often take as unity. This is easily seen by considering the mathematical behavior of $a(t - r/c)$. Evidently, if we add a little time Δt, we get the same value for $a(t - r/c)$ as we would have if we had subtracted a little distance: $\Delta r = -c\,\Delta t$.

Stated another way: if we add a little time Δt, we can restore $a(t - r/c)$ to its former value by *adding* a little distance $\Delta r = c\,\Delta t$. That is, as time goes on *the field moves as a wave outward from the source.* That is the reason why we sometimes say light is propagated as waves. It is equivalent to saying that the field is delayed, or to saying that the electric field is moving outward as time goes on.

An interesting special case is that where the charge q is moving up and down in an oscillatory manner. The case which we studied experimentally in the last chapter was one in which the displacement x at any time t was equal to a certain constant x_0, the magnitude of the oscillation, times $\cos\omega t$. Then the acceleration is

$$a = -\omega^2 x_0 \cos\omega t = a_0 \cos\omega t, \tag{29.2}$$

where a_0 is the maximum acceleration, $-\omega^2 x_0$. Putting this formula into (29.1), we find

$$E = -q\sin\theta\,\frac{a_0\cos\omega(t - r/c)}{4\pi\epsilon_0 r c^2}. \tag{29.3}$$

Now, ignoring the angle θ and the constant factors, let us see what that looks like as a function of position or as a function of time.

29-2 Energy of radiation

First of all, at any particular moment or in any particular place, the strength of the field varies inversely as the distance r, as we mentioned previously. Now we must point out that the *energy* content of a wave, or the energy effects that such an electric field can have, are proportional to the *square* of the field, because if, for instance, we have some kind of a charge or an oscillator in the electric field, then if we let the field act on the oscillator, it makes it move. If this is a linear oscillator, the acceleration, velocity, and displacement produced by the electric field acting on the charge are all proportional to the field. So the kinetic energy which is developed in the charge is proportional to the *square* of the field. So we shall take it that the energy that a field can deliver to a system is proportional somehow to the square of the field.

This means that the energy that the source can deliver decreases as we get farther away; in fact, it varies *inversely as the square of the distance.* But that

Begin with Section 29-2. Note that electromagnetic waves are referred to frequently, the discussion applies to any classical wave. The speed of light is denoted c. If some other classical wave is under consideration, substitute its wave speed, v.

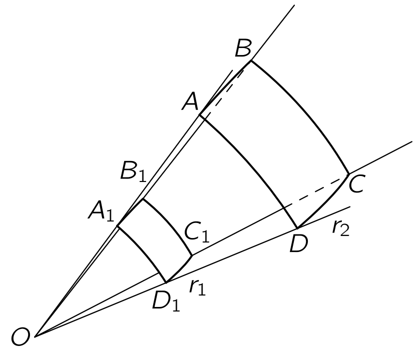

Fig. 29-4. The energy flowing within the cone $OABCD$ is independent of the distance r at which it is measured.

has a very simple interpretation: if we wanted to pick up all the energy we could from the wave in a certain cone at a distance r_1 (Fig. 29-4), and we do the same at another distance r_2, we find that the amount of energy per unit area at any one place goes inversely as the square of r, but the area of the surface intercepted by the cone goes *directly* as the square of r. So the energy that we can take out of the wave within a given conical angle is the same, no matter how far away we are! In particular, the total energy that we could take out of the whole wave by putting absorbing oscillators all around is a certain fixed amount. So the fact that the amplitude of E varies as $1/r$ is the same as saying that there is an energy flux which is never lost, an energy which goes on and on, spreading over a greater and greater effective area. Thus we see that after a charge has oscillated, it has lost some energy which it can never recover; the energy keeps going farther and farther away without diminution. So if we are far enough away that our basic approximation is good enough, the charge cannot recover the energy which has been, as we say, radiated away. Of course the energy still exists somewhere, and is available to be picked up by other systems. We shall study this energy "loss" further in Chapter 32.

Let us now consider more carefully how the wave (29.3) varies as a function of time at a given place, and as a function of position at a given time. Again we ignore the $1/r$ variation and the constants.

29-3 Sinusoidal waves

First let us fix the position r, and watch the field as a function of time. It is oscillatory at the angular frequency ω. The angular frequency ω can be defined as the *rate of change of phase with time* (radians per second). We have already studied such a thing, so it should be quite familiar to us by now. The *period* is the time needed for one oscillation, one complete cycle, and we have worked that out too; it is $2\pi/\omega$, because ω times the period is one cycle of the cosine.

Now we introduce a new quantity which is used a great deal in physics. This has to do with the opposite situation, in which we fix t and look at the wave as a function of distance r. Of course we notice that, as a function of r, the wave (29.3) is also oscillatory. That is, aside from $1/r$, which we are ignoring, we see that E oscillates as we change the position. So, in analogy with ω, we can define a quantity called the *wave number*, symbolized as k. This is defined as *the rate of change of phase with distance* (radians per meter). That is, as we move in space at a fixed time, the phase changes.

There is another quantity that corresponds to the period, and we might call it the period in space, but it is usually called the wavelength, symbolized λ. The wavelength is the distance occupied by one complete cycle. It is easy to see, then, that the wavelength is $2\pi/k$, because k times the wavelength would be the number of radians that the whole thing changes, being the product of the rate of change of the radians per meter, times the number of meters, and we must make a 2π change for one cycle. So $k\lambda = 2\pi$ is exactly analogous to $\omega t_0 = 2\pi$.

Now in our particular wave there is a definite relationship between the frequency and the wavelength, but the above definitions of k and ω are actually quite general. That is, the wavelength and the frequency may not be related in the same way in other physical circumstances. However, in our circumstance the rate of change of phase with distance is easily determined, because if we call $\phi = \omega(t - r/c)$ the phase, and differentiate (partially) with respect to distance r, the rate of change, $\partial\phi/\partial r$, is

$$\left|\frac{\partial\phi}{\partial r}\right| = k = \frac{\omega}{c}. \tag{29.4}$$

There are many ways to represent the same thing, such as

$$\lambda = ct_0 \qquad (29.5) \qquad\qquad \lambda\nu = c \qquad (29.7)$$

$$\omega = ck \qquad (29.6) \qquad\qquad \omega\lambda = 2\pi c \qquad (29.8)$$

Why is the wavelength equal to c times the period? That's very easy, of course, because if we sit still and wait for one period to elapse, the waves, travelling at the speed c, will move a distance ct_0, and will of course have moved over just one wavelength.

In a physical situation other than that of light, k is not necessarily related to ω in this simple way. If we call the distance along an axis x, then the formula for a cosine wave moving in a direction x with a wave number k and an angular frequency ω will be written in general as $\cos(\omega t - kx)$.

Now that we have introduced the idea of wavelength, we may say something more about the circumstances in which (29.1) is a legitimate formula. We recall that the field is made up of several pieces, one of which varies inversely as r, another part which varies inversely as r^2, and others which vary even faster. It would be worth while to know in what circumstances the $1/r$ part of the field is the most important part, and the other parts are relatively small. Naturally, the answer is "if we go 'far enough' away," because terms which vary inversely as the square ultimately become negligible compared with the $1/r$ term. How far is "far enough"? The answer is, qualitatively, that the other terms are of order λ/r smaller than the $1/r$ term. Thus, so long as we are beyond a few wavelengths, (29.1) is an excellent approximation to the field. Sometimes the region beyond a few wavelengths is called the "wave zone."

29-4 Two dipole radiators

Next let us discuss the mathematics involved in combining the effects of two oscillators to find the net field at a given point. This is very easy in the few cases that we considered in the previous chapter. We shall first describe the effects qualitatively, and then more quantitatively. Let us take the simple case, where the oscillators are situated with their centers in the same horizontal plane as the detector, and the line of vibration is vertical.

Figure 29-5(a) represents the top view of two such oscillators, and in this particular example they are half a wavelength apart in a N–S direction, and are oscillating together in the same phase, which we call zero phase. Now we would like to know the intensity of the radiation in various directions. By the intensity we mean the amount of energy that the field carries past us per second, which is proportional to the square of the field, averaged in time. So the thing to look at, when we want to know how bright the light is, is the square of the electric field, not the electric field itself. (The electric field tells the strength of the force felt by a stationary charge, but the amount of energy that is going past, in watts per square meter, is proportional to the square of the electric field. We shall derive

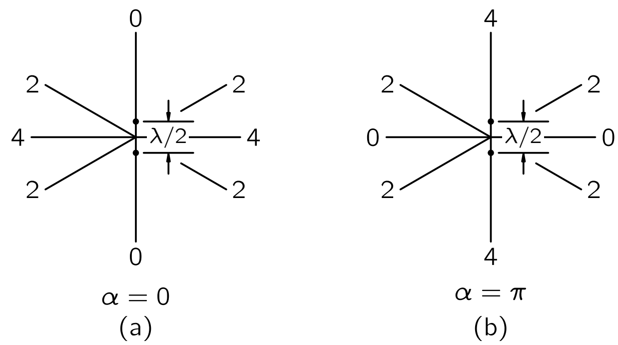

Fig. 29-5. The intensities in various directions from two dipole oscillators one-half wavelength apart. Left: in phase ($\alpha = 0$). Right: one-half period out of phase ($\alpha = \pi$).

the constant of proportionality in Chapter 31.) If we look at the array from the W side, both oscillators contribute equally and in phase, so the electric field is twice as strong as it would be from a single oscillator. Therefore *the intensity is four times as strong as it would be if there were only one oscillator*. (The numbers in Fig. 29-5 represent how strong the intensity would be in this case, compared with what it would be if there were only a single oscillator of unit strength.) Now, in either the N or S direction along the line of the oscillators, since they are half a wavelength apart, the effect of one oscillator turns out to be out of phase by exactly half an oscillation from the other, and therefore the fields add to zero. At a certain particular intermediate angle (in fact, at 30°) the intensity is 2, and it falls off, 4, 2, 0, and so forth. We have to learn how to find these numbers at other angles. It is a question of adding two oscillations with different phases.

Let us quickly look at some other cases of interest. Suppose the oscillators are again one-half a wavelength apart, but the phase α of one is set half a period behind the other in its oscillation (Fig. 29-5b). In the W direction the intensity is now zero, because one oscillator is "pushing" when the other one is "pulling." But in the N direction the signal from the near one comes at a certain time, and that of the other comes half a period later. But the latter was *originally* half a period behind in timing, and therefore it is now exactly *in time* with the first one, and so the intensity in this direction is 4 units. The intensity in the direction at 30° is still 2, as we can prove later.

Now we come to an interesting case which shows up a possibly useful feature. Let us remark that one of the reasons that phase relations of oscillators are interesting is for beaming radio transmitters. For instance, if we build an antenna system and want to send a radio signal, say, to Hawaii, we set the antennas up as in Fig. 29-5(a) and we broadcast with our two antennas in phase, because Hawaii is to the west of us. Then we decide that tomorrow we are going to broadcast toward Alberta, Canada. Since that is north, not west, all we have to do is to reverse the phase of one of our antennas, and we can broadcast to the north. So we can build antenna systems with various arrangements. Ours is one of the simplest possible ones; we can make them much more complicated, and by changing the phases in the various antennas we can send the beams in various directions and send most of the power in the direction in which we wish to transmit, without ever moving the antenna! In both of the preceding cases, however, while we are broadcasting toward Alberta we are wasting a lot of power on Easter Island, and it would be interesting to ask whether it is possible to send it in only *one* direction. At first sight we might think that with a pair of antennas of this nature the result is always going to be symmetrical. So let us consider a case that comes out unsymmetrical, to show the possible variety.

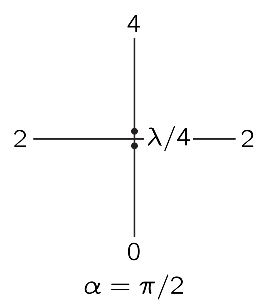

Fig. 29-6. A pair of dipole antennas giving maximum power in one direction.

If the antennas are separated by one-quarter wavelength, and if the N one is one-fourth period behind the S one in time, then what happens (Fig. 29-6)? In the W direction we get 2, as we will see later. In the S direction we get *zero*, because the signal from S comes at a certain time; that from N comes 90° later in *time*, but it is already 90° behind in its built-in phase, therefore it arrives, altogether, 180° out of phase, and there is no effect. On the other hand, in the N direction, the N signal arrives earlier than the S signal by 90° in time, because it is a quarter wavelength closer. But its phase is set so, that it is oscillating 90° *behind* in time, which just compensates the delay difference, and therefore the two signals appear *together* in phase, making the field strength twice as large, and the energy four times as great.

Thus, by using some cleverness in spacing and phasing our antennas, we can send the power all in one direction. But still it is distributed over a great range of angles. Can we arrange it so that it is focused still more sharply in a particular direction? Let us consider the case of Hawaii again, where we are sending the beam east and west but it is spread over quite an angle, because even at 30° we are still getting half the intensity—we are wasting the power. Can we do better than that? Let us take a situation in which the separation is ten wavelengths (Fig. 29-7), which is more nearly comparable to the situation in which we experimented in the previous chapter, with separations of several wavelengths rather than a small fraction of a wavelength. Here the picture is quite different.

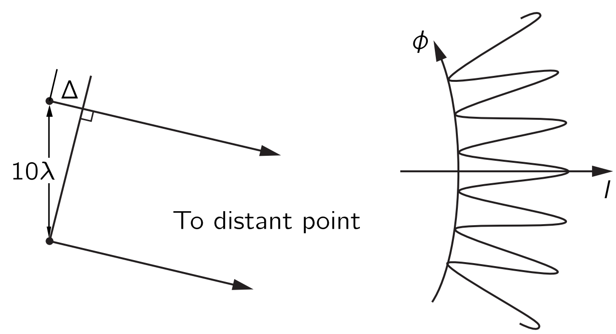

Fig. 29-7. The intensity pattern for two dipoles separated by 10λ.

If the oscillators are ten wavelengths apart (we take the in-phase case to make it easy), we see that in the E–W direction, they are in phase, and we get a strong intensity, four times what we would get if one of them were there alone. On the other hand, at a very small angle away, the arrival times differ by 180° and the intensity is zero. To be precise, if we draw a line from each oscillator to a distant point and the difference Δ in the two distances is $\lambda/2$, half an oscillation, then they will be out of phase. So this first null occurs when that happens. (The figure is not drawn to scale; it is only a rough sketch.) This means that we do indeed have a very sharp beam in the direction we want, because if we just move over a little bit we lose all our intensity. Unfortunately for practical purposes, if we were thinking of making a radio broadcasting array and we doubled the distance Δ, then we would be a whole cycle out of phase, which is the same as being exactly *in* phase again! Thus we get many successive maxima and minima, just as we found with the $2\frac{1}{2}\lambda$ spacing in Chapter 28.

Now how can we arrange to get rid of all these extra maxima, or "lobes," as they are called? We could get rid of the unwanted lobes in a rather interesting

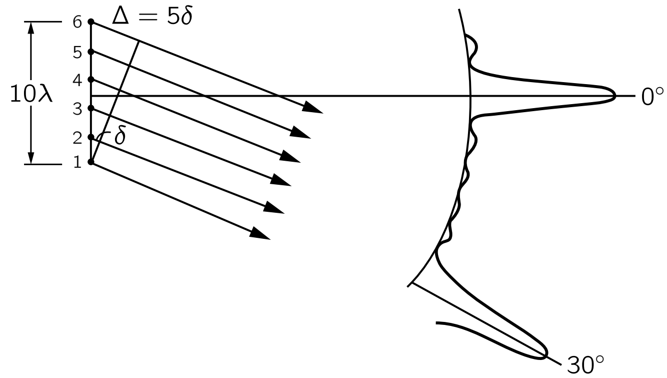

Fig. 29-8. A six-dipole antenna array and part of its intensity pattern.

way. Suppose that we were to place another set of antennas between the two that we already have. That is, the outside ones are still 10λ apart, but between them, say every 2λ, we have put another antenna, and we drive them all in phase. There are now six antennas, and if we looked at the intensity in the E–W direction, it would, of course, be much higher with six antennas than with one. The field would be six times and the intensity thirty-six times as great (the square of the field). We get 36 units of intensity in that direction. Now if we look at neighboring points, we find a zero as before, roughly, but if we go farther, to where we used to get a big "bump," we get a much smaller "bump" now. Let us try to see why.

The reason is that although we might expect to get a big bump when the distance Δ is exactly equal to the wavelength, it is true that dipoles 1 and 6 are then in phase and are cooperating in trying to get some strength in that direction. But numbers 3 and 4 are roughly $\frac{1}{2}$ a wavelength out of phase with 1 and 6, and although 1 and 6 push together, 3 and 4 push together too, but in opposite phase. Therefore there is very little intensity in this direction—but there is something; it does not balance exactly. This kind of thing keeps on happening; we get very little bumps, and we have the strong beam in the direction where we want it. But in this particular example, something else will happen: namely, since the distance between successive dipoles is 2λ, it is possible to find an angle where the distance δ between *successive dipoles* is exactly one wavelength, so that the effects from all of them are in phase again. Each one is delayed relative to the next one by 360°, so they all come back in phase, and we have another strong beam in that direction! It is easy to avoid this in practice because it is possible to put the dipoles closer than one wavelength apart. If we put in more antennas, closer than one wavelength apart, then this cannot happen. But the fact that this *can* happen at certain angles, if the spacing is bigger than one wavelength, is a very interesting and useful phenomenon in other applications—not radio broadcasting, but in *diffraction gratings*.

29-5 The mathematics of interference

Now we have finished our analysis of the phenomena of dipole radiators qualitatively, and we must learn how to analyze them quantitatively. To find the effect of two sources at some particular angle in the most general case, where the two oscillators have some intrinsic relative phase α from one another and the strengths A_1 and A_2 are not equal, we find that we have to add two cosines having the same frequency, but with different phases. It is very easy to find this phase difference; it is made up of a delay due to the difference in distance, and the intrinsic, built-in phase of the oscillation. Mathematically, we have to find the sum R of two waves: $R = A_1 \cos(\omega t + \phi_1) + A_2 \cos(\omega t + \phi_2)$. How do we do it?

It is really very easy, and we presume that we already know how to do it. However, we shall outline the procedure in some detail. First, we can, if we are clever with mathematics and know enough about cosines and sines, simply work it out. The easiest such case is the one where A_1 and A_2 are equal, let us say they are both equal to A. In those circumstances, for example (we could call this the trigonometric method of solving the problem), we have

$$R = A[\cos(\omega t + \phi_1) + \cos(\omega t + \phi_2)]. \tag{29.9}$$

Once, in our trigonometry class, we may have learned the rule that

$$\cos A + \cos B = 2 \cos \tfrac{1}{2}(A + B) \cos \tfrac{1}{2}(A - B). \tag{29.10}$$

If we know that, then we can immediately write R as

$$R = 2A \cos \tfrac{1}{2}(\phi_1 - \phi_2) \cos(\omega t + \tfrac{1}{2}\phi_1 + \tfrac{1}{2}\phi_2). \tag{29.11}$$

So we find that we have an oscillatory wave with a new phase and a new amplitude. In general, the result *will* be an oscillatory wave with a new amplitude A_R, which we may call the resultant amplitude, oscillating at the same frequency but with a phase difference ϕ_R, called the resultant phase. In view of this, our particular case has the following result: that the resultant amplitude is

$$A_R = 2A\cos\tfrac{1}{2}(\phi_1 - \phi_2), \tag{29.12}$$

and the resultant phase is the average of the two phases, and we have completely solved our problem.

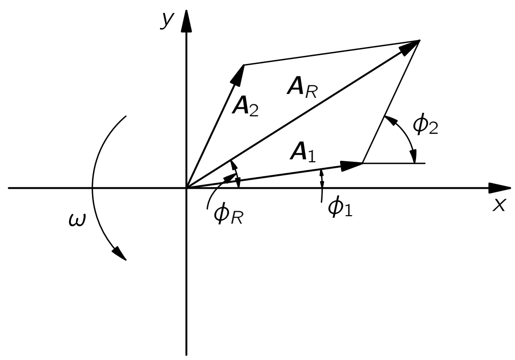

Fig. 29-9. A geometrical method for combining two cosine waves. The entire diagram is thought of as rotating counterclockwise with angular frequency ω.

Now suppose that we cannot remember that the sum of two cosines is twice the cosine of half the sum times the cosine of half the difference. Then we may use another method of analysis which is more geometrical. Any cosine function of ωt can be considered as the horizontal projection of a *rotating vector*. Suppose there were a vector $\boldsymbol{A}_1$ of length A_1 rotating with time, so that its angle with the horizontal axis is $\omega t + \phi_1$. (We shall leave out the ωt in a minute, and see that it makes no difference.) Suppose that we take a snapshot at the time $t = 0$, although, in fact, the picture is rotating with angular velocity ω (Fig. 29-9). The projection of $\boldsymbol{A}_1$ along the horizontal axis is precisely $A_1\cos(\omega t + \phi_1)$. Now at $t = 0$ the second wave could be represented by another vector, $\boldsymbol{A}_2$, of length A_2 and at an angle ϕ_2, and also rotating. They are both rotating with the same angular velocity ω, and therefore the *relative* positions of the two are fixed. The system goes around like a rigid body. The horizontal projection of $\boldsymbol{A}_2$ is $A_2\cos(\omega t + \phi_2)$. But we know from the theory of vectors that if we add the two vectors in the ordinary way, by the parallelogram rule, and draw the resultant vector $\boldsymbol{A}_R$, the x-component of the resultant is the sum of the x-components of the other two vectors. That solves our problem. It is easy to check that this gives the correct result for the special case we treated above, where $A_1 = A_2 = A$. In this case, we see from Fig. 29-9 that $\boldsymbol{A}_R$ lies midway between $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ and makes an angle $\frac{1}{2}(\phi_2 - \phi_1)$ with each. Therefore we see that $A_R = 2A\cos\frac{1}{2}(\phi_2 - \phi_1)$, as before. Also, as we see from the triangle, the phase of $\boldsymbol{A}_R$, as it goes around, is the average angle of $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ when the two amplitudes are equal. Clearly, we can also solve for the case where the amplitudes are not equal, just as easily. We can call that the *geometrical* way of solving the problem.

There is still another way of solving the problem, and that is the *analytical* way. That is, instead of having actually to draw a picture like Fig. 29-9, we can write something down which says the same thing as the picture: instead of drawing the vectors, we write a *complex number* to represent each of the vectors. The real parts of the complex numbers are the actual physical quantities. So in our particular case the waves could be written in this way: $A_1e^{i(\omega t+\phi_1)}$ [the real part of this is $A_1\cos{(\omega t + \phi_1)}$] and $A_2e^{i(\omega t+\phi_2)}$. Now we can add the two:

$$R = A_1e^{i(\omega t+\phi_1)} + A_2e^{i(\omega t+\phi_2)} = (A_1e^{i\phi_1} + A_2e^{i\phi_2})e^{i\omega t} \tag{29.13}$$

or

$$\hat{R} = A_1e^{i\phi_1} + A_2e^{i\phi_2} = A_Re^{i\phi_R}. \tag{29.14}$$

This solves the problem that we wanted to solve, because it represents the result as a complex number of magnitude A_R and phase ϕ_R.

To see how this method works, let us find the amplitude A_R which is the "length" of $\hat{R}$. To get the "length" of a complex quantity, we always multiply the quantity by its complex conjugate, which gives the length squared. The complex conjugate is the same expression, but with the sign of the i's reversed. Thus we have

$$A_R^2 = (A_1e^{i\phi_1} + A_2e^{i\phi_2})(A_1e^{-i\phi_1} + A_2e^{-i\phi_2}). \tag{29.15}$$

In multiplying this out, we get $A_1^2 + A_2^2$ (here the e's cancel), and for the cross terms we have

$$A_1A_2(e^{i(\phi_1-\phi_2)} + e^{i(\phi_2-\phi_1)}).$$

Now

$$e^{i\theta} + e^{-i\theta} = \cos\theta + i\sin\theta + \cos\theta - i\sin\theta.$$

That is to say, $e^{i\theta} + e^{-i\theta} = 2\cos\theta$. Our final result is therefore

$$A_R^2 = A_1^2 + A_2^2 + 2A_1A_2\cos{(\phi_2 - \phi_1)}. \tag{29.16}$$

As we see, this agrees with the length of $\boldsymbol{A}_R$ in Fig. 29-9, using the rules of trigonometry.

Thus the sum of the two effects has the intensity A_1^2 we would get with one of them alone, plus the intensity A_2^2 we would get with the other one alone, plus a correction. This correction we call the *interference effect.* It is really only the difference between what we get simply by adding the intensities, and what actually happens. We call it interference whether it is positive or negative. (Interference in ordinary language usually suggests opposition or hindrance, but in physics we often do not use language the way it was originally designed!) If the interference term is positive, we call that case *constructive* interference, horrible though it may sound to anybody other than a physicist! The opposite case is called *destructive* interference.

Now let us see how to apply our general formula (29.16) for the case of two oscillators to the special situations which we have discussed qualitatively. To apply this general formula, it is only necessary to find what phase difference, $\phi_1 - \phi_2$, exists between the signals arriving at a given point. (It depends only on the phase difference, of course, and not on the phase itself.) So let us consider the case where the two oscillators, of equal amplitude, are separated by some distance d and have an intrinsic relative phase α. (When one is at phase zero, the phase of the other is α.) Then we ask what the intensity will be in some azimuth direction θ from the E–W line. [Note that this is *not* the same θ as appears in (29.1). We are torn between using an unconventional symbol like $\not{U}$, or the conventional symbol θ (Fig. 29-10).] The phase relationship is found by noting that the difference in distance from P to the two oscillators is $d\sin\theta$, so that the phase difference contribution from this is the number of wavelengths in $d\sin\theta$, multiplied by 2π. (Those who are more sophisticated might want to multiply the wave number k, which is the rate of change of phase with distance, by $d\sin\theta$;

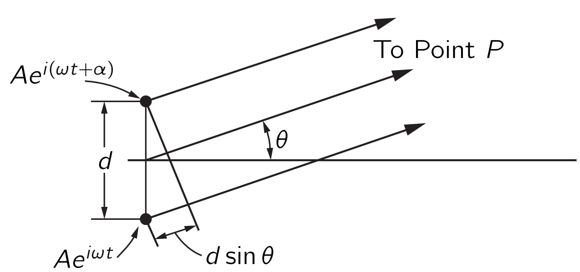

Fig. 29-10. Two oscillators of equal amplitude, with a phase difference α between them.

it is exactly the same.) The phase difference due to the distance difference is thus $2\pi d \sin\theta/\lambda$, but, due to the timing of the oscillators, there is an additional phase α. So the phase difference at arrival would be

$$\phi_2 - \phi_1 = \alpha + 2\pi d \sin\theta/\lambda. \tag{29.17}$$

This takes care of all the cases. Thus all we have to do is substitute this expression into (29.16) for the case $A_1 = A_2$, and we can calculate all the various results for two antennas of equal intensity.

Now let us see what happens in our various cases. The reason we know, for example, that the intensity is 2 at 30° in Fig. 29-5 is the following: the two oscillators are $\frac{1}{2}\lambda$ apart, so at 30°, $d\sin\theta = \lambda/4$. Thus $\phi_2 - \phi_1 = 2\pi\lambda/4\lambda = \pi/2$, and so the interference term is zero. (We are adding two vectors at 90°.) The result is the hypotenuse of a 45° right-angle triangle, which is $\sqrt{2}$ times the unit amplitude; squaring it, we get twice the intensity of one oscillator alone. All the other cases can be worked out in this same way.

I have included two more sections from the next chapter, but stop your reading here for now.

Use what you learned in Sections 5 and 6 of Churchill, Brown, and Verhey to do the first problem on the following page.

21

Electromagnetic Radiation: Interference

Refer to *The Feynman Lectures on Physics*, Vol. I, Chapters 28 and 29.

21.1 Interpret the following two problems in complex numbers geometrically, and show that the absolute value of A in each case is as given:

(a)

$$A = re^{i\theta/2} + re^{-i\theta/2}$$
$$|A| = 2r\cos(\theta/2)$$

(b)

$$A = \sum_{n=0}^{N} re^{in\theta}$$
$$|A| = r\frac{\sin\left(\frac{N+1}{2}\theta\right)}{\sin(\theta/2)}$$

21.2 A charge q traverses a circular path of radius a at an angular velocity ω.

(a) Evaluate the electric field $\boldsymbol{E}(t)$ at a great distance R from the charge, at an angle θ with respect to the axis of the circular path.

(b) Find the intensity of the radiation $I(\theta)$ at a great distance R on the axis ($\theta = 0$) and in the plane of the circle ($\theta = \pi/2$).

Assume that $\omega a \ll c$ and $a \ll R$.

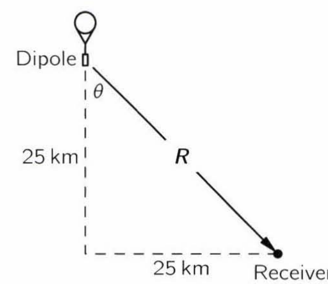

Figure 21-1

21.3 The power per unit area delivered by an electromagnetic wave is proportional to the mean-square electric field strength.

(a) If the total power radiated by an oscillating charge is P_{total}, how much power P falls on a unit area normal to the radius vector $\boldsymbol{R}$ at an angle θ with respect to the axis of oscillation?

(b) Evaluate P in W m^{-2} for a vertically oriented dipole suspended from a cosmic ray radiosonde balloon at an altitude of 25 km and at a horizontal distance of 25 km from the receiver, as shown in Fig. 21-1, if the transmitter is radiating 0.5 W total.

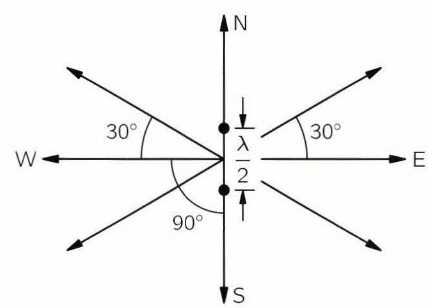

Figure 21-2

21.4 Two vertical antennas are arranged as shown in Fig. 21-2 and are driven in phase. The antennas are driven so that one would, if alone, radiate a certain intensity I_0 in all horizontal directions, and the other, an intensity $2I_0$. What should be the observed intensity I in the various directions shown in the figure?

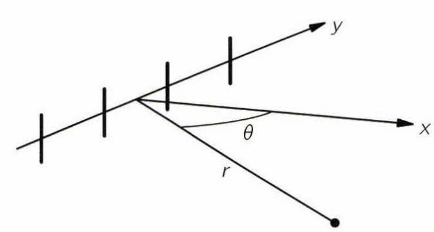

Figure 21-3

21.5 Four identical dipole radiators are aligned parallel to one another and are equally spaced along a line at a distance 2.50 cm apart. They are driven at a frequency of 3.00×10^9 Hz and are phased so that, starting from one end, each successive dipole lags the preceding one by 90°. Find the intensity pattern of the radiation $I(\theta)$ at a great distance in the equatorial plane (perpendicular to the dipole axes), and sketch this function in the polar coordinates shown in Fig. 21-3. Such a diagram is called the *radiation pattern* or *lobe pattern* of an antenna system.

30

Diffraction

30-1 The resultant amplitude due to n equal oscillators

This chapter is a direct continuation of the previous one, although the name has been changed from *Interference* to *Diffraction*. No one has ever been able to define the difference between interference and diffraction satisfactorily. It is just a question of usage, and there is no specific, important physical difference between them. The best we can do, roughly speaking, is to say that when there are only a few sources, say two, interfering, then the result is usually called interference, but if there is a large number of them, it seems that the word diffraction is more often used. So, we shall not worry about whether it is interference or diffraction, but continue directly from where we left off in the middle of the subject in the last chapter.

Thus we shall now discuss the situation where there are n equally spaced oscillators, all of equal amplitude but different from one another in phase, either because they are driven differently in phase, or because we are looking at them at an angle such that there is a difference in time delay. For one reason or another, we have to add something like this:

$$R = A[\cos\omega t + \cos(\omega t + \phi) + \cos(\omega t + 2\phi) + \cdots + \cos(\omega t + (n-1)\phi)], \tag{30.1}$$

where ϕ is the phase difference between one oscillator and the next one, as seen in a particular direction. Specifically, $\phi = \alpha + 2\pi d \sin\theta/\lambda$. Now we must add all the terms together. We shall do this geometrically. The first one is of length A, and it has zero phase. The next is also of length A and it has a phase equal to ϕ. The next one is again of length A and it has a phase equal to 2ϕ, and so on. So we are evidently going around an equiangular polygon with n sides (Fig. 30-1).

Now the vertices, of course, all lie on a circle, and we can find the net amplitude most easily if we find the radius of that circle. Suppose that Q is the center of

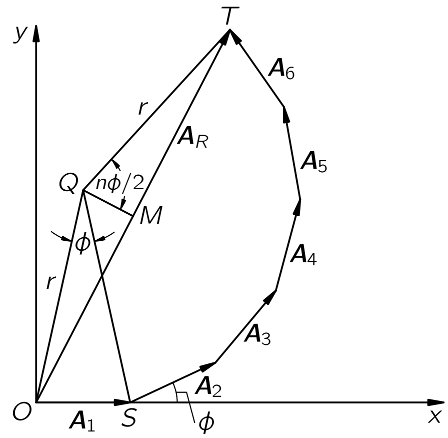

Fig. 30-1. The resultant amplitude of $n = 6$ equally spaced sources with net successive phase differences ϕ.

the circle. Then we know that the angle OQS is just a phase angle ϕ. (This is because the radius QS bears the same geometrical relation to $\boldsymbol{A}_2$ as QO bears to $\boldsymbol{A}_1$, so they form an angle ϕ between them.) Therefore the radius r must be such that $A = 2r\sin\phi/2$, which fixes r. But the large angle OQT is equal to $n\phi$, and we thus find that $A_R = 2r\sin n\phi/2$. Combining these two results to eliminate r, we get

$$A_R = A\,\frac{\sin n\phi/2}{\sin\phi/2}. \tag{30.2}$$

The resultant intensity is thus

$$I = I_0\,\frac{\sin^2 n\phi/2}{\sin^2\phi/2}. \tag{30.3}$$

Now let us analyze this expression and study some of its consequences. In the first place, we can check it for $n = 1$. It checks: $I = I_0$. Next, we check it for $n = 2$: writing $\sin\phi = 2\sin\phi/2\cos\phi/2$, we find that $A_R = 2A\cos\phi/2$, which agrees with (29.12).

Now the idea that led us to consider the addition of several sources was that we might get a much stronger intensity in one direction than in another; that the nearby maxima which would have been present if there were only two sources will have gone down in strength. In order to see this effect, we plot the curve that comes from (30.3), taking n to be enormously large and plotting the region near $\phi = 0$. In the first place, if ϕ is exactly 0, we have 0/0, but if ϕ is infinitesimal, the ratio of the two sines squared is simply n^2, since the sine and the angle are approximately equal. Thus the intensity of the maximum of the curve is equal to n^2 times the intensity of one oscillator. That is easy to see, because if they are all in phase, then the little vectors have no relative angle and all n of them add up so the amplitude is n times, and the intensity n^2 times, stronger.

As the phase ϕ increases, the ratio of the two sines begins to fall off, and the first time it reaches zero is when $n\phi/2 = \pi$, because $\sin\pi = 0$. In other words, $\phi = 2\pi/n$ corresponds to the first minimum in the curve (Fig. 30-2). In terms of what is happening with the arrows in Fig. 30-1, the first minimum occurs when all the arrows come back to the starting point; that means that the total accumulated angle in all the arrows, the total phase difference between the first and last oscillator, must be 2π to complete the circle.

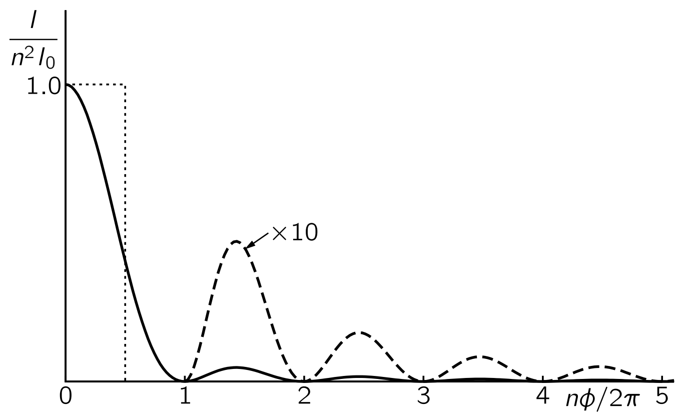

Fig. 30-2. The intensity as a function of phase angle for a large number of oscillators of equal strength.

Now we go to the next maximum, and we want to see that it is really much smaller than the first one, as we had hoped. We shall not go precisely to the maximum position, because both the numerator and the denominator of (30.3) are variant, but $\sin\phi/2$ varies quite slowly compared with $\sin n\phi/2$ when n is large, so when $\sin n\phi/2 = 1$ we are very close to the maximum. The next maximum of $\sin^2 n\phi/2$ comes at $n\phi/2 = 3\pi/2$, or $\phi = 3\pi/n$. This corresponds to the arrows having traversed the circle one and a half times. On putting $\phi = 3\pi/n$ into the formula to find the size of the maximum, we find that $\sin^2 3\pi/2 = 1$ in the numerator (because that is why we picked this angle), and in the denominator we have $\sin^2 3\pi/2n$. Now if n is sufficiently large, then this angle is very small and the sine is equal to the angle; so for all practical purposes, we can put $\sin 3\pi/2n = 3\pi/2n$. Thus we find that the intensity at this maximum is $I = I_0(4n^2/9\pi^2)$. But n^2I_0 was the maximum intensity, and so we have $4/9\pi^2$ times the maximum intensity, which is about 0.045, less than 5 percent, of the maximum intensity! Of course there are decreasing intensities farther out. So we have a very sharp central maximum with very weak subsidiary maxima on the sides.

It is possible to prove that the area of the whole curve, including all the little bumps, is equal to $2\pi nI_0$, or twice the area of the dotted rectangle in Fig. 30-2.

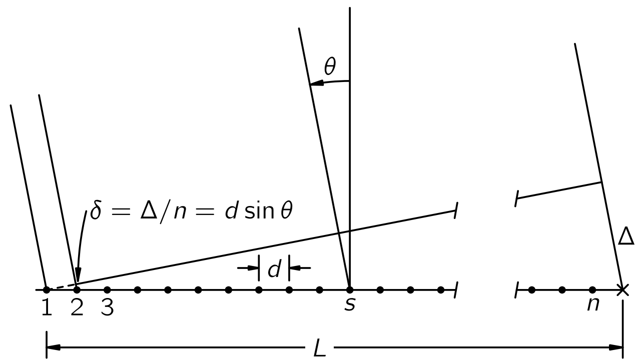

Fig. 30-3. A linear array of n equal oscillators, driven with phases $\alpha_s = s\alpha$.

Now let us consider further how we may apply Eq. (30.3) in different circumstances, and try to understand what is happening. Let us consider our sources to be all on a line, as drawn in Fig. 30-3. There are n of them, all spaced by a distance d, and we shall suppose that the intrinsic relative phase, one to the next, is α. Then if we are observing in a given direction θ from the normal, there is an additional phase $2\pi d \sin\theta/\lambda$ because of the time delay between each successive two, which we talked about before. Thus

$$\begin{aligned}\phi &= \alpha + 2\pi d \sin\theta/\lambda \\ &= \alpha + kd\sin\theta.\end{aligned} \tag{30.4}$$

First, we shall take the case $\alpha = 0$. That is, all oscillators are in phase, and we want to know what the intensity is as a function of the angle θ. In order to find out, we merely have to put $\phi = kd\sin\theta$ into formula (30.3) and see what happens. In the first place, there is a maximum when $\phi = 0$. That means that when all the oscillators are in phase there is a strong intensity in the direction $\theta = 0$. On the other hand, an interesting question is, where is the first minimum? That occurs when $\phi = 2\pi/n$. In other words, when $2\pi d\sin\theta/\lambda = 2\pi/n$, we get the first minimum of the curve. If we get rid of the 2π's so we can look at it a little better, it says that

$$nd\sin\theta = \lambda. \tag{30.5}$$

Now let us understand physically why we get a minimum at that position. nd is the total length L of the array. Referring to Fig. 30-3, we see that $nd\sin\theta = L\sin\theta = \Delta$. What (30.5) says is that when Δ is equal to *one wavelength*, we get a minimum. Now why do we get a minimum when $\Delta = \lambda$? Because the contributions of the various oscillators are then uniformly distributed in phase from $0°$ to $360°$. The arrows (Fig. 30-1) are going around a whole circle—we are adding equal vectors in all directions, and such a sum is zero. So when we have an angle such that $\Delta = \lambda$, we get a minimum. That is the first minimum.

There is another important feature about formula (30.3), which is that if the angle ϕ is increased by any multiple of 2π, it makes no difference to the formula. So we will get other strong maxima at $\phi = 2\pi$, 4π, 6π, and so forth. Near each of these great maxima the pattern of Fig. 30-2 is repeated. We may ask ourselves, what is the geometrical circumstance that leads to these other great maxima? The condition is that $\phi = 2\pi m$, where m is any integer. That is, $2\pi d\sin\theta/\lambda = 2\pi m$. Dividing by 2π, we see that

$$d\sin\theta = m\lambda. \tag{30.6}$$

This looks like the other formula, (30.5). No, that formula was $nd\sin\theta = \lambda$. The difference is that here we have to look at the *individual sources*, and when we say $d\sin\theta = m\lambda$, that means that we have an angle θ such that $\delta = m\lambda$. In other words, each source is now contributing a certain amount, and successive ones are out of phase by a whole multiple of $360°$, and therefore are contributing *in phase*, because out of phase by $360°$ is the same as being in phase. So they all contribute in phase and produce just as good a maximum as the one for $m = 0$ that we discussed before. The subsidiary bumps, the whole shape of the pattern, is just like the one near $\phi = 0$, with exactly the same minima on each side, etc. Thus such an array will send beams in various directions—each beam having a strong central maximum and a certain number of weak "side lobes." The various strong beams are referred to as the zero-order beam, the first-order beam, etc., according to the value of m. m is called the *order* of the beam.

We call attention to the fact that if d is less than λ, Eq. (30.6) can have no solution except $m = 0$, so that if the spacing is too small there is only one possible beam, the zero-order one centered at $\theta = 0$. (Of course, there is also a beam in the opposite direction.) In order to get subsidiary great maxima, we must have the spacing d of the array greater than one wavelength.

30-2 The diffraction grating

In technical work with antennas and wires it is possible to arrange that all the phases of the little oscillators, or antennas, are equal. The question is whether and how we can do a similar thing with light. We cannot at the present time literally make little optical-frequency radio stations and hook them up with infinitesimal wires and drive them all with a given phase. But there is a very easy way to do what amounts to the same thing.

Suppose that we had a lot of parallel wires, equally spaced at a spacing d, and a radiofrequency source very far away, practically at infinity, which is generating an electric field which arrives at each one of the wires at the same phase (it is so far away that the time delay is the same for all of the wires). (One can work out cases with curved arrays, but let us take a plane one.) Then the external electric field will drive the electrons up and down in each wire. That is, the field which is coming from the original source will shake the electrons up and down, and in moving, these represent *new generators*. This phenomenon is called scattering: a light wave from some source can induce a motion of the electrons in a piece of material, and these motions generate their own waves. Therefore all that is necessary is to set up a lot of wires, equally spaced, drive them with a radiofrequency source far away, and we have the situation that we want, without a whole lot of special wiring. If the incidence is normal, the phases will be equal, and we will get exactly the circumstance we have been discussing. Therefore, if the wire spacing is greater than the wavelength, we will get a strong intensity of scattering in the normal direction, and in certain other directions given by (30.6).

This can also be done with light! Instead of wires, we use a flat piece of glass and make notches in it such that each of the notches scatters a little differently than the rest of the glass. If we then shine light on the glass, each one of the notches will represent a source, and if we space the lines very finely, but not closer than a wavelength (which is technically almost impossible anyway), then we would expect a miraculous phenomenon: the light not only will pass straight through, but there will also be a strong beam at a finite angle, depending on the spacing of the notches! Such objects have actually been made and are in common use—they are called *diffraction gratings*.

In one of its forms, a diffraction grating consists of nothing but a plane glass sheet, transparent and colorless, with scratches on it. There are often several hundred scratches to the millimeter, *very* carefully arranged so as to be equally spaced. The effect of such a grating can be seen by arranging a projector so as to throw a narrow, vertical line of light (the image of a slit) onto a screen. When we put the grating into the beam, with its scratches vertical, we see that the line is still there but, in addition, on each side we have *another* strong patch of light which is *colored*. This, of course, is the slit image spread out over a wide angular range, because the angle θ in (30.6) depends upon λ, and lights of different colors, as we know, correspond to different frequencies, and therefore different wavelengths. The longest visible wavelength is red, and since $d\sin\theta = \lambda$, that requires a larger θ. And we do, in fact, find that red is at a greater angle out from the central image! There should also be a beam on the other side, and indeed we see one on the screen. Then, there might be another solution of (30.6) when $m = 2$. We do see that there is something vaguely there—very weak—and there are even other beams beyond.

We have just argued that all these beams ought to be of the same strength, but we see that they actually are not and, in fact, not even the first ones on the right and left are equal! The reason is that the grating has been carefully built to do just this. How? If the grating consists of very fine notches, infinitesimally wide, spaced evenly, then all the intensities would indeed be equal. But, as a matter of fact, although we have taken the simplest case, we could also have considered an array of *pairs* of antennas, in which each member of the pair has a certain strength and some relative phase. In this case, it is possible to get intensities which are different in the different orders. A grating is often made with little "sawtooth" cuts instead of little symmetrical notches. By carefully arranging the, "sawteeth," more light may be sent into one particular order of spectrum than into the others. In a practical grating, we would like to have as much light as possible in one of the orders. This may seem a complicated point to bring in, but it is a very clever thing to do, because it makes the grating more useful.

So far, we have taken the case where all the phases of the sources are equal. But we also have a formula for ϕ when the phases differ from one to the next by an angle α. That requires wiring up our antennas with a slight phase shift between each one. Can we do that with light? Yes, we can do it very easily, for suppose that there were a source of light at infinity, *at an angle* such that the light is coming in at an angle θ_{in}, and let us say that we wish to discuss the scattered beam, which is leaving at an angle θ_{out} (Fig. 30-4). The θ_{out} is the

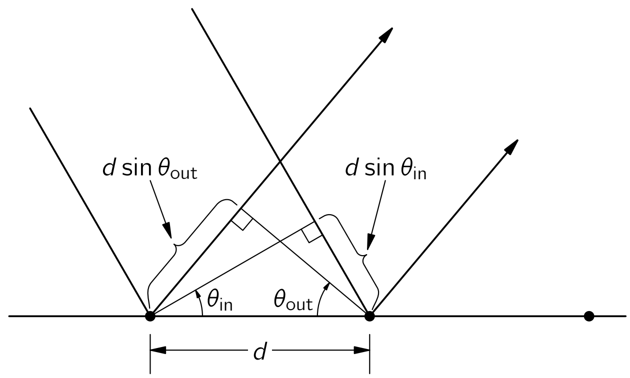

Fig. 30-4. The path difference for rays scattered from adjacent rulings of a grating is $d \sin\theta_{out} - d \sin\theta_{in}$.

same θ as we have had before, but the θ_{in} is merely a means for arranging that the phase of each source is different: the light coming from the distant driving source first hits one scratch, then the next, then the next, and so on, with a phase shift from one to the other, which, as we see, is $\alpha = -2\pi d \sin\theta_{in}/\lambda$. Therefore we have the formula for a grating in which light both comes in and goes out at an angle:

$$\phi = 2\pi d \sin\theta_{out}/\lambda - 2\pi d \sin\theta_{in}/\lambda. \tag{30.7}$$

Let us try to find out where we get strong intensity in these circumstances. The condition for strong intensities is, of course, that ϕ should be a multiple of 2π. There are several interesting points to be noted.

One case of rather great interest is that which corresponds to $m = 0$, where d is less than λ; in fact, this is the only solution. In this case we see that $\sin\theta_{out} = \sin\theta_{in}$, which means that the light comes out in the *same direction* as the light which was exciting the grating. We might think that the light "goes right through." No, it is *different light* that we are talking about. The light that goes right through is from the original source; what we are talking about is the new light *which is generated by scattering.* It turns out that the scattered light is going in the same direction as the original light, in fact it can interfere with it—a feature which we will study later.

There is another solution for this same case. For a given θ_{in}, θ_{out} may be the *supplement* of θ_{in}. So not only do we get a beam in the same direction as the incoming beam but also one in another direction, which, if we consider it carefully, is such that the *angle of incidence is equal to the angle of scattering.* This we call the reflected beam.

So we begin to understand the basic machinery of reflection: the light that comes in generates motions of the atoms in the reflector, and the reflector then regenerates a *new wave*, and one of the solutions for the direction of scattering, the *only* solution if the spacing of the scatterers is small compared with one wavelength, is that the angle at which the light comes out is equal to the angle at which it comes in!

Next, we discuss the special case when $d \to 0$. That is, we have just a solid piece of material, so to speak, but of finite length. In addition, we want the phase shift from one scatterer to the next to go to zero. In other words, we put more and more antennas between the other ones, so that each of the phase differences is getting smaller, but the number of antennas is increasing in such a way that the total phase difference, between one end of the line and the other, is constant. Let us see what happens to (30.3) if we keep the difference in phase $n\phi$ from one end to the other constant (say $n\phi = \Phi$), letting the number go to infinity and the phase shift ϕ of each one go to zero. But now ϕ is so small that $\sin\phi = \phi$, and if we also recognize $n^2 I_0$ as I_m, the maximum intensity at the center of the beam, we find

$$I = 4I_m \sin^2 \tfrac{1}{2}\Phi/\Phi^2. \tag{30.8}$$

This limiting case is what is shown in Fig. 30-2.

In such circumstances we find the same general kind of a picture as for finite spacing with $d > \lambda$; all the side lobes are practically the same as before, but there are no higher-order maxima. If the scatterers are all in phase, we get a maximum in the direction $\theta_{\text{out}} = 0$, and a minimum when the distance Δ is equal to λ, just as for finite d and n. So we can even analyze a *continuous* distribution of scatterers or oscillators, by using integrals instead of summing.

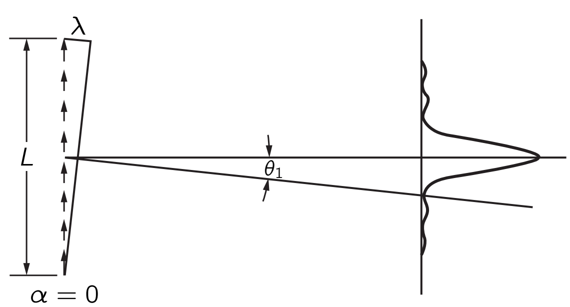

Fig. 30-5. The intensity pattern of a continuous line of oscillators has a single strong maximum and many weak "side lobes."

As an example, suppose there were a long line of oscillators, with the charge oscillating along the direction of the line (Fig. 30-5). From such an array the greatest intensity is perpendicular to the line. There is a little bit of intensity up and down from the equatorial plane, but it is very slight. With this result, we can handle a more complicated situation. Suppose we have a set of such lines, each producing a beam only in a plane perpendicular to the line. To find the intensity in various directions from a series of long wires, instead of infinitesimal wires, is the same problem as it was for infinitesimal wires, so long as we are in the central plane perpendicular to the wires; we just add the contribution from each of the long wires. That is why, although we actually analyzed only tiny antennas, we might as well have used a grating with long, narrow slots. Each of the long slots produces an effect only in its own direction, not up and down, but they are all set next to each other horizontally, so they produce interference that way.

Thus we can build up more complicated situations by having various distributions of scatterers in lines, planes, or in space. The first thing we did was to consider scatterers in a line, and we have just extended the analysis to strips; we can work it out by just doing the necessary summations, adding the contributions from the individual scatterers. The principle is always the same.

30-3 Resolving power of a grating

We are now in a position to understand a number of interesting phenomena. For example, consider the use of a grating for separating wavelengths. We noticed that the whole spectrum was spread out on the screen, so a grating can be used as an instrument for separating light into its different wavelengths. One of the interesting questions is: supposing that there were two sources of slightly different frequency, or slightly different wavelength, how close together in wavelength could they be such that the grating would be unable to tell that there were really two different wavelengths there? The red and the blue were clearly separated. But when one wave is red and the other is slightly redder, very close, how close can they be? This is called the *resolving power* of the grating, and one way of analyzing the problem is as follows. Suppose that for light of a certain color we happen to have the maximum of the diffracted beam occurring at a certain angle. If we vary the wavelength the phase $2\pi d \sin\theta/\lambda$ is different, so of course the maximum occurs at a different angle. That is why the red and blue are spread out. How different in angle must it be in order for us to be able to see it? If the two maxima are exactly on top of each other, of course we cannot see