

## Waveguides

The Feynman lectures on physics, by Richard  
Feynman. Addison-Wesley publishing co. Mass

NOTICE  
THIS MATERIAL MAY BE  
PROTECTED BY COPYRIGHT  
LAW (TITLE 17 U.S. CODE)

PROPERTY Dine

### 24-1 The transmission line

In the last chapter we studied what happened to the lumped elements of circuits when they were operated at very high frequencies, and we were led to see that a resonant circuit could be replaced by a cavity with the fields resonating inside. Another interesting technical problem is the connection of one object to another, so that electromagnetic energy can be transmitted between them. In low-frequency circuits the connection is made with wires, but this method doesn't work very well at high frequencies because the circuits would radiate energy into all the space around them, and it is hard to control where the energy will go. The fields spread out around the wires; the currents and voltages are not "guided" very well by the wires. In this chapter we want to look into the ways that objects can be interconnected at high frequencies. At least, that's one way of presenting our subject.

Another way is to say that we have been discussing the behavior of waves in free space. Now it is time to see what happens when oscillating fields are confined in one or more dimensions. We will discover the interesting new phenomenon when the fields are confined in only two dimensions and allowed to go free in the third dimension, they propagate in waves. These are "guided waves"—the subject of this chapter.

We begin by working out the general theory of the *transmission line*. The ordinary power transmission line that runs from tower to tower over the countryside radiates away some of its power, but the power frequencies (50–60 cycles/sec) are so low that this loss is not serious. The radiation could be stopped by surrounding the line with a metal pipe, but this method would not be practical for power lines because the voltages and currents used would require a very large, expensive, and heavy pipe. So simple "open lines" are used.

For somewhat higher frequencies—say a few kilocycles—radiation can already be serious. However, it can be reduced by using "twisted-pair" transmission lines, as is done for short-run telephone connections. At higher frequencies, however, the radiation soon becomes intolerable, either because of power losses or because the energy appears in other circuits where it isn't wanted. For frequencies from a few kilocycles to some hundreds of megacycles, electromagnetic signals and power are usually transmitted via coaxial lines consisting of a wire inside a cylindrical "outer conductor" or "shield." Although the following treatment will apply to a transmission line of two parallel conductors of any shape, we will carry it out referring to a coaxial line.

We take the simplest coaxial line that has a central conductor, which we suppose is a thin hollow cylinder, and an outer conductor which is another thin cylinder on the same axis as the inner conductor, as in Fig. 24-1. We begin by figuring out approximately how the line behaves at relatively low frequencies. We have already described some of the low-frequency behavior when we said earlier that two such conductors had a certain amount of inductance per unit length or a certain capacity per unit length. We can, in fact, describe the low-frequency behavior of any transmission line by giving its inductance per unit length,  $L_0$  and its capacity per unit length,  $C_0$ . Then we can analyze the line as the limiting case of the  $L$ - $C$  filter as discussed in Section 22-6. We can make a filter which imitates the line by taking small series elements  $L_0 \Delta x$  and small shunt capacities  $C_0 \Delta x$ , where  $\Delta x$  is an element of length of the line. Using our results for the infinite filter, we see that there would be a propagation of electric

### 24-1 The transmission line

### 24-2 The rectangular waveguide

### 24-3 The cutoff frequency

### 24-4 The speed of the guided waves

### 24-5 Observing guided waves

### 24-6 Waveguide plumbing

### 24-7 Waveguide modes

### 24-8 Another way of looking at the guided waves

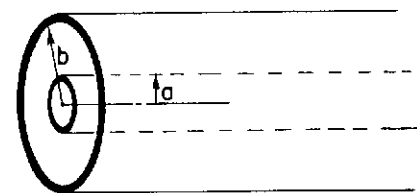


Fig. 24-1. A coaxial transmission line.

signals along the line. Rather than following that approach, however, we would now rather look at the line from the point of view of a differential equation.

Suppose that we see what happens at two neighboring points along the transmission line, say at the distances  $x$  and  $x + \Delta x$  from the beginning of the line. Let's call the voltage difference between the two conductors  $V(x)$ , and the current along the "hot" conductor  $I(x)$  (see Fig. 24-2). If the current in the line is varying, the inductance will give us a voltage drop across the small section of line from  $x$  to  $x + \Delta x$  in the amount

$$\Delta V = V(x + \Delta x) - V(x) = -L_0 \Delta x \frac{dI}{dt}.$$

Or, taking the limit as  $\Delta x \rightarrow 0$ , we get

$$\frac{\partial V}{\partial x} = -L_0 \frac{\partial I}{\partial t}. \quad (24.1)$$

The changing current gives a gradient of the voltage.

Referring again to the figure, if the voltage at  $x$  is changing, there must be some charge supplied to the capacity in that region. If we take the small piece of line between  $x$  and  $x + \Delta x$ , the charge on it is  $q = C_0 \Delta x V$ . The time rate-of-change of this charge is  $C_0 \Delta x dV/dt$ , but the charge changes only if the current  $I(x)$  into the element is different from the current  $I(x + \Delta x)$  out. Calling the difference  $\Delta I$ , we have

$$\Delta I = -C_0 \Delta x \frac{dV}{dt}.$$

Taking the limit as  $\Delta x \rightarrow 0$ , we get

$$\frac{\partial I}{\partial x} = -C_0 \frac{\partial V}{\partial t}. \quad (24.2)$$

So the conservation of charge implies that the gradient of the current is proportional to the time rate-of-change of the voltage.

Equations (24.1) and (24.2) are then the basic equations of a transmission line. If we wish, we could modify them to include the effects of resistance in the conductors or of leakage of charge through the insulation between the conductors, but for our present discussion we will just stay with the simple example.

The two transmission line equations can be combined by differentiating one with respect to  $t$  and the other with respect to  $x$  and eliminating either  $V$  or  $I$ . Then we have either

$$\frac{\partial^2 V}{\partial x^2} = C_0 L_0 \frac{\partial^2 V}{\partial t^2} \quad (24.3)$$

or

$$\frac{\partial^2 I}{\partial x^2} = C_0 L_0 \frac{\partial^2 I}{\partial t^2}. \quad (24.4)$$

Once more we recognize the wave equation in  $x$ . For a uniform transmission line, the voltage (and current) propagates along the line as a wave. The voltage along the line must be of the form  $V(x, t) = f(x - vt)$  or  $V(x, t) = g(x + vt)$ , or a sum of both. Now what is the velocity  $v$ ? We know that the coefficient of the  $\partial^2/\partial t^2$  term is just  $1/v^2$ , so

$$v = \frac{1}{\sqrt{L_0 C_0}}. \quad (24.5)$$

We will leave it for you to show that the voltage *for each wave* in a line is proportional to the current of that wave and that the constant of proportionality is just the characteristic impedance  $z_0$ . Calling  $V_+$  and  $I_+$  the voltage and current for a wave going in the plus  $x$ -direction, you should get

$$V_+ = z_0 I_+. \quad (24.6)$$

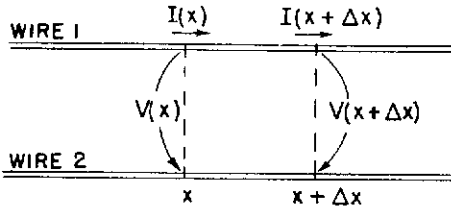


Fig. 24-2. The currents and voltages of a transmission line.

Similary, for the wave going toward minus  $x$  the relation is

$$V_- = -z_0 I_-.$$

The characteristic impedance—as we found out from our filter equations—is given by

$$z_0 = \sqrt{\frac{L_0}{C_0}}, \quad (24.7)$$

and is, therefore, a pure resistance.

To find the propagation speed  $v$  and the characteristic impedance  $z_0$  of a transmission line, we have to know the inductance and capacity per unit length. We can calculate them easily for a coaxial cable, so we will see how that goes. For the inductance we follow the ideas of Section 17-8, and set  $\frac{1}{2}LI^2$  equal to the magnetic energy which we get by integrating  $\epsilon_0 c^2 B^2/2$  over the volume. Suppose that the central conductor carries the current  $I$ ; then we know that  $B = I/2\pi\epsilon_0 c^2 r$ , where  $r$  is the distance from the axis. Taking as a volume element a cylindrical shell of thickness  $dr$  and of length  $l$ , we have for the magnetic energy

$$U = \frac{\epsilon_0 c^2}{2} \int_a^b \left( \frac{I}{2\pi\epsilon_0 c^2 r} \right)^2 l 2\pi r dr,$$

where  $a$  and  $b$  are the radii of the inner and outer conductors, respectively. Carrying out the integral, we get

$$U = \frac{I^2 l}{4\pi\epsilon_0 c^2} \ln \frac{b}{a}. \quad (24.8)$$

Setting the energy equal to  $\frac{1}{2}LI^2$ , we find

$$L = \frac{l}{2\pi\epsilon_0 c^2} \ln \frac{b}{a}. \quad (24.9)$$

It is, as it should be, proportional to the length  $l$  of the line, so the inductance per unit length  $L_0$  is

$$L_0 = \frac{\ln(b/a)}{2\pi\epsilon_0 c^2}. \quad (24.10)$$

We have worked out the charge on a cylindrical condenser (see Section 12-2). Now, dividing the charge by the potential difference, we get

$$C = \frac{2\pi\epsilon_0 l}{\ln(b/a)}.$$

The capacity per unit length  $C_0$  is  $C/l$ . Combining this result with Eq. (24.10), we see that the product  $L_0 C_0$  is just equal to  $1/c^2$ , so  $v = 1/\sqrt{L_0 C_0}$  is equal to  $c$ . The wave travels down the line with the speed of light. We point out that this result depends on our assumptions: (a) that there are no dielectrics or magnetic materials in the space between the conductors, and (b) that the currents are all on the surfaces of the conductors (as they would be for perfect conductors). We will see later that for good conductors at high frequencies, all currents distribute themselves on the surfaces as they would for a perfect conductor, so this assumption is then valid.

Now it is interesting that so long as assumptions (a) and (b) are correct, the product  $L_0 C_0$  is equal to  $1/c^2$  for *any* parallel pair of conductors—even, say, for a hexagonal inner conductor anywhere inside an elliptical outer conductor. So long as the cross section is constant and the space between has no material, waves are propagated at the velocity of light.

No such general statement can be made about the characteristic impedance. For the coaxial line, it is

$$z_0 = \frac{\ln(b/a)}{2\pi\epsilon_0 c}. \quad (24.11)$$

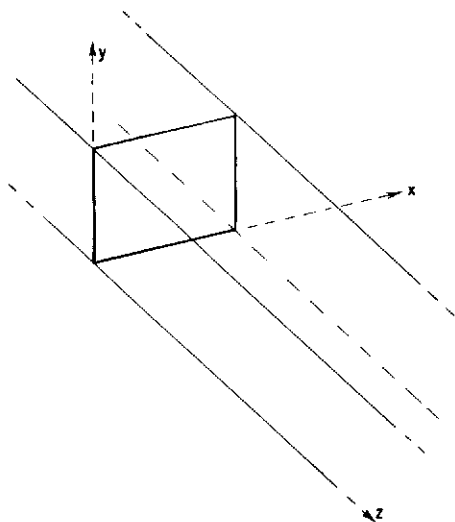


Fig. 24-3. Coordinates chosen for the rectangular waveguide.

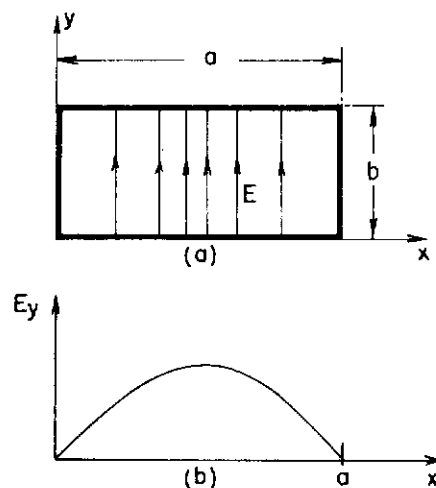


Fig. 24-4. The electric field in the waveguide at some value of  $z$ .

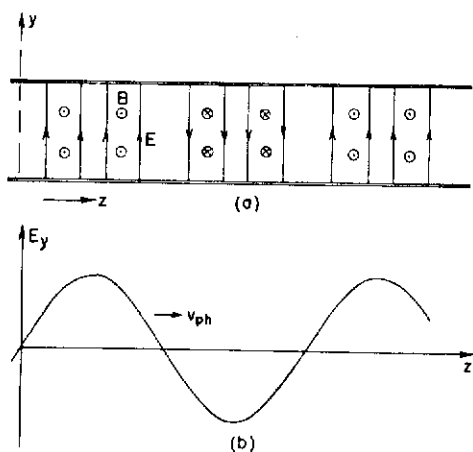


Fig. 24-5. The  $z$ -dependence of the field in the waveguide.

The factor  $1/\epsilon_0 c$  has the dimensions of a resistance and is equal to  $120\pi$  ohms. The geometric factor  $\ln(b/a)$  depends only logarithmically on the dimensions, so for the coaxial line—and most lines—the characteristic impedance has typical values of from 50 ohms or so to a few hundred ohms.

## 24-2 The rectangular waveguide

The next thing we want to talk about seems, at first sight, to be a striking phenomenon: if the central conductor is removed from the coaxial line, it can still carry electromagnetic power. In other words, at high enough frequencies a hollow tube will work just as well as one with wires. It is related to the mysterious way in which a resonant circuit of a condenser and inductance gets replaced by nothing but a can at high frequencies.

Although it may seem to be a remarkable thing when one has been thinking in terms of a transmission line as a distributed inductance and capacity, we all know that electromagnetic waves can travel along inside a hollow metal pipe. If the pipe is straight, we can see through it! So certainly electromagnetic waves go through a pipe. But we also know that it is not possible to transmit low-frequency waves (power or telephone) through the inside of a single metal pipe. So it must be that electromagnetic waves will go through if their wavelength is short enough. Therefore we want to discuss the limiting case of the longest wavelength (or the lowest frequency) that can get through a pipe of a given size. Since the pipe is then being used to carry waves, it is called a *waveguide*.

We will begin with a rectangular pipe, because it is the simplest case to analyze. We will first give a mathematical treatment and come back later to look at the problem in a much more elementary way. The more elementary approach, however, can be applied easily only to a rectangular guide. The basic phenomena are the same for a general guide of arbitrary shape, so the mathematical argument is fundamentally more sound.

Our problem, then, is to find what kind of waves can exist inside a rectangular pipe. Let's first choose some convenient coordinates; we take the  $z$ -axis along the length of the pipe, and the  $x$ - and  $y$ -axes parallel to the two sides, as shown in Fig. 24-3.

We know that when light waves go down the pipe, they have a transverse electric field; so suppose we look first for solutions in which  $E$  is perpendicular to  $z$ , say with only a  $y$ -component,  $E_y$ . This electric field will have some variation across the guide; in fact, it must go to zero at the sides parallel to the  $y$ -axis, because the currents and charges in a conductor always adjust themselves so that there is no tangential component of the electric field at the surface of a conductor. So  $E_y$  will vary with  $x$  in some arch, as shown in Fig. 24-4. Perhaps it is the Bessel function we found for a cavity? No, because the Bessel function has to do with cylindrical geometries. For a rectangular geometry, waves are usually simple harmonic functions, so we should try something like  $\sin k_x x$ .

Since we want waves that propagate down the guide, we expect the field to alternate between positive and negative values as we go along in  $z$ , as in Fig. 24-5, and these oscillations will travel along the guide with some velocity  $v$ . If we have oscillations at some definite frequency  $\omega$ , we would guess that the wave might vary with  $z$  like  $\cos(\omega t - k_z z)$ , or to use the more convenient mathematical form, like  $e^{i(\omega t - k_z z)}$ . This  $z$ -dependence represents a wave travelling with the speed  $v = \omega/k_z$  (see Chapter 29, Vol. I).

So we might guess that the wave in the guide would have the following mathematical form:

$$E_y = E_0 \sin k_x x e^{i(\omega t - k_z z)}. \quad (24.12)$$

Let's see whether this guess satisfies the correct field equations. First, the electric field should have no tangential components at the conductors. Our field satisfies this requirement; it is perpendicular to the top and bottom faces and is zero at the two side faces. Well, it is if we choose  $k_x$  so that one-half a cycle of

$\sin k_x x$  just fits in the width of the guide—that is, if

$$k_x a = \pi. \quad (24.13)$$

There are other possibilities, like  $k_x a = 2\pi, 3\pi, \dots$ , or, in general,

$$k_x a = n\pi, \quad (24.14)$$

where  $n$  is any integer. These represent various complicated arrangements of the field, but for now let's take only the simplest one, where  $k_x = \pi/a$ , where  $a$  is the width of the inside of the guide.

Next, the divergence of  $\mathbf{E}$  must be zero in the free space inside the guide, since there are no charges there. Our  $\mathbf{E}$  has only a  $y$ -component, and it doesn't change with  $y$ , so we do have that  $\nabla \cdot \mathbf{E} = 0$ .

Finally, our electric field must agree with the rest of Maxwell's equations in the free space inside the guide. That is the same thing as saying that it must satisfy the wave equation

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial y^2} + \frac{\partial^2 E_y}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 E_y}{\partial t^2} = 0. \quad (24.15)$$

We have to see whether our guess, Eq. (24.12), will work. The second derivative of  $E_y$  with respect to  $x$  is just  $-k_x^2 E_y$ . The second derivative with respect to  $y$  is zero, since nothing depends on  $y$ . The second derivative with respect to  $z$  is  $-k_z^2 E_y$ , and the second derivative with respect to  $t$  is  $-\omega^2 E_y$ . Equation (24.15) then says that

$$k_x^2 E_y + k_z^2 E_y - \frac{\omega^2}{c^2} E_y = 0.$$

Unless  $E_y$  is zero everywhere (which is not very interesting), this equation is correct if

$$k_x^2 + k_z^2 - \frac{\omega^2}{c^2} = 0. \quad (24.16)$$

We have already fixed  $k_x$ , so this equation tells us that there can be waves of the type we have assumed if  $k_z$  is related to the frequency  $\omega$  so that Eq. (24.16) is satisfied—in other words, if

$$k_z = \sqrt{(\omega^2/c^2) - (\pi^2/a^2)}. \quad (24.17)$$

The waves we have described are propagated in the  $z$ -direction with this value of  $k_z$ .

The wave number  $k_z$  we get from Eq. (24.17) tells us, for a given frequency  $\omega$ , the speed with which the nodes of the wave propagate down the guide. The phase velocity is

$$v = \frac{\omega}{k_z}. \quad (24.18)$$

You will remember that the wavelength  $\lambda$  of a travelling wave is given by  $\lambda = 2\pi v/\omega$ , so  $k_z$  is also equal to  $2\pi/\lambda_g$ , where  $\lambda_g$  is the wavelength of the oscillations along the  $z$ -direction—the “guide wavelength.” The wavelength in the guide is different, of course, from the free-space wavelength of electromagnetic waves of the same frequency. If we call the free-space wavelength  $\lambda_0$ , which is equal to  $2\pi c/\omega$ , we can write Eq. (24.17) as

$$\lambda_g = \frac{\lambda_0}{\sqrt{1 - (\lambda_0/2a)^2}}. \quad (24.19)$$

Besides the electric fields there are magnetic fields that will travel with the wave, but we will not bother to work out an expression for them right now. Since  $c^2 \nabla \times \mathbf{B} = \partial \mathbf{E} / \partial t$ , the lines of  $\mathbf{B}$  will circulate around the regions in which  $\partial \mathbf{E} / \partial t$  is largest, that is, halfway between the maximum and minimum of  $\mathbf{E}$ . The loops of  $\mathbf{B}$  will lie parallel to the  $xz$ -plane and between the crests and troughs of  $\mathbf{E}$ , as shown in Fig. 24-6.

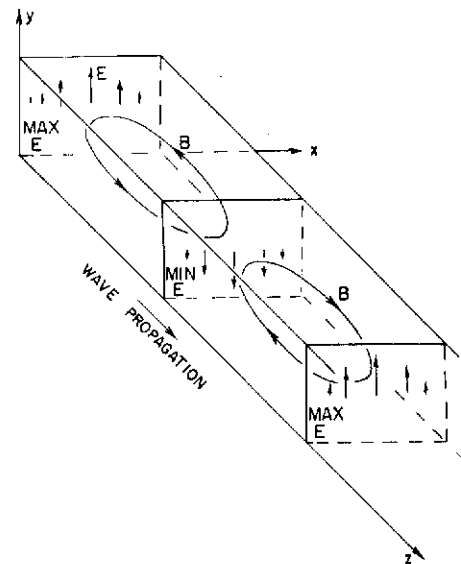


Fig. 24-6. The magnetic field in the waveguide.

### 24-3 The cutoff frequency

In solving Eq. (24.16) for  $k_z$ , there should really be two roots—one plus and one minus. We should write

$$k_z = \pm \sqrt{(\omega^2/c^2) - (\pi^2/a^2)}. \quad (24.20)$$

The two signs simply mean that there can be waves which propagate with a negative phase velocity (toward  $-z$ ), as well as waves which propagate in the positive direction in the guide. Naturally, it should be possible for waves to go in either direction. Since both types of waves can be present at the same time, there will be the possibility of standing-wave solutions.

Our equation for  $k_z$  also tells us that higher frequencies give larger values of  $k_z$ , and therefore smaller wavelengths, until in the limit of large  $\omega$ ,  $k$  becomes equal to  $\omega/c$ , which is the value we would expect for waves in free space. The light we “see” through a pipe still travels at the speed  $c$ . But now notice that if we go toward low frequencies, something strange happens. At first the wavelength gets longer and longer, but if  $\omega$  gets too small the quantity inside the square root of Eq. (24.20) suddenly becomes negative. This will happen as soon as  $\omega$  gets to be less than  $\pi c/a$ —or when  $\lambda_0$  becomes greater than  $2a$ . In other words, when the frequency gets smaller than a certain critical frequency  $\omega_c = \pi c/a$ , the wave number  $k_z$  (and also  $\lambda_g$ ) becomes imaginary and we haven’t got a solution any more. Or do we? Who said that  $k_z$  has to be real? What if it does come out imaginary? Our field equations are still satisfied. Perhaps an imaginary  $k_z$  also represents a wave.

Suppose  $\omega$  is less than  $\omega_c$ ; then we can write

$$k_z = \pm ik', \quad (24.21)$$

where  $k'$  is a positive real number:

$$k' = \sqrt{(\pi^2/a^2) - (\omega^2/c^2)}. \quad (24.22)$$

If we now go back to our expression, Eq. (24.12), for  $E_y$ , we have

$$E_y = E_0 \sin k_x x e^{i(\omega t \mp ik' z)}, \quad (24.23)$$

which we can write as

$$E_y = E_0 \sin k_x x e^{\pm k' z} e^{i\omega t}. \quad (24.24)$$

This expression gives an  $E$ -field that oscillates with time as  $e^{i\omega t}$  but which varies with  $z$  as  $e^{\pm k' z}$ . It decreases or increases with  $z$  smoothly as a real exponential. In our derivation we didn’t worry about the sources that started the waves, but there must, of course, be a source someplace in the guide. The sign that goes with  $k'$  must be the one that makes the field decrease with increasing distance from the source of the waves.

So for frequencies below  $\omega_c = \pi c/a$ , waves do *not* propagate down the guide; the oscillating fields penetrate into the guide only a distance of the order of  $1/k'$ . For this reason, the frequency  $\omega_c$  is called the “cutoff frequency” of the guide. Looking at Eq. (24.22), we see that for frequencies just a little below  $\omega_c$ , the number  $k'$  is small and the fields can penetrate a long distance into the guide. But if  $\omega$  is much less than  $\omega_c$ , the exponential coefficient  $k'$  is equal to  $\pi/a$  and the field dies off extremely rapidly, as shown in Fig. 24-7. The field decreases by  $1/e$  in the distance  $a/\pi$ , or in only about one-third of the guide width. The fields penetrate very little distance from the source.

We want to emphasize an interesting feature of our analysis of the guided waves—the appearance of the imaginary wave number  $k_z$ . Normally, if we solve an equation in physics and get an imaginary number, it doesn’t mean anything physical. For *waves*, however, an imaginary wave number *does* mean something. The wave equation is still satisfied; it only means that the solution gives exponentially decreasing fields instead of propagating waves. So in any wave problem where  $k$  becomes imaginary for some frequency, it means that the form of the wave changes—the sine wave changes into an exponential.

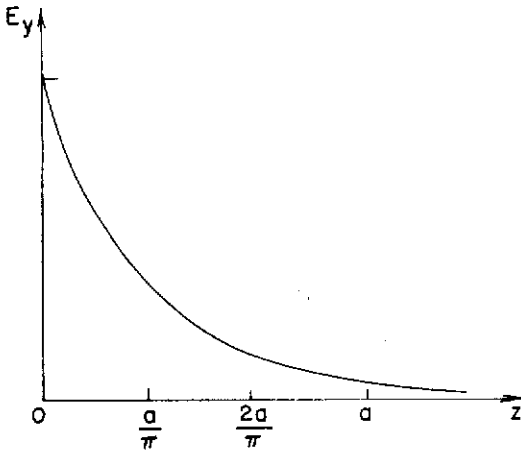


Fig. 24-7. The variation of  $E_y$  with  $z$  for  $\omega \ll \omega_c$ .

#### 24-4 The speed of the guided waves

The wave velocity we have used above is the phase velocity, which is the speed of a node of the wave; it is a function of frequency. If we combine Eqs. (24.17) and (24.18), we can write

$$v_{\text{phase}} = \frac{c}{\sqrt{1 - (\omega_c/\omega)^2}}. \quad (24.25)$$

For frequencies above cutoff—where travelling waves exist— $\omega_c/\omega$  is less than one, and  $v_{\text{phase}}$  is real and *greater than* the speed of light. We have already seen in Chapter 48 of Vol. I that *phase* velocities greater than light are possible, because it is just the nodes of the wave which are moving and not energy or information. In order to know how fast *signals* will travel, we have to calculate the speed of pulses or modulations made by the interference of a wave of one frequency with one or more waves of slightly different frequencies (see Chapter 48, Vol. I). We have called the speed of the envelope of such a group of waves the group velocity; it is not  $\omega/k$  but  $d\omega/dk$ :

$$v_{\text{group}} = \frac{d\omega}{dk}. \quad (24.26)$$

Taking the derivative of Eq. (24.17) with respect to  $\omega$  and inverting to get  $d\omega/dk$ , we find that

$$v_{\text{group}} = c \sqrt{1 - (\omega_c/\omega)^2}, \quad (24.27)$$

which is less than the speed of light.

The geometric mean of  $v_{\text{phase}}$  and  $v_{\text{group}}$  is just  $c$ , the speed of light:

$$v_{\text{phase}} v_{\text{group}} = c^2. \quad (24.28)$$

This is curious, because we have seen a similar relation in quantum mechanics. For a particle with any velocity—even relativistic—the momentum  $p$  and energy  $U$  are related by

$$U^2 = p^2 c^2 + m^2 c^4. \quad (24.29)$$

But in quantum mechanics the energy is  $\hbar\omega$ , and the momentum is  $\hbar/\lambda$ , which is equal to  $\hbar k$ ; so Eq. (24.29) can be written

$$\frac{\omega^2}{c^2} = k^2 + \frac{m^2 c^2}{\hbar^2}, \quad (24.30)$$

or

$$k = \sqrt{(\omega^2/c^2) - (m^2 c^2/\hbar^2)}, \quad (24.31)$$

which looks very much like Eq. (24.17) . . . Interesting!

The group velocity of the waves is also the speed at which energy is transported along the guide. If we want to find the energy flow down the guide, we can get it from the energy density times the group velocity. If the root mean square electric field is  $E_0$ , then the average density of electric energy is  $\epsilon_0 E_0^2/2$ . There is also some energy associated with the magnetic field. We will not prove it here, but in any cavity or guide the magnetic and electric energies are equal, so the total electromagnetic energy density is  $\epsilon_0 E_0^2$ . The power  $dU/dt$  transmitted by the guide is then

$$\frac{dU}{dt} = \epsilon_0 E_0^2 a b v_{\text{group}}. \quad (24.32)$$

(We will see later another, more general way of getting the energy flow.)

#### 24-5 Observing guided waves

Energy can be coupled into a waveguide by some kind of an "antenna." For example, a little vertical wire or "stub" will do. The presence of the guided waves can be observed by picking up some of the electromagnetic energy with a little receiving "antenna," which again can be a little stub of wire or a small loop.

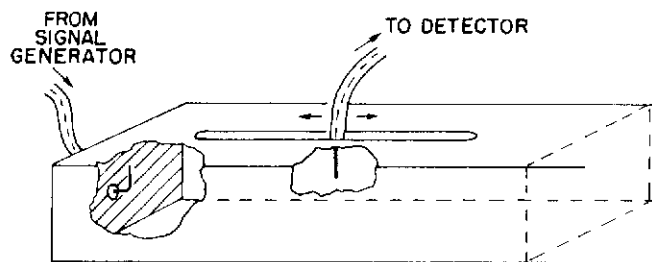


Fig. 24-8. A waveguide with a driving stub and a pickup probe.

In Fig. 24-8, we show a guide with some cutaways to show a driving stub and a pickup "probe". The driving stub can be connected to a signal generator via a coaxial cable, and the pickup probe can be connected by a similar cable to a detector. It is usually convenient to insert the pickup probe via a long thin slot in the guide, as shown in Fig. 24-8. Then the probe can be moved back and forth along the guide to sample the fields at various positions.

If the signal generator is set at some frequency  $\omega$  greater than the cutoff frequency  $\omega_c$ , there will be waves propagated down the guide from the driving stub. These will be the only waves present if the guide is infinitely long, which can effectively be arranged by terminating the guide with a carefully designed absorber in such a way that there are no reflections from the far end. Then, since the detector measures the time average of the fields near the probe, it will pick up a signal which is independent of the position along the guide; its output will be proportional to the power being transmitted.

If now the far end of the guide is finished off in some way that produces a reflected wave—as an extreme example, if we closed it off with a metal plate—there will be a reflected wave in addition to the original forward wave. These two waves will interfere and produce a standing wave in the guide similar to the standing waves on a string which we discussed in Chapter 49 of Vol. I. Then, as the pickup probe is moved along the line, the detector reading will rise and fall periodically, showing a maximum in the fields at each loop of the standing wave and a minimum at each node. The distance between two successive nodes (or loops) is just  $\lambda_g/2$ . This gives a convenient way of measuring the guide wavelength. If the frequency is now moved closer to  $\omega_c$ , the distances between nodes increase, showing that the guide wavelength increases as predicted by Eq. (24.19).

Suppose now the signal generator is set at a frequency just a little below  $\omega_c$ . Then the detector output will decrease gradually as the pickup probe is moved down the guide. If the frequency is set somewhat lower, the field strength will fall rapidly, following the curve of Fig. 24-7, and showing that waves are not propagated.

#### 24-6 Waveguide plumbing

An important practical use of waveguides is for the transmission of high-frequency power, as, for example, in coupling the high-frequency oscillator or output amplifier of a radar set to an antenna. In fact, the antenna itself usually consists of a parabolic reflector fed at its focus by a waveguide flared out at the end to make a "horn" that radiates the waves coming along the guide. Although high frequencies can be transmitted along a coaxial cable, a waveguide is better for transmitting large amounts of power. First, the maximum power that can be transmitted along a line is limited by the breakdown of the insulation (solid or gas) between the conductors. For a given amount of power, the field strengths in a guide are usually less than they are in a coaxial cable, so higher powers can be transmitted before breakdown occurs. Second, the power losses in the coaxial cable are usually greater than in a waveguide. In a coaxial cable there must be insulating material to support the central conductor, and there is an energy loss in this material—particularly at high frequencies. Also, the current densities on the central conductor are quite high, and since the losses go as the *square* of the current density, the lower currents that appear on the walls of the guide result in lower



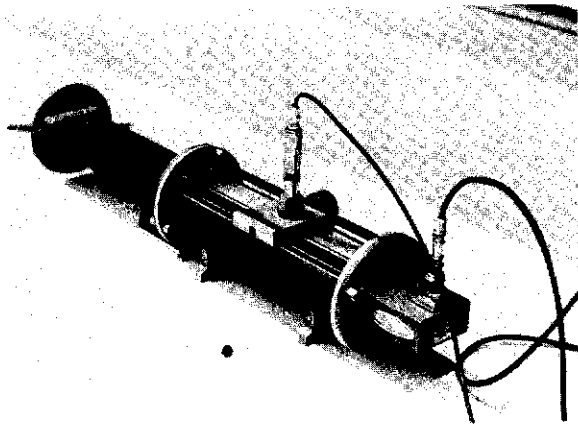


Fig. 24-9. Sections of waveguide connected with flanges.

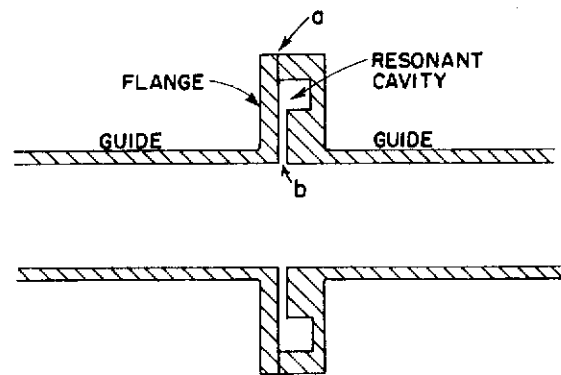


Fig. 24-10. A low-loss connection between two sections of waveguide.

energy losses. To keep these losses to a minimum, the inner surfaces of the guide are often plated with a material of high conductivity, such as silver.

The problem of connecting a "circuit" with waveguides is quite different from the corresponding circuit problem at low frequencies, and is usually called microwave "plumbing." Many special devices have been developed for the purpose. For instance, two sections of waveguide are usually connected together by means of flanges, as can be seen in Fig. 24-9. Such connections can, however, cause serious energy losses, because the surface currents must flow across the joint, which may have a relatively high resistance. One way to avoid such losses is to make the flanges as shown in the cross section drawn in Fig. 24-10. A small space is left between the adjacent sections of the guide, and a groove is cut in the face of one of the flanges to make a small cavity of the type shown in Fig. 23-16(c). The dimensions are chosen so that this cavity is resonant at the frequency being used. This resonant cavity presents a high "impedance" to the currents, so relatively little current flows across the metallic joints (at *a* in Fig. 24-10). The high guide currents simply charge and discharge the "capacity" of the gap (at *b* in the figure), where there is little dissipation of energy.

Suppose you want to stop a waveguide in a way that won't result in reflected waves. Then you must put something at the end that imitates an infinite length of guide. You need a "termination" which acts for the guide like the characteristic impedance does for a transmission line—something that absorbs the arriving waves without making reflections. Then the guide will act as though it went on forever. Such terminations are made by putting inside the guide some wedges of resistance material carefully designed to absorb the wave energy while generating almost no reflected waves.

If you want to connect *three* things together—for instance, one source to two different antennas—then you can use a "T" like the one shown in Fig. 24-11. Power fed in at the center section of the "T" will be split and go out the two side arms (and there may also be some reflected waves). You can see qualitatively from the sketches in Fig. 24-12 that the fields would spread out when they get to the end of the input section and make electric fields that will start waves going out the two arms. Depending on whether electric fields in the guide are parallel or perpendicular to the "top" of the "T," the fields at the junction would be roughly as shown in (a) or (b) of Fig. 24-12.

Finally, we would like to describe a device called an "unidirectional coupler," which is very useful for telling what is going on after you have connected a complicated arrangement of waveguides. Suppose you want to know which way the waves are going in a particular section of guide—you might be wondering, for instance, whether or not there is a strong reflected wave. The unidirectional coupler takes out a small fraction of the power of a guide if there is a wave going one way, but none if the wave is going the other way. By connecting the output of the coupler to a detector, you can measure the "one-way" power in the guide.

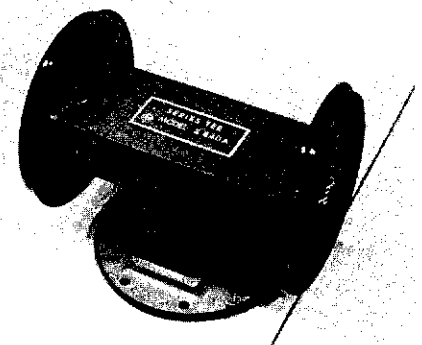


Fig. 24-11. A waveguide "T." (The flanges have plastic end caps to keep the inside clean while the "T" is not being used.)

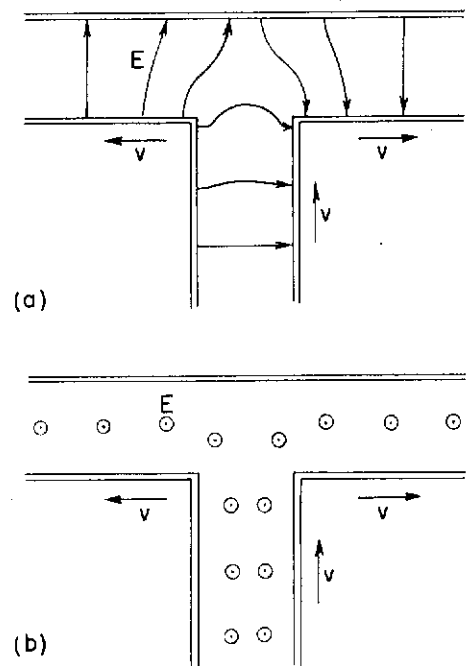


Fig. 24-12. The electric fields in a waveguide "T" for two possible field orientations.

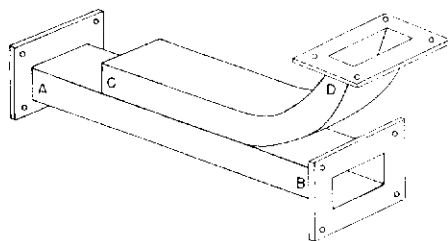


Fig. 24-13. A unidirectional coupler.

Figure 24-13 is a drawing of a unidirectional coupler; a piece of waveguide  $AB$  has another piece of waveguide  $CD$  soldered to it along one face. The guide  $CD$  is curved away so that there is room for the connecting flanges. Before the guides are soldered together, two (or more) holes have been drilled in each guide (matching each other) so that some of the fields in the main guide  $AB$  can be coupled into the secondary guide  $CD$ . Each of the holes acts like a little antenna that produces a wave in the secondary guide. If there were only one hole, waves would be sent in both directions and would be the same no matter which way the wave was going in the primary guide. But when there are *two* holes with a separation space equal to one-quarter of the guide wavelength, they will make two sources  $90^\circ$  out of phase. Do you remember that we considered in Chapter 29 of Vol. I the interference of the waves from two antennas spaced  $\lambda/4$  apart and excited  $90^\circ$  out of phase in time? We found that the waves subtract in one direction and add in the opposite direction. The same thing will happen here. The wave produced in the guide  $CD$  will be going in the same direction as the wave in  $AB$ .

If the wave in the primary guide is travelling from  $A$  toward  $B$ , there will be a wave at the output  $D$  of the secondary guide. If the wave in the primary guide goes from  $B$  toward  $A$ , there will be a wave going toward the end  $C$  of the secondary guide. This end is equipped with a termination, so that this wave is absorbed and there is no wave at the output of the coupler.

## 24-7 Waveguide modes

The wave we have chosen to analyze is a special solution of the field equations. There are many more. Each solution is called a waveguide "mode." For example, our  $x$ -dependence of the field was just one-half a cycle of a sine wave. There is an equally good solution with a full cycle; then the variation of  $E_y$  with  $x$  is as shown in Fig. 24-14. The  $k_x$  for such a mode is twice as large, so the cutoff frequency is much higher. Also, in the wave we studied  $E$  has only a  $y$ -component, but there are other modes with more complicated electric fields. If the electric field has components only in  $x$  and  $y$ —so that the total electric field is always at right angles to the  $z$ -direction—the mode is called a "transverse electric" (or TE) mode. The magnetic field of such modes will always have a  $z$ -component. It turns out that if  $E$  has a component in the  $z$ -direction (along the direction of propagation), then the magnetic field will always have only transverse components. So such fields are called transverse magnetic (TM) modes. For a rectangular guide, all the other modes have a higher cutoff frequency than the simple TE mode we have described. It is, therefore, possible—and usual—to use a guide with a frequency just above the cutoff for this lowest mode but below the cutoff frequency for all the others, so that just the one mode is propagated. Otherwise, the behavior gets complicated and difficult to control.

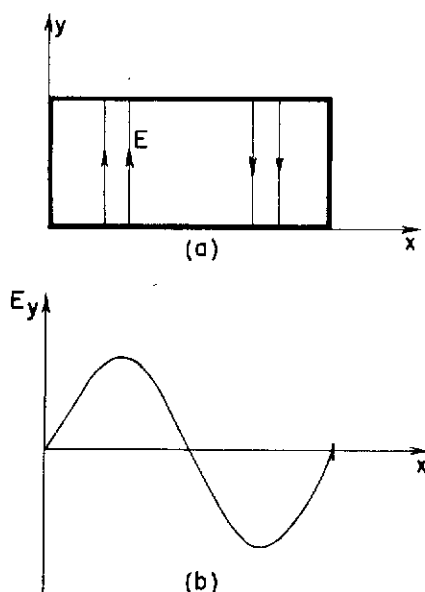


Fig. 24-14. Another possible variation of  $E_y$  with  $x$ .

## 24-8 Another way of looking at the guided waves

We want now to show you another way of understanding why a waveguide attenuates the fields rapidly for frequencies below the cutoff frequency  $\omega_c$ . Then you will have a more "physical" idea of why the behavior changes so drastically between low and high frequencies. We can do this for the rectangular guide by analyzing the fields in terms of reflections—or images—in the walls of the guide. The approach only works for rectangular guides, however; that's why we started with the more mathematical analysis which works, in principle, for guides of any shape.

For the mode we have described, the vertical dimension (in  $y$ ) had no effect, so we can ignore the top and bottom of the guide and imagine that the guide is extended indefinitely in the vertical direction. We imagine then that the guide just consists of two vertical plates with the separation  $a$ .

Let's say that the source of the fields is a vertical wire placed in the middle of the guide, with the wire carrying a current that oscillates at the frequency  $\omega$ . In the absence of the guide walls such a wire would radiate cylindrical waves.

Now we consider that the guide walls are perfect conductors. Then, just as in electrostatics, the conditions at the surface will be correct if we add to the field of the wire the field of one or more suitable image wires. The image idea works just as well for electrodynamics as it does for electrostatics, provided, of course, that we also include the retardations. We know that is true because we have often seen a mirror producing an image of a light source. And a mirror is just a "perfect" conductor for electromagnetic waves with optical frequencies.

Now let's take a horizontal cross section, as shown in Fig. 24-15, where  $W_1$  and  $W_2$  are the two guide walls and  $S_0$  is the source wire. We call the direction of the current in the wire positive. Now if there were only one wall, say  $W_1$ , we could remove it if we placed an image source (with opposite polarity) at the position marked  $S_1$ . But with both walls in place there will also be an image of  $S_0$  in the wall  $W_2$ , which we show as the image  $S_2$ . This source, too, will have an image in  $W_1$ , which we call  $S_3$ . Now both  $S_1$  and  $S_3$  will have images in  $W_2$  at the positions marked  $S_4$  and  $S_6$ , and so on. For our two plane conductors with the source halfway between, the fields are the same as those produced by an infinite line of sources, all separated by the distance  $a$ . (It is, in fact just what you would see if you looked at a wire placed halfway between two parallel mirrors.) For the fields to be zero at the walls, the polarity of the currents in the images must alternate from one image to the next. In other words, they oscillate  $180^\circ$  out of phase. The waveguide field is, then, just the superposition of the fields of such an infinite set of line sources.

We know that if we are close to the sources, the field is very much like the static fields. We considered in Section 7-5 the static field of a grid of line sources and found that it is like the field of a charged plate except for terms that decrease exponentially with the distance from the grid. Here the average source strength is zero, because the sign alternates from one source to the next. Any fields which exist should fall off exponentially with distance. Close to the source, we see the field mainly of the nearest source; at large distances, many sources contribute and their average effect is zero. So now we see why the waveguide below cutoff frequency gives an exponentially decreasing field. At low frequencies, in particular, the static approximation is good, and it predicts a rapid attenuation of the fields with distance.

Now we are faced with the opposite question: Why are waves propagated at all? That is the mysterious part! The reason is that at high frequencies the retardation of the fields can introduce additional changes in phase which can cause the fields of the out-of-phase sources to add instead of cancelling. In fact, in Chapter 29 of Vol. I we have already studied, just for this problem, the fields generated by an array of antennas or by an optical grating. There we found that when several radio antennas are suitably arranged, they can give an interference pattern that has a strong signal in some direction but no signal in another.

Suppose we go back to Fig. 24-15 and look at the fields which arrive at a large distance from the array of image sources. The fields will be strong only in certain directions which depend on the frequency—only in those directions for which the fields from all the sources add in phase. At a reasonable distance from the sources the field propagates in these special directions as plane waves. We have sketched such a wave in Fig. 24-16, where the solid lines represent the wave crests and the dashed lines represent the troughs. The wave direction will be the one for which the difference in the retardation for two neighboring sources to the crest of a wave corresponds to one-half a period of oscillation. In other words, the difference between  $r_2$  and  $r_0$  in the figure is one-half of the free-space wavelength:

$$r_2 - r_0 = \frac{\lambda_0}{2}.$$

The angle  $\theta$  is then given by

$$\sin \theta = \frac{\lambda_0}{2a}. \quad (24.33)$$

There is, of course, another set of waves travelling downward at the symmetric angle with respect to the array of sources. The complete waveguide field (not too

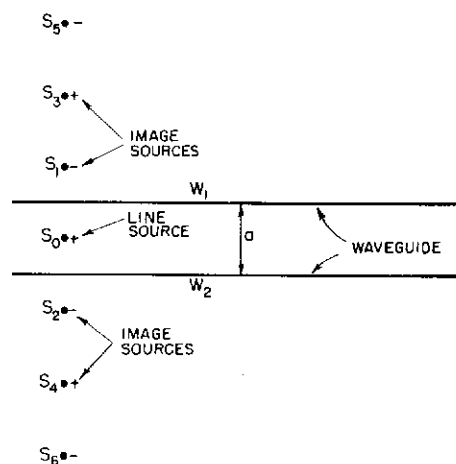


Fig. 24-15. The line source  $S_0$  between the conducting plane walls  $W_1$  and  $W_2$ . The walls can be replaced by the infinite sequence of image sources.

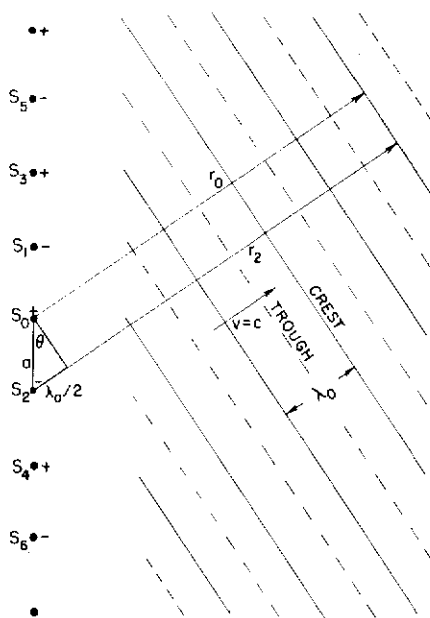


Fig. 24-16. One set of coherent waves from an array of line sources.

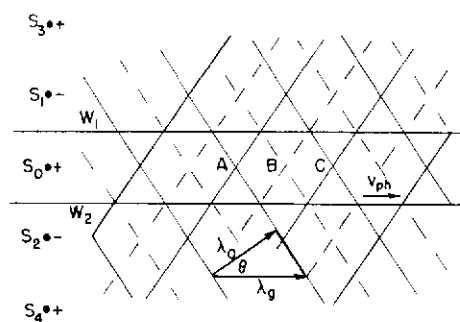


Fig. 24-17. The waveguide field can be viewed as the superposition of two trains of plane waves.

close to the source) is the superposition of these two sets of waves, as shown in Fig. 24-17. The actual fields are really like this, of course, only between the two walls of the waveguide.

At points like *A* and *C*, the crests of the two wave patterns coincide, and the field will have a maximum; at points like *B*, both waves have their peak negative value, and the field has its minimum (largest negative) value. As time goes on the field in the guide appears to be travelling along the guide with a wavelength  $\lambda_g$ , which is the distance from *A* to *C*. That distance is related to  $\theta$  by

$$\cos \theta = \frac{\lambda_0}{\lambda_g}. \quad (24.34)$$

Using Eq. (24.33) for  $\theta$ , we get that

$$\lambda_g = \frac{\lambda_0}{\cos \theta} = \frac{\lambda_0}{\sqrt{1 - (\lambda_0/2a)^2}}, \quad (24.35)$$

which is just what we found in Eq. (24.19).

Now we see why there is only wave propagation above the cutoff frequency  $\omega_0$ . If the free-space wavelength is longer than  $2a$ , there is no angle where the waves shown in Fig. 24-16 can appear. The necessary constructive interference appears suddenly when  $\lambda_0$  drops below  $2a$ , or when  $\omega$  goes above  $\omega_0 = \pi c/a$ .

If the frequency is high enough, there can be two or more possible directions in which the waves will appear. For our case, this will happen if  $\lambda_0 < \frac{2}{3}a$ . In general, however, it could also happen when  $\lambda_0 < a$ . These additional waves correspond to the higher guide modes we have mentioned.

It has also been made evident by our analysis why the phase velocity of the guided waves is greater than  $c$  and why this velocity depends on  $\omega$ . As  $\omega$  is changed, the angle of the free waves of Fig. 24-16 changes, and therefore so does the velocity along the guide.

Although we have described the guided wave as the superposition of the fields of an infinite array of line sources, you can see that we would arrive at the same result if we imagined two sets of free-space waves being continually reflected back and forth between two perfect mirrors—remembering that a reflection means a reversal of phase. These sets of reflecting waves would all cancel each other unless they were going at just the angle  $\theta$  given in Eq. (24.33). There are many ways of looking at the same thing.

### 32-4 The complex index of refraction

We want to look now at the consequences of our result, Eq. (32.33). First, we notice that  $\alpha$  is complex, so the index  $n$  is going to be a complex number. What does that mean? Let's say that we write  $n$  as the sum of a real and an imaginary part:

$$n = n_R - in_I, \quad (32.35)$$

where  $n_R$  and  $n_I$  are real functions of  $\omega$ . We write  $in_I$  with a minus sign, so that  $n_I$  will be a positive quantity in all ordinary optical materials. (In ordinary inactive materials—that are not, like lasers, light sources themselves— $\gamma$  is a positive number, and that makes the imaginary part of  $n$  negative.) Our plane wave of Eq. (32.21) is written in terms of  $n$  as

$$E_x = E_0 e^{-i\omega(t - nz/c)}.$$

Writing  $n$  as in Eq. (32.35), we would have

$$E_x = E_0 e^{-\omega n_I z/c} e^{i\omega(t - n_R z/c)}. \quad (32.36)$$

The term  $e^{i\omega(t - n_R z/c)}$  represents a wave travelling with the speed  $c/n_R$ , so  $n_R$  represents what we normally think of as the index of refraction. But the *amplitude* of this wave is

$$E_0 e^{-\omega n_I z/c},$$

which decreases exponentially with  $z$ . A graph of the strength of the electric field at some instant as a function of  $z$  is shown in Fig. 32-1, for  $n_I \approx n_R/2\pi$ . The imaginary part of the index represents the attenuation of the wave due to the energy losses in the atomic oscillators. The *intensity* of the wave is proportional to the square of the amplitude, so

$$\text{Intensity} \propto e^{-2\omega n_I z/c}.$$

This is often written as

$$\text{Intensity} \propto e^{-\beta z},$$

where  $\beta = 2\omega n_I/c$  is called the *absorption coefficient*. Thus we have in Eq. (32.33) not only the theory of the index of refraction of materials, but the theory of their absorption of light as well.

In what we usually consider to be transparent material, the quantity  $c/\omega n_I$ —which has the dimensions of a length—is quite large in comparison with the thickness of the material.

### 32-5 The index of a mixture

There is another prediction of our theory of the index of refraction that we can check against experiment. Suppose we consider a mixture of two materials. The index of the mixture is not the average of the two indexes, but should be given in terms of the sum of the two polarizabilities, as in Eq. (32.34). If we ask about the index of, say, a sugar solution, the total polarizability is the sum of the polarizability of the water and that of the sugar. Each must, of course, be calculated using for  $N$  the number per unit volume of the molecules of the particular kind. In other words, if a given solution has  $N_1$  molecules of water, whose polarizability is  $\alpha_1$ , and  $N_2$  molecules of sucrose ( $\text{C}_{12}\text{H}_{22}\text{O}_{11}$ ), whose polarizability is  $\alpha_2$ , we should have that

$$3 \left( \frac{n^2 - 1}{n^2 + 2} \right) = N_1 \alpha_1 + N_2 \alpha_2. \quad (32.37)$$

We can use this formula to test our theory against experiment by measuring the index for various concentrations of sucrose in water. We are making several assumptions here, however. Our formula assumes that there is no chemical action when the sucrose is dissolved and that the disturbances to the individual atomic

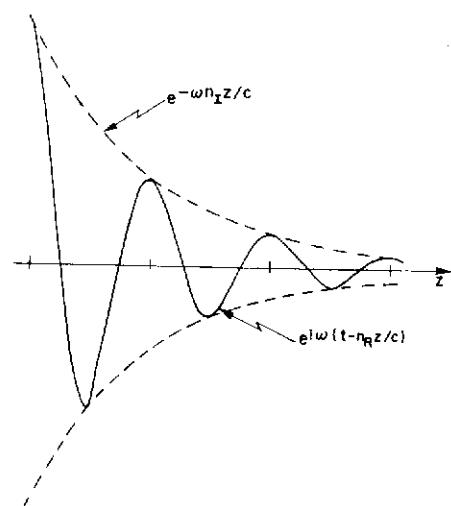


Fig. 32-1. A graph of  $E_x$  for some instant  $t$ , if  $n_I \approx n_R/2\pi$ .

Table 32-2

Refractive index of sucrose solutions, and comparison with predictions of Eq. (32.37).

Data from Handbook								
A	B	C	D	E	F	G	H	J
Fraction of sucrose by weight	density (gm/cm <sup>3</sup> )	$n$ at 20°C	Moles of sucrose <sup>d</sup> per liter, $N_2/N_0$	Moles of water <sup>c</sup> per liter, $N_1/N_0$	$3 \left( \frac{n^2 - 1}{n^2 + 2} \right)$	$N_1\alpha_1$	$N_2\alpha_2$	$N_0\alpha_2$ (gm/liter)
0 <sup>a</sup>	0.9982	1.333	0	55.5	0.617	$\frac{0.617}{0.487}$	0	—
0.30	1.1270	1.3811	0.970	43.8	0.698	0.487	0.211	0.213
0.50	1.2296	1.4200	1.798	34.15	0.759	0.379	0.380	0.211
0.85	1.4454	1.5033	3.59	12.02	0.886	0.1335	0.752	0.210
1.00 <sup>b</sup>	1.588	1.5577 <sup>c</sup>	4.64	0	0.960	0	0.960	0.207

<sup>a</sup> pure water<sup>c</sup> average (see text)<sup>e</sup> molecular weight of water = 18<sup>b</sup> sugar crystals<sup>d</sup> molecular weight of sucrose = 342

oscillators are not too different for various concentrations. So our result is certainly only approximate. Anyway, let's see how good it is.

We have picked the example of a sugar solution because there is a good table of measurements of the index of refraction in the *Handbook of Chemistry and Physics* and also because sugar is a molecular crystal that goes into solution without ionizing or otherwise changing its chemical state.

We give in the first three columns of Table 32-2 the data from the handbook. Column A is the percent of sucrose by weight, column B is the measured density (gm/cm<sup>3</sup>), and column C is the measured index of refraction for light whose wavelength is 589.3 millimicrons. For pure sugar we have taken the measured index of sugar crystals. The crystals are not isotropic, so the measured index is different along different directions. The handbook gives three values:

$$n_1 = 1.5376, \quad n_2 = 1.5651, \quad n_3 = 1.5705.$$

We have taken the average.

Now we could try to compute  $n$  for each concentration, but we don't know what value to take for  $\alpha_1$  or  $\alpha_2$ . Let's test the theory this way: We will assume that the polarizability of water ( $\alpha_1$ ) is the same at all concentrations and compute the polarizability of sucrose by using the experiment of values for  $n$  and solving Eq. (38.27) for  $\alpha_2$ . If the theory is correct, we should get the same  $\alpha_2$  for all concentrations.

First, we need to know  $N_1$  and  $N_2$ : let's express them in terms of Avogadro's number,  $N_0$ . Let's take one liter (1000 cm<sup>3</sup>) for our unit of volume. Then  $N_i/N_0$  is the weight per liter divided by the gram-molecular weight. And the weight per liter is the density (multiplied by 1000 to get grams per liter) times the fractional weight of either the sucrose or the water. In this way, we get  $N_2/N_0$  and  $N_1/N_0$  as in columns D and E of the table.

In column F we have computed  $3(n^2 - 1)/(n^2 + 2)$  from the experimental values of  $n$  in column C. For pure water,  $3(n^2 - 1)/(n^2 + 2)$  is 0.617, which is equal to just  $N_1\alpha_1$ . We can then fill in the rest of Column G, since for each row G/E may be in the same ratio—namely, 0.617:55.5. Subtracting column G from column F, we get the contribution  $N_2\alpha_2$  of the sucrose, shown in column H. Dividing these entries by the values of  $N_2/N_0$  in column D, we get the value of  $N_0\alpha_2$  shown in column J.

From our theory we would expect all the values of  $N_0\alpha_2$  to be the same. They are not exactly equal, but pretty close. We can conclude that our ideas are fairly correct. Even more, we find that the polarizability of the sugar molecule doesn't seem to depend much on its surroundings—its polarizability is nearly the same in a dilute solution as it is in the crystal.

### 32-6 Waves in metals

The theory we have worked out in this chapter for solid materials can also be applied to good conductors, like metals, with very little modification. In metals some of the electrons have no binding force holding them to any particular atom; it is these "free" electrons which are responsible for the conductivity. There are other electrons which are bound, and the theory above is directly applicable to them. Their influence, however, is usually swamped by the effects of the conduction electrons. We will consider now only the effects of the free electrons.

If there is no restoring force on an electron—but still some resistance to its motion—its equation of motion differs from Eq. (32.1) only because the term in  $\omega_0^2 x$  is lacking. So all we have to do is set  $\omega_0^2 = 0$  in the rest of our derivations—except that there is one more difference. The reason that we had to distinguish between the average field and the local field in a dielectric is that in an insulator each of the dipoles is fixed in position, so that it has a definite relationship to the position of the others. But because the conduction electrons in a metal move around all over the place, the field on them *on the average* is just the average field  $E$ . So the correction we made to Eq. (32.5) by using Eq. (32.28) should *not* be made for conduction electrons. Therefore the formula for the index of refraction for metals should look like Eq. (32.27), except with  $\omega_0$  set equal to zero, namely,

$$n^2 = 1 + \frac{Nq_e^2}{m\epsilon_0} \frac{1}{-\omega^2 + i\gamma\omega}. \quad (32.38)$$

This is only the contribution from the conduction electrons, which we will assume is the major term for metals.

Now we even know how to find what value to use for  $\gamma$ , because it is related to the conductivity of the metal. In Chapter 43 of Volume I we discussed how the conductivity of a metal comes from the diffusion of the free electrons through the crystal. The electrons go on a jagged path from one scattering to the next, and between scatterings they move freely except for an acceleration due to any average electric field (as shown in Fig. 32-2). We found in Chapter 43 of Volume I that the average drift velocity is just the acceleration times the average time  $\tau$  between collisions. The acceleration is  $q_e E/m$ , so

$$v_{\text{drift}} = \frac{q_e E}{m} \tau. \quad (32.39)$$

This formula assumed that  $E$  was constant, so that  $v_{\text{drift}}$  was a steady velocity. Since there is no average acceleration, the drag force is equal to the applied force. We have defined  $\gamma$  by saying that  $\gamma m v$  is the drag force [see Eq. (32.1)], which is  $q_e E$ ; therefore we have that

$$\gamma = \frac{1}{\tau}. \quad (32.40)$$

Although we cannot easily measure  $\tau$  directly, we can determine it by measuring the conductivity of the metal. It is found experimentally that an electric field  $E$  in a metal produces a current with the density  $j$  proportional to  $E$  (for isotropic materials):

$$j = \sigma E.$$

The proportionality constant  $\sigma$  is called the *conductivity*. This is just what we expect from Eq. (32.39) if we set

$$j = Nq_e v_{\text{drift}}.$$

Then

$$\sigma = \frac{Nq_e^2}{m} \tau. \quad (32.41)$$

So  $\tau$ —and therefore  $\gamma$ —can be related to the observed electrical conductivity. Using Eqs. (32.40) and (32.41), we can rewrite our formula for the index, Eq.

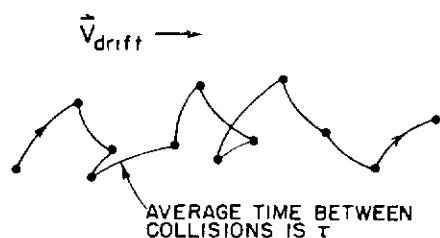


Fig. 32-2. The motion of a free electron.

(32.38), in the following form:

$$n^2 = 1 + \frac{\sigma/\epsilon_0}{i\omega(1 + i\omega\tau)}, \quad (32.42)$$

where

$$\tau = \frac{1}{\gamma} = \frac{m\sigma}{Nq_e^2}. \quad (32.43)$$

This is a convenient formula for the index of refraction of metals.

### 32-7 Low-frequency and high-frequency approximations; the skin depth and the plasma frequency

Our result, Eq. (32.42), for the index of refraction for metals predicts quite different characteristics for wave propagation at different frequencies. Let's first see what happens at very *low* frequencies. If  $\omega$  is small enough, we can approximate Eq. (32.42) by

$$n^2 = -i \frac{\sigma}{\epsilon_0\omega}. \quad (32.44)$$

Now, as you can check by taking the square,\*

$$\sqrt{-i} = \frac{1-i}{\sqrt{2}};$$

so for low frequencies,

$$n = \sqrt{\sigma/2\epsilon_0\omega} (1 - i). \quad (32.45)$$

The real and imaginary parts of  $n$  have the same magnitude. With such a large imaginary part to  $n$ , the wave is rapidly attenuated in the metal. Referring to Eq. (32.36), the amplitude of a wave going in the  $z$ -direction decreases as

$$\exp[-\sqrt{\sigma\omega/2\epsilon_0c^2} z]. \quad (32.46)$$

Let's write this as

$$e^{-z/\delta}, \quad (32.47)$$

where  $\delta$  is then the distance in which the wave amplitude decreases by the factor  $e^{-1} = 1/2.72$ —or roughly one-third. The amplitude of such a wave as a function of  $z$  is shown in Fig. 32-3. Since electromagnetic waves will penetrate into a metal only this distance,  $\delta$  is called the *skin depth*. It is given by

$$\delta = \sqrt{2\epsilon_0c^2/\sigma\omega}. \quad (32.48)$$

Now what do we mean by "low" frequencies? Looking at Eq. (32.42), we see that it can be approximated by Eq. (32.44) only if  $\omega\tau$  is much less than one and if  $\omega\epsilon_0/\sigma$  is also much less than one—that is, our low-frequency approximation applies when

$$\omega \ll \frac{1}{\tau}$$

and

$$\omega \ll \frac{\sigma}{\epsilon_0}. \quad (32.49)$$

Let's see what frequencies these correspond to for a typical metal like copper. We compute  $\tau$  by using Eq. (32.43), and  $\sigma/\epsilon_0$ , by using the measured conductivity. We take the following data from a handbook:

$$\sigma = 5.76 \times 10^7 \text{ (ohm-meter)}^{-1},$$

$$\text{atomic weight} = 63.5 \text{ grams},$$

$$\text{density} = 8.9 \text{ grams/cm}^3,$$

$$\text{Avogadro's number} = 6.02 \times 10^{23} \text{ (gram atomic weight)}^{-1}.$$

\* Or writing  $-i = e^{-i\pi/2}$ ;  $\sqrt{-i} = e^{-i\pi/4} = \cos \pi/4 - i \sin \pi/4$ , which gives the same result.

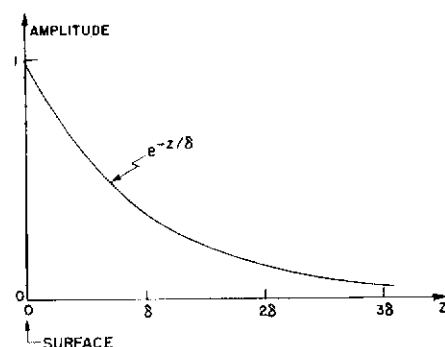


Fig. 32-3. The amplitude of a transverse electromagnetic wave as a function of distance into a metal.



If we assume that there is one free electron per atom, then the number of electrons per cubic meter is

$$N = 8.5 \times 10^{28} \text{ meter}^{-3}.$$

Using

$$q_e = 1.6 \times 10^{-19} \text{ coulomb},$$

$$\epsilon_0 = 8.85 \times 10^{-12} \text{ farad-meter}^{-1},$$

$$m = 9.11 \times 10^{-31} \text{ kgm},$$

we get

$$\tau = 2.4 \times 10^{-14} \text{ sec},$$

$$\frac{1}{\tau} = 4.1 \times 10^{13} \text{ sec}^{-1},$$

$$\frac{\sigma}{\epsilon_0} = 6.5 \times 10^{18} \text{ sec}^{-1}.$$

So for frequencies less than about  $10^{12}$  cycles per second, copper will have the "low-frequency" behavior we describe (that means for waves whose free-space wavelength is longer than 0.3 millimeters—very short radio waves!).

For these waves, the skin depth in copper is

$$\delta = \sqrt{\frac{0.028 \text{ m}^2\text{-sec}^{-1}}{\omega}}.$$

For microwaves of 10,000 megacycles per second (3-cm waves)

$$\delta = 6.7 \times 10^{-4} \text{ cm}.$$

The wave penetrates a very small distance.

We can see from this why in studying cavities (or waveguides) we needed to worry only about the fields inside the cavity, and not in the metal or outside the cavity. Also, we see why the losses in a cavity are reduced by a thin plating of silver or gold. The losses come from the current, which are appreciable only in a thin layer equal to the skin depth.

Suppose we look now at the index of a metal like copper at high frequencies. For very high frequencies  $\omega\tau$  is much greater than one, and Eq. (32.42) is well approximated by

$$n^2 = 1 - \frac{\sigma}{\epsilon_0\omega^2\tau}. \quad (32.50)$$

For waves of high frequencies the index of a metal becomes real—and less than one! This is also evident from Eq. (32.38) if the dissipation term with  $\gamma$  is neglected, as can be done for very large  $\omega$ . Equation (32.38) gives

$$n^2 = 1 - \frac{Nq_e^2}{m\epsilon_0\omega^2}, \quad (32.51)$$

which is, of course, the same as Eq. (32.50). We have seen before the quantity  $Nq_e^2/m\epsilon_0$ , which we called the square of the plasma frequency (Section 7-3):

$$\omega_p^2 = \frac{Nq_e^2}{\epsilon_0 m},$$

so we can write Eq. (32.50) or Eq. (32.51) as

$$n^2 = 1 - \left(\frac{\omega_p}{\omega}\right)^2.$$

The plasma frequency is a kind of "critical" frequency.

For  $\omega < \omega_p$  the index of a metal has an imaginary part, and waves are attenuated; but for  $\omega \gg \omega_p$  the index is real, and the metal becomes transparent. You know, of course, that metals are reasonably transparent to x-rays. But some metals are even transparent in the ultraviolet. In Table 32-3 we give for

several metals the experimental observed wavelength at which they begin to become transparent. In the second column we give the calculated critical wavelength  $\lambda_p = 2\pi c/\omega_p$ . Considering that the experimental wavelength is not too well defined, the fit of the theory is fairly good.

You may wonder why the plasma frequency  $\omega_p$  should have anything to do with the propagation of electromagnetic waves in metals. The plasma frequency came up in Chapter 7 as the natural frequency of *density* oscillations of the free electrons. (A clump of electrons is repelled by electric forces, and the inertia of the electrons leads to an oscillation of density.) So *longitudinal* plasma waves are resonant at  $\omega_p$ . But we are now talking about *transverse* electromagnetic waves, and we have found that transverse waves are absorbed for frequencies below  $\omega_p$ . (It's an interesting and *not* accidental coincidence.)

Although we have been talking about wave propagation in metals, you appreciate by this time the universality of the phenomena of physics—that it doesn't make any difference whether the free electrons are in a metal or whether they are in the plasma of the ionosphere of the earth, or in the atmosphere of a star. To understand radio propagation in the ionosphere, we can use the same expressions—using, of course, the proper values for  $N$  and  $\tau$ . We can see now why long radio waves are absorbed or reflected by the ionosphere, whereas short waves go right through. (Short waves must be used for communication with satellites.)

We have talked about the high- and low-frequency extremes for wave propagation in metals. For the in-between frequencies the full-blown formula of Eq. (32.42) must be used. In general, the index will have real and imaginary parts; the wave is attenuated as it propagates into the metal. For very thin layers, metals are somewhat transparent even at optical frequencies. As an example, special goggles for people who work around high-temperature furnaces are made by evaporating a thin layer of gold on glass. The visible light is transmitted fairly well—with a strong green tinge—but the infrared is strongly absorbed.

Finally, it cannot have escaped the reader that many of these formulas resemble in some ways those for the dielectric constant  $\kappa$  discussed in Chapter 10. The dielectric constant  $\kappa$  measures the response of the material to a constant field, that is, for  $\omega = 0$ . If you look carefully at the definition of  $n$  and  $\kappa$  you see that  $\kappa$  is simply the limit of  $n^2$  as  $\omega \rightarrow 0$ . Indeed, placing  $\omega = 0$  and  $n^2 = \kappa$  in equations of this chapter will reproduce the equations of the theory of the dielectric constant of Chapter 11.

Table 32-3\*

Wavelengths below which the metal becomes transparent

Metal	$\lambda$ (experimental)	$\lambda_p = 2\pi c/\omega_p$
Li	1550 Å	1550 Å
Na	2100	2090
K	3150	2870
Rb	3400	3220

\* From: C. Kittel, *Introduction to Solid State Physics*, John Wiley and Sons, Inc., New York, 2nd ed., 1956, p. 266.