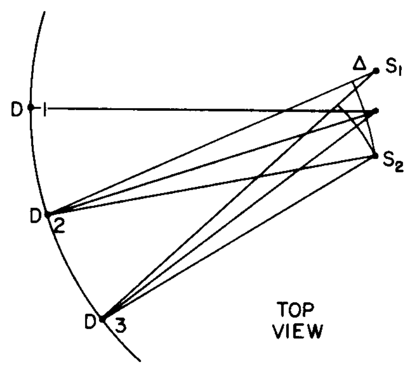

Fig. 28–3. Illustration of interference of sources.

Now comes an interesting possibility. Suppose we make the charges in S_1 and S_2 both accelerate up and down, but delay the timing of S_2 so that they are 180° out of phase. Then the field produced by S_1 will be in one direction and the field produced by S_2 will be in the opposite direction at any instant, and therefore we should get *no* effect at point 1. The phase of oscillation is neatly adjustable by means of a pipe which is carrying the signal to S_2. By changing the length of this pipe we change the time it takes the signal to arrive at S_2 and thus we change the phase of that oscillation. By adjusting this length, we can indeed find a place where there is no more signal left, in spite of the fact that both S_1 and S_2 are moving! The fact that they are both moving can be checked, because if we cut one out, we can see the motion of the other. So the two of them together can produce *zero* if everything is adjusted correctly.

Now, it is very interesting to show that the addition of the two fields is in fact a *vector* addition. We have just checked it for up and down motion, but let us check two nonparallel directions. First, we restore S_1 and S_2 to the same phase; that is, they are again moving together. But now we turn S_1 through 90°, as shown in Fig. 28–4. Now we should have at point 1 the sum of two effects, one of which is vertical and the other horizontal. The electric field is the vector sum of these two in-phase signals—they are both strong at the same time and go through zero together; the total field should be a signal R at 45°. If we turn D to get the maximum noise, it should be at about 45°, and not vertical. And if we turn it at right angles to that direction, we should get zero, which is easy to measure. Indeed, we observe just such behavior!

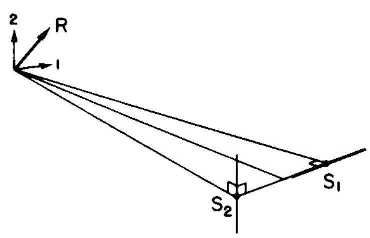

Fig. 28–4. Illustration of the vector character of the combination of sources.

Now, how about the retardation? How can we demonstrate that the signal is retarded? We could, with a great deal of equipment, measure the time at which it arrives, but there is another, very simple way. Referring again to Fig. 28–3, suppose that S_1 and S_2 are in phase. They are both shaking together, and they produce equal electric fields at point 1. But suppose we go to a certain place 2 which is closer to S_2 and farther from S_1. Then, in accordance with the principle that the acceleration should be retarded by an amount equal to r/c, if the retardations are not equal, the signals are no longer in phase. Thus it should be possible to find a position at which the distances of D from S_1 and S_2 differ by some amount Δ, in such a manner that there is no net signal. That is, the distance Δ is to be the distance light goes in one-half an oscillation of the generator. We may go still further, and find a point where the difference is greater by a whole cycle; that is to say, the signal from the first antenna reaches point 3 with a delay in time that is greater than that of the second antenna by just the length of time it takes for the electric current to oscillate once, and therefore the two electric fields produced at 3 are in phase again. At point 3 the signal is strong again.

This completes our discussion of the experimental verification of some of the important features of Eq. (28.6). Of course we have not really checked the $1/r$ variation of the electric field strength, or the fact that there is also a magnetic field that goes along with the electric field. To do so would require rather sophisticated techniques and would hardly add to our understanding at this point. In any case, we have checked those features that are of the greatest importance for our later applications, and we shall come back to study some of the other properties of electromagnetic waves next year.