Chapter 8. Radiation, Scattering, Interference, and Diffraction

This chapter continues the discussion of electromagnetic wave propagation, now focusing on the results of wave incidence on various objects of more complex shapes. Depending on the shape, the resulting wave pattern is called either "scattering", or "diffraction", or "interference". However, as the reader will see, the boundaries between these effects may be blurry, and their basic mathematical description may be conveniently based on the same key calculation – the electric-dipole radiation of a spherical wave by a localized source. Naturally, I will start the chapter from this calculation, deriving it from an even more general result – the "retarded-potential" solution of the Maxwell equations.

8.1. Retarded potentials

Let us start by finding the general solution of the macroscopic Maxwell equations (6.99) in a dispersion-free, linear, uniform, isotropic medium characterized by frequency-independent real ε and μ.[1] The easiest way to perform this calculation is to use the scalar (ϕ) and vector (**A**) potentials defined by Eqs. (6.7):

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \qquad \mathbf{B} = \nabla \times \mathbf{A}. \tag{8.1}$$

As was discussed in Sec. 6.8, by imposing upon the potentials the Lorenz gauge condition (6.117),

$$\nabla \cdot \mathbf{A} + \frac{1}{v^2}\frac{\partial \phi}{\partial t} = 0, \qquad \text{with } v^2 \equiv \frac{1}{\varepsilon\mu}, \tag{8.2}$$

which does not affect the fields **E** and **B**, the Maxwell equations are reduced to a pair of very similar, simple equations (6.118) for the potentials:

$$\nabla^2\phi - \frac{1}{v^2}\frac{\partial^2\phi}{\partial t^2} = -\frac{\rho}{\varepsilon}, \tag{8.3a}$$

$$\nabla^2\mathbf{A} - \frac{1}{v^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = -\mu\,\mathbf{j}. \tag{8.3b}$$

Let us find the general solution of these equations, for now thinking of the densities $\rho(\mathbf{r}, t)$ and $\mathbf{j}(\mathbf{r}, t)$ of the stand-alone charges and currents as known functions. (This will not prevent the results from being valid for the cases when $\rho(\mathbf{r}, t)$ and $\mathbf{j}(\mathbf{r}, t)$ should be calculated self-consistently.) The idea of such a solution may be borrowed from electro- and magnetostatics. Indeed, for the stationary case ($\partial/\partial t = 0$), the solutions of Eqs. (8.3) are given by a ready generalization of, respectively, Eqs. (1.38) and (5.28) to a uniform, linear medium:

$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon}\int \rho(\mathbf{r}')\frac{d^3r'}{|\mathbf{r}-\mathbf{r}'|}, \tag{8.4a}$$

$$\mathbf{A}(\mathbf{r}) \equiv \frac{\mu}{4\pi}\int \mathbf{j}(\mathbf{r}')\frac{d^3r'}{|\mathbf{r}-\mathbf{r}'|}. \tag{8.4b}$$

[1] When necessary (e.g., at the discussion of the Cherenkov radiation in Sec. 10.5), it will be not too hard to generalize these results to a dispersive medium.

As we know, these expressions may be derived by, first, calculating the potential of a point source, and then using the linear superposition principle for a system of such sources.

Let us do the same for the time-dependent case, starting from the field induced by a time-dependent point charge at the origin:[2]

$$\rho(\mathbf{r},t) = q(t)\delta(\mathbf{r}), \tag{8.5}$$

In this case, Eq. (3a) is homogeneous everywhere but the origin:

$$\nabla^2\phi - \frac{1}{v^2}\frac{\partial^2\phi}{\partial t^2} = 0, \quad \text{for } r \neq 0. \tag{8.6}$$

Due to the spherical symmetry of the problem, it is natural to look for a spherically symmetric solution to this equation.[3] Thus, we may simplify the Laplace operator correspondingly (as was repeatedly done earlier in this course), so Eq. (6) becomes

$$\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right]\phi = 0, \quad \text{for } r \neq 0. \tag{8.7}$$

By introducing a new variable $\chi(r, t) \equiv r\phi(r, t)$, Eq. (7) is reduced to the 1D wave equation

$$\left(\frac{\partial^2}{\partial r^2} - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right)\chi = 0, \quad \text{for } r \neq 0. \tag{8.8}$$

From discussions in Chapter 7,[4] we know that its general solution may be represented as

$$\chi(r,t) = \chi_{\text{out}}\left(t - \frac{r}{v}\right) + \chi_{\text{in}}\left(t + \frac{r}{v}\right), \tag{8.9}$$

where χ_{in} and χ_{out} are (so far) arbitrary functions of one variable. The physical sense of $\phi_{\text{out}} = \chi_{\text{out}}/r$ is a spherical wave propagating from our source (located at $r = 0$) to outer space, i.e. exactly the solution we are looking for. On the other hand, $\phi_{\text{in}} = \chi_{\text{in}}/r$ describes a spherical wave that could be created by some distant spherically-symmetric source, that converged exactly on our charge located at the origin – evidently not the effect we want to consider here. Discarding this term, and returning to $\phi = \chi/r$, we get

$$\phi(r,t) = \frac{1}{r}\chi_{\text{out}}\left(t - \frac{r}{v}\right), \quad \text{for } r \neq 0. \tag{8.10}$$

In order to calculate the function χ_{out}, let us consider the solution (10) at distances r so small ($r << vt$) that the time-derivative term in Eq. (3a), with the right-hand side (5),

$$\nabla^2\phi - \frac{1}{v^2}\frac{\partial^2\phi}{\partial t^2} = -\frac{q(t)}{\varepsilon}\delta(\mathbf{r}), \tag{8.11}$$

[2] Admittedly, this expression does *not* satisfy the continuity equation (4.5), but this deficiency will be corrected imminently, at the linear superposition stage – see Eq. (17) below.

[3] Let me emphasize that this is *not* the general solution to Eq. (6). For example, it does not describe the possible waves created by other sources, that pass by the considered charge $q(t)$. However, such fields are irrelevant to our current task: to calculate the field *induced* by the charge $q(t)$. The solution becomes general when it is integrated (as it will be) over all relevant charges.

[4] See also CM Sec. 6.3.

is much smaller than the spatial derivative term (which diverges at $r \to 0$) . Then Eq. (11) is reduced to the Poisson equation, whose solution (4a), for the source (5), is

$$\phi(r \to 0, t) = \frac{q(t)}{4\pi\varepsilon r}. \tag{8.12}$$

Now requiring the two solutions, Eqs. (10) and (12), to coincide at $r << vt$, we get $\chi_{out}(t) = q(t)/4\pi\varepsilon r$, so Eq. (10) becomes

$$\phi(r,t) = \frac{1}{4\pi\varepsilon r} q\left(t - \frac{r}{v}\right). \tag{8.13}$$

Just as was repeatedly done in statics, this result may be readily generalized for the arbitrary position $\mathbf{r}'$ of the point charge:

$$\rho(\mathbf{r},t) = q(t)\delta(\mathbf{r} - \mathbf{r}') \equiv q(t)\delta(\mathbf{R}), \tag{8.14}$$

where R is the distance between the field observation point $\mathbf{r}$ and the source position point $\mathbf{r}'$, i.e. the length of the vector,

$$\mathbf{R} \equiv \mathbf{r} - \mathbf{r}', \tag{8.15}$$

connecting these points – see Fig. 1.

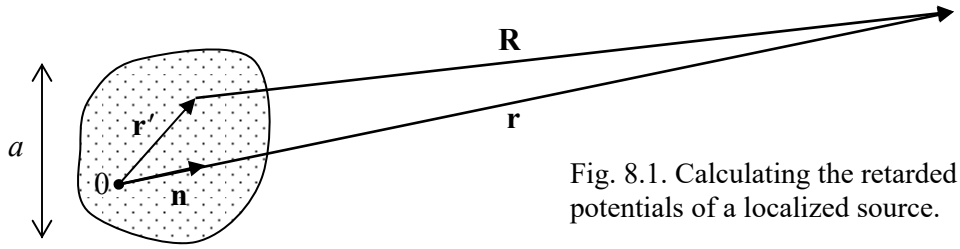


Fig. 8.1. Calculating the retarded potentials of a localized source.

Obviously, now Eq. (13) becomes

$$\phi(\mathbf{r},t) = \frac{1}{4\pi\varepsilon R} q\left(t - \frac{R}{v}\right). \tag{8.16}$$

Finally, we may use the linear superposition principle to write, for the arbitrary charge distribution,

Retarded scalar potential

$$\phi(\mathbf{r},t) = \frac{1}{4\pi\varepsilon}\int \rho\left(\mathbf{r}', t - \frac{R}{v}\right)\frac{d^3r'}{R}, \tag{8.17a}$$

where the integration is extended over all charges of the system under analysis. Solving Eq. (4b) absolutely similarly, for the vector potential we get[5]

Retarded vector potential

$$\mathbf{A}(\mathbf{r},t) = \frac{\mu}{4\pi}\int \mathbf{j}\left(\mathbf{r}', t - \frac{R}{v}\right)\frac{d^3r'}{R}. \tag{8.17b}$$

[5] As should be clear from the analogy of Eqs. (17) with their stationary forms (4), which were discussed, respectively, in Chapters 1 and 5, in the Gaussian units the retarded potential formulas are valid with the coefficient $1/4\pi$ dropped in Eq. (17a), and replaced with $1/c$ in Eq. (17b).

(Now nothing prevents the functions $\rho(\mathbf{r}, t)$ and $\boldsymbol{j}(\mathbf{r}, t)$ from satisfying the continuity equation.)

The solutions expressed by Eqs. (17) are traditionally called the *retarded potentials*, the name signifying the fact that the observed fields are "retarded" (in the meaning "delayed") in time by $\Delta t = R/v$ relative to the source variations – physically, because of the finite speed v of the electromagnetic wave propagation. Note that, very remarkably, these simple expressions are *exact* solutions of the macroscopic Maxwell equations (again, in a uniform, linear, dispersion-free medium) for an *arbitrary* distribution of stand-alone charges and currents. They also may be considered as the *general* solutions of these equations, provided that the integration has been extended over all field sources in the Universe – or at least over those ones that affect our observations.

Note also that due to the mathematical similarity of the microscopic and macroscopic Maxwell equations, Eqs. (17) are valid, with the coefficient replacement $\varepsilon \to \varepsilon_0$ and $\mu \to \mu_0$, for the exact, rather than the macroscopic fields, provided that the functions $\rho(\mathbf{r}, t)$ and $\mathbf{j}(\mathbf{r}, t)$ describe not only stand-alone but all charges and currents in the system. (Alternatively, this statement may be formulated as the validity of Eqs. (17), with the same coefficient replacement, in free space.)

Finally, note that Eqs. (17) may be plugged into Eqs. (1), giving (after an explicit differentiation) the so-called *Jefimenko equations*[6] for fields $\mathbf{E}$ and $\mathbf{B}$ – similar in structure to Eqs. (17), but more cumbersome. Conceptually, the existence of such equations is good news, because they are free from the gauge ambiguity pertinent to the potentials ϕ and $\mathbf{A}$. However, the practical value of these explicit expressions for the fields is not overly high: for all applications I am aware of, it is easier to use Eqs. (17) to calculate the particular expressions for the potentials first, and only then calculate the fields from Eqs. (1). Let me now present an (arguably, the most important) example of this approach.

8.2. Electric dipole radiation

Consider again the problem that was discussed in electrostatics (Sec. 3.1), namely the field of a localized source with linear dimensions $a << r$ (see Fig. 1 again), but now with time-dependent charge and/or current distributions. Using all the arguments of that discussion, in particular the condition expressed by Eq. (3.1), $r' << r$, we may apply the Taylor expansion (3.3), truncated to two leading terms,

$$f(\mathbf{R}) = f(\mathbf{r}) - \mathbf{r}' \cdot \nabla f(\mathbf{r}) + \dots , \tag{8.18}$$

to the scalar function $f(\mathbf{R}) \equiv R$ (for which $\nabla f(\mathbf{r}) = \nabla R = \mathbf{n}$, where $\mathbf{n} \equiv \mathbf{r}/r$ is the unit vector directed toward the observation point – see Fig. 1) to approximate the distance R as

$$R \approx r - \mathbf{r}' \cdot \mathbf{n} . \tag{8.19}$$

In each of the retarded potential formulas (17), R participates in two places: in the denominator and in the source's time argument. If ρ and $\mathbf{j}$ change in time on the scale $\sim 1/\omega$, where ω is some characteristic frequency, then any change of the argument $(t - R/v)$ on that time scale, for example due to a change of R on the spatial scale $\sim v/\omega = 1/k$, may substantially change these functions. Thus, the expansion (19) may be applied to R in the argument $(t - R/v)$ only if $ka << 1$, i.e. if the system's size a

[6] They were published by O. D. Jefimenko only in 1966, but the Fourier representation of the same result was obtained much earlier (in 1912) by G. A. Scott.

is much smaller than the radiation wavelength $\lambda = 2\pi/k$. On the other hand, the function $1/R$ changes relatively slowly, and for it even the first term of the expansion (19) gives a good approximation as soon as $a << r, R$. In the latter approximation alone, Eq. (17a) yields

$$\phi(\mathbf{r},t) \approx \frac{1}{4\pi\varepsilon r}\int \rho\left(\mathbf{r}', t-\frac{R}{v}\right)d^3r' \equiv \frac{1}{4\pi\varepsilon r}Q\left(t-\frac{R}{v}\right), \tag{8.20}$$

where $Q(t)$ is the net electric charge of the localized system. Due to the charge conservation, this charge cannot change with time, so the approximation (20) describes just a static Coulomb field of our localized source, rather than a radiated wave.

Let us, however, apply a similar approximation to the vector potential (17b):

$$\mathbf{A}(\mathbf{r},t) \approx \frac{\mu}{4\pi r}\int \mathbf{j}\left(\mathbf{r}', t-\frac{R}{v}\right)d^3r'. \tag{8.21}$$

According to Eq. (5.87), the right-hand side of this expression vanishes in statics, but in dynamics, this is no longer true. For example, if the current is due to some non-relativistic motion[7] of a system of point charges q_k, we can write

$$\int \mathbf{j}(\mathbf{r}',t)d^3r' = \sum_k q_k \dot{\mathbf{r}}_k(t) = \frac{d}{dt}\sum_k q_k \mathbf{r}_k(t) \equiv \dot{\mathbf{p}}(t), \tag{8.22}$$

where $\mathbf{p}(t)$ is the dipole moment of the localized system, defined by Eq. (3.6). Now, after the integration, we may keep only the first term of the approximation (19) in the argument $(t - R/v)$ as well, getting

$$\mathbf{A}(\mathbf{r},t) \approx \frac{\mu}{4\pi r}\dot{\mathbf{p}}\left(t-\frac{r}{v}\right), \qquad \text{for } a << R, \frac{1}{k}. \tag{8.23}$$

Let us analyze what exactly this result describes. The second of Eqs. (1) allows us to calculate the magnetic field by the spatial differentiation of $\mathbf{A}$. At large distances $r >> \lambda$ (i.e. in the so-called *far-field zone*), where Eq. (23) describes a locally-plane wave, the dominating contribution to this derivative is given by the dipole moment factor:

Far-field wave

$$\mathbf{B}(\mathbf{r},t) = \frac{\mu}{4\pi r}\nabla\times\dot{\mathbf{p}}\left(t-\frac{r}{v}\right) = -\frac{\mu}{4\pi rv}\mathbf{n}\times\ddot{\mathbf{p}}\left(t-\frac{r}{v}\right). \tag{8.24}$$

This expression means that the magnetic field, at the observation point, is perpendicular to the vectors $\mathbf{n}$ and (the retarded value of) $\ddot{\mathbf{p}}$, and its magnitude is

$$B = \frac{\mu}{4\pi rv}\ddot{p}\left(t-\frac{r}{v}\right)\sin\Theta, \qquad \text{i.e. } H = \frac{1}{4\pi rv}\ddot{p}\left(t-\frac{r}{v}\right)\sin\Theta, \tag{8.25}$$

where Θ is the angle between those two vectors – see Fig. 2.[8]

[7] For relativistic particles, moving with velocities of the order of speed of light, one has to be more careful. As the result, I will postpone the discussion of their radiation until Chapter 10, i.e. until after the detailed discussion of special relativity in Chapter 9.

[8] From the first of Eqs. (1) for the electric field, in the first approximation (23), we would get $-\partial\mathbf{A}/\partial t = -(1/4\pi\varepsilon vr)\ddot{\mathbf{p}}(t - r/v) = -(Z/4\pi r)\ddot{\mathbf{p}}(t - r/v)$. The transverse component of this vector (see Fig. 2) is the proper electric field $\mathbf{E}$

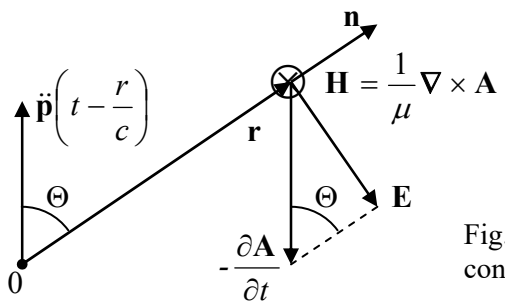


Fig. 8.2. Far-fields of a localized source, contributing to its electric dipole radiation.

The most important feature of this result is that the time-dependent field decreases very slowly (only as $1/r$) with the distance from the source, so the radial component of the corresponding Poynting vector (7.9b),[9]

$$S_r = ZH^2 = \frac{Z}{(4\pi vr)^2}\left[\ddot{p}\left(t-\frac{r}{v}\right)\right]^2 \sin^2\Theta\,, \tag{8.26}$$

Instant power density

drops only as $1/r^2$, i.e. the full instant power $\mathscr{P}$ of the emitted wave,

$$\mathscr{P} \equiv \oint_{4\pi} S_r r^2 d\Omega = \frac{Z}{(4\pi v)^2}\ddot{p}^2 2\pi\int_0^{\pi}\sin^3\Theta\, d\Theta = \frac{Z}{6\pi v^2}\ddot{p}^2\,. \tag{8.27}$$

Larmor formula

does not depend on the distance from the source – as it should for radiation. [10]

This is the famous *Larmor formula*[11] for the *electric dipole radiation*; it is the dominating component of radiation by a localized system of charges – unless $\ddot{\mathbf{p}} = 0$. Please notice its angular dependence: the radiation vanishes at the axis of the retarded vector $\ddot{\mathbf{p}}$ (where $\Theta = 0$), and reaches its maximum in the plane normal to that axis.

In order to find the average power, Eq. (27) has to be averaged over a sufficiently long time. In particular, if the source is monochromatic, $\mathbf{p}(t) = \mathrm{Re}[\mathbf{p}_\omega \exp\{-i\omega t\}]$, with a time-independent vector amplitude $\mathbf{p}_\omega$, such averaging may be carried out just over one period, giving an extra factor 2 in the denominator:

$$\overline{\mathscr{P}} = \frac{Z\omega^4}{12\pi v^2}|p_\omega|^2. \tag{8.28}$$

Average radiation power

The easiest application of this formula is to a point charge oscillating, with frequency ω, along a straight line (which we may take for the z-axis), with amplitude a. In this case, $\mathbf{p} = qz(t)\mathbf{n}_z = qa\,\mathrm{Re}[\exp\{-i\omega t\}]\mathbf{n}_z$, and if the charge velocity amplitude, $a\omega$, is much less than the electromagnetic wave's speed v, we may use Eq. (28) with $p_\omega = qa$, giving

---

$= Z\mathbf{H}\times\mathbf{n}$ of the radiated wave, while its longitudinal component is exactly compensated by $(-\nabla\phi)$ in the *next* term of the Taylor expansion of Eq. (17a) in small parameter $ka \sim a/\lambda << 1$.

[9] Note the "doughnut" dependence of S_r on the direction $\mathbf{n}$, frequently used to visualize the dipole radiation.

[10] In the Gaussian units, for free space ($v = c$), Eq. (27) reads $\mathscr{P} = (2/3c^3)\ddot{p}^2$.

[11] Named after Joseph Larmor, who was the first to derive this formula (in 1897) for the particular case of a single point charge q moving with acceleration $\ddot{\mathbf{r}}$, when $\ddot{\mathbf{p}} = q\ddot{\mathbf{r}}$.

$$\overline{\mathscr{P}} = \frac{Zq^2a^2\omega^4}{12\pi v^2}. \tag{8.29}$$

Applied to a classical picture of an electron (with $q = -e \approx 1.6\times10^{-19}$C), initially rotating about an atom's nucleus at an atomic distance $a \sim 10^{-10}$ m, Eq. (29) shows[12] that the energy loss due to the dipole radiation is so large that it would cause the electron to collapse on the nucleus in just $\sim10^{-11}$ s. In the beginning of the 1900s, this result was one of the main arguments for the development of quantum mechanics, which prevents such a collapse of electrons for their lowest-energy (ground) quantum state.

Another useful application of Eq. (28) is the radio wave radiation by a short, straight, symmetric antenna which is fed, for example, by a TEM transmission line such as a coaxial cable – see Fig. 3.

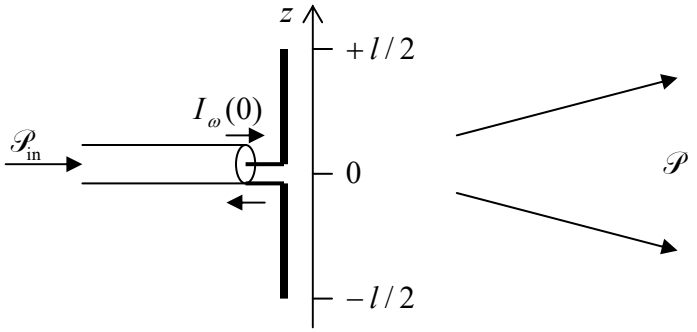


Fig. 8.3. The dipole antenna.

The exact solution of this problem is rather complicated because the law $I_\omega(z)$ of the current variation along the antenna's length should be calculated self-consistently with the distribution of the electromagnetic field induced by the current in the surrounding space. (Unfortunately, this fact is not mentioned in some textbooks.) However, the current should be largest in the feeding point (in Fig. 3, taken for $z = 0$) and vanish at the antenna's ends ($z = \pm l/2$), and hence we may guess that at $l << \lambda$, the linear function

$$I_\omega(z) = I_\omega(0)\left(1 - \frac{2}{l}|z|\right), \tag{8.30}$$

should be a good approximation of the actual distribution – as it indeed is. Now we can use the continuity equation $\partial Q/\partial t = I$, i.e. $-i\omega Q_\omega = I_\omega$, to calculate the complex amplitude $Q_\omega(z) = iI_\omega(z)\mathrm{sgn}(z)/\omega$ of the electric charge $Q(z, t) = \mathrm{Re}[Q_\omega \exp\{-i\omega t\}]$ of the wire's segment $[0, z]$, and from it, the amplitude of the charge's linear density:

$$\lambda_\omega(z) \equiv \frac{dQ_\omega(z)}{d|z|} = -i\frac{2I_\omega(0)}{\omega l}\mathrm{sgn}\, z\,. \tag{8.31}$$

From here, the dipole moment's amplitude is

$$p_\omega = 2\int_0^{l/2}\lambda_\omega(z)z\,dz = -i\frac{I_\omega(0)}{2\omega}l\,, \tag{8.32}$$

so Eq. (28) yields

[12] Actually, the formula needs a numerical coefficient adjustment to account for the electron's orbital (rather than linear) motion – the task left for the reader's exercise. However, this adjustment does not affect the order-of-magnitude estimate given above.

$$\overline{\mathscr{P}} = Z\frac{\omega^4}{12\pi v^2}\frac{|I_\omega(0)|^2}{4\omega^2}l^2 = \frac{Z(kl)^2}{24\pi}\frac{|I_\omega(0)|^2}{2}, \tag{8.33}$$

where $k = \omega/v$. The analogy between this result and the dissipation power $\mathscr{P} = \mathrm{Re}Z\,|\,I_\omega|^2/2$ in a lumped linear circuit element, enables the interpretation of the first fraction in the last form of Eq. (33) as the real part of the antenna's impedance:

$$\mathrm{Re}\,Z_A = Z\frac{(kl)^2}{24\pi}, \tag{8.34}$$

as felt by the transmission line.

According to Eq. (7.118), the wave traveling along the line toward the antenna is fully radiated, i.e. not reflected back, only if Z_A equals to Z_W of the line. As we know from Sec. 7.5 (and the solution of the related problems), for typical TEM lines, $Z_W \sim Z_0$, while Eq. (34), which is only valid in the limit $kl << 1$, shows that for the radiation into free space ($Z = Z_0$), $\mathrm{Re}Z_A$ is much less than Z_0. Hence to reach the impedance matching condition $Z_W = Z_A$, the antenna's length should be increased – as a more involved theory shows, to $l \approx \lambda/2$. However, in many cases, practical considerations make short antennas necessary. The example most often met nowadays is the cell phone antennas, which use frequencies close to 1 or 2 GHz, with free-space wavelengths λ between 15 and 30 cm, i.e. much larger than the phone size.[13] The quadratic dependence of the antenna's efficiency on l, following from Eq. (34), explains why every millimeter counts in the design of such antennas, and why the designs are carefully optimized using software packages for the (virtually exact) numerical solution of the Maxwell equations for the specific shape of the antenna and other phone parts.[14]

To conclude this section, let me note that if the wave source is not monochromatic, so $\mathbf{p}(t)$ should be represented as a Fourier series,

$$\mathbf{p}(t) = \mathrm{Re}\sum_\omega \mathbf{p}_\omega e^{-i\omega t}, \tag{8.35}$$

the terms corresponding to the interference of spectral components with different frequencies ω are averaged out at the time averaging of the Poynting vector, and the *average* radiated power is just a sum of contributions (28) from all substantial frequency components.

8.3. Wave scattering

The Larmor formula may be used as the basis of the theory of *scattering* – the phenomenon illustrated by Fig. 4. Generally, scattering is a complex problem. However, in many cases it allows the so-called *Born approximation*,[15] in which the scattered wave field's effect on the scattering object is assumed to be much weaker than that of the incident wave, and is neglected.

[13] The situation will be partly remedied by the planned transfer of wireless mobile technology to the next generations, with the signal frequencies gradually moving up.

[14] A partial list of popular software packages of this kind includes both publicly available codes such as Nec2 (whose various versions are available online, e.g., at http://www.qsl.net/4nec2/), and proprietary packages – such as *Momentum* from Agilent Technologies (now owned by Hewlett-Packard), *FEKO* from EM Software & Systems, and *XFdtd* from Remcom.

[15] Named after Max Born, one of the founding fathers of quantum mechanics. However, the basic idea of this approach was developed much earlier (in 1881) by Lord Rayleigh – born John William Strutt.

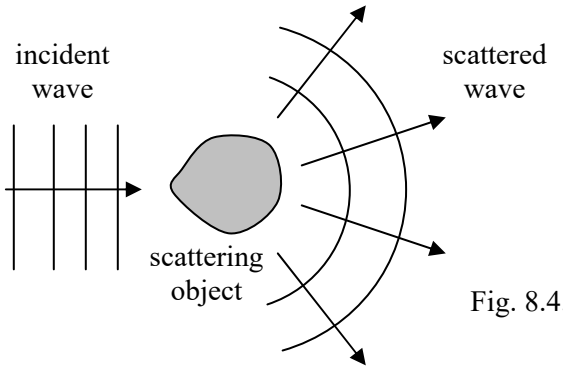


Fig. 8.4. Wave scattering (schematically).

As the first example of this approach, let us consider the scattering of a plane wave, propagating in free space ($Z = Z_0$, $v = c$), by an otherwise free[16] charged particle whose motion may be described by non-relativistic classical mechanics. (This requires, in particular, the incident wave to be not too powerful, so the speed of the charge's motion induced by the wave remains much lower than c.) As was already discussed at the derivation of Eq. (7.32), in this case, the magnetic component of the Lorentz force (5.10) is negligible in comparison with the force $\mathbf{F}_e = q\mathbf{E}$ exerted by the wave's electric field. Thus, assuming that the incident wave is linearly polarized along some axis x, the equation of the particle's motion in the Born approximation is just $m\ddot{x} = qE(t)$, so for the x-component $p_x = qx$ of its dipole moment we can write

$$\ddot{p} = q\ddot{x} = \frac{q^2}{m}E(t). \tag{8.36}$$

As we already know from Sec. 2, oscillations of the dipole moment lead to radiation of a wave with a wide angular distribution of intensity; in our case, this is the scattered wave – see Fig. 4. Its full power may be found by plugging Eq. (36) into Eq. (27):

$$\mathscr{P} = \frac{Z_0}{6\pi c^2}\ddot{p}^2 = \frac{Z_0 q^4}{6\pi c^2 m^2}E^2(t), \tag{8.37}$$

so for the average power, we get

$$\overline{\mathscr{P}} = \frac{Z_0 q^4}{12\pi c^2 m^2}\left|E_\omega\right|^2. \tag{8.38}$$

Since this power is proportional to the incident wave's intensity S, it is customary to characterize the scattering ability of an object by the ratio,

Full cross-section: definition

$$\sigma \equiv \frac{\overline{\mathscr{P}}}{\overline{S_{\text{incident}}}} \equiv \frac{\overline{\mathscr{P}}}{\left|E_\omega\right|^2/2Z_0}, \tag{8.39}$$

which has the dimensionality of area, and is called the *total cross-section* of scattering.[17] For this measure, Eq. (38) yields the famous result

[16] As Eq. (7.30) shows, this calculation is also valid for an oscillator with a low own frequency, $\omega_0 << \omega$.

[17] This definition parallels those accepted in the classical and quantum theories of *particle* scattering – see, e.g., respectively, CM Sec. 3.5 and QM Sec. 3.3.

$$\sigma = \frac{Z_0^2 q^4}{6\pi c^2 m^2} = \frac{\mu_0^2 q^4}{6\pi m^2}, \tag{8.40}$$

which is called the *Thomson scattering formula*,[18] especially when applied to an electron. This relation is most frequently represented in the form[19]

$$\sigma = \frac{8\pi}{3} r_c^2, \qquad \text{with } r_c \equiv \frac{q^2}{4\pi\varepsilon_0}\cdot\frac{1}{mc^2}. \tag{8.41}$$

Thomson scattering

This constant r_c is called the *classical radius of the particle* (or sometimes the "Thomson scattering length"); for the electron ($q = -e$, $m = m_e$) it is close to 2.82×10^{-15} m. Its possible interpretation is evident from Eq. (41) for r_c: at that distance between two similar particles, the potential energy $q^2/4\pi\varepsilon_0 r$ of their electrostatic interaction is equal to the particle's rest-mass energy mc^2.[20]

Now we have to go back and establish the conditions at which the Born approximation, when the field of the scattered wave is negligible, is indeed valid for a point-object scattering. Since the scattered wave's intensity described by Eq. (26) diverges at $r \to 0$ as $1/r^2$, according to the definition (39) of the cross-section, it may become comparable to S_{incident} at $r^2 \sim \sigma$. However, Eq. (38) itself is only valid if $r >> \lambda$, so the Born approximation does not lead to a contradiction only if

$$\sigma << \lambda^2. \tag{8.42}$$

For the Thompson scattering by an electron, this condition means $\lambda >> r_c \sim 3\times10^{-15}$ m and is fulfilled for all frequencies up to very hard γ-rays with photon energies ~100 MeV.

Possibly the most notable feature of the result (40) is its independence of the wave frequency. As it follows from its derivation, particularly from Eq. (37), this independence is intimately related to the unbound character of charge motion. For bound charges, say for electrons in gas molecules, this result is only valid if the wave frequency ω is much higher than the frequencies ω_j of most important quantum transitions. In the opposite limit, $\omega << \omega_j$, the result is dramatically different. Indeed, in this limit we may approximate the molecule's dipole moment by its static value (3.48):

$$\mathbf{p} = \alpha\mathbf{E}. \tag{8.43}$$

In the Born approximation, and in the absence of the molecular field effects mentioned in Sec. 3.3, **E** in this expression is just the incident wave's field, and we can use Eq. (28) to calculate the power of the wave scattered by a single molecule:

---

[18] Named after Sir Joseph John ("JJ") Thomson, the discoverer of the electron – and isotopes as well! He should not be confused with his son, G. P. Thomson, who discovered (simultaneously with C. Davisson and L. Germer) quantum-mechanical wave properties of the same electron.

[19] In the Gaussian units, this formula looks like $r_c = q^2/mc^2$ (giving, of course, the same numerical values: for the electron, $r_c \approx 2.82\times10^{-13}$ cm). This *classical* quantity should not be confused with the particle's *Compton wavelength* $\lambda_C \equiv 2\pi\hbar/mc$ (for the electron, close to 2.24×10^{-12} m), which naturally arises in *quantum* electrodynamics – see a brief discussion in the next chapter, and also QM Sec. 1.1.

[20] It is fascinating how smartly the *relativistic* expression mc^2 sneaked into the result (40)-(41), which was obtained using the *non-relativistic* equation (36) of the particle motion. This was possible because the calculation engaged electromagnetic waves, which propagate with the speed of light, and whose quanta (*photons*), as a result, may be frequently treated as relativistic (moreover, ultra-relativistic) particles – see the next chapter.

$$\overline{\mathcal{P}} = \frac{Z_0 \omega^4}{4\pi c^2} \alpha^2 \left|E_\omega\right|^2 . \tag{8.44}$$

Now, using the last form of the definition (39) of the cross-section, we get a very simple result,

$$\sigma = \frac{Z_0^2 \omega^4}{6\pi c^2} \alpha^2 , \tag{8.45}$$

showing that in contrast to Eq. (40), at low frequencies σ changes as fast as ω^4.

Now let us explore the effect of such *Rayleigh scattering* on wave propagation in a gas, with a relatively low volumic density n. We may expect (and will prove in the next section) that due to the randomness of molecule positions, the waves scattered by individual molecules may be treated as *incoherent* ones, so the total scattering power may be calculated just as the sum of those scattered by each molecule. We can use this fact to write the balance of the incident's wave intensity in a small volume dV of length (along the incident wave direction) dz, and area A across it. Since such a segment includes $ndV = nAdz$ molecules, and according to Eq. (39), each of them scatters power $S\sigma = \mathcal{P}\sigma/A$, the total scattered power is $n\mathcal{P}\sigma dz$; hence the incident power's change is

$$d\mathcal{P} \equiv -n\sigma\mathcal{P}\, dz . \tag{8.46}$$

Comparing this equation with the definition (7.213) of the wave attenuation constant, applied to the scattering, [21]

$$d\mathcal{P} \equiv -\alpha_{\text{scat}}\mathcal{P}\, dz . \tag{8.47}$$

we see that this effect gives the following contribution to attenuation: $\alpha_{\text{scat}} = n\sigma$. From here, using Eq. (3.50) to write $\alpha = \varepsilon_0(\kappa - 1)/n$, where κ is the dielectric constant, and Eq. (45) for σ, we get

Rayleigh scattering

$$\boxed{\alpha_{\text{scat}} = \frac{k^4}{6\pi n}(\kappa - 1)^2, \qquad \text{where } k \equiv \frac{2\pi}{\lambda_0} = \frac{\omega}{c}.} \tag{8.48}$$

This is the famous *Rayleigh scattering formula*, which in particular explains the colors of blue sky and red sunsets. Indeed, through the visible light spectrum, ω changes almost two-fold; as a result, the scattering of blue components of sunlight is an order of magnitude higher than that of its red components. For the air near the Earth's surface, $\kappa - 1 \approx 6\times10^{-4}$, and $n \sim 2.5\times10^{25}$ m^{-3} – see Sec. 3.3. Plugging these numbers into Eq. (48), we see that the effective length $l_{\text{scat}} \equiv 1/\alpha_{\text{scat}}$ of scattering is ~30 km for the blue light and ~200 km for the red light.[22] The effective thickness h of the Earth's atmosphere is ~10 km, so the Sun looks just a bit yellowish during most of the day. However, an elementary geometry shows that at sunset, the light has to pass the length $l \sim (R_E h)^{1/2} \approx 300$ km to reach an Earth-surface observer; as a result, the blue components of the Sun's light spectrum are almost completely scattered out, and even the red components are weakened very substantially.

[21] I am sorry for using the same letter (α) for both the molecular polarizability and the wave attenuation, but both notations are traditional. Hopefully, the subscript "scat" marking α in the latter meaning minimizes the possibility of confusion.

[22] These values are approximate because both n and $(\kappa - 1)$ vary through the atmosphere's thickness.

8.4. Interference and diffraction

Now let us discuss scattering by objects with a size of the order of, or even larger than λ. For such extended objects, the phase difference factors (neglected above) step in, leading in particular to the important effects of *interference* and *diffraction*. These effects show up not as much in the total power of the scattered radiation, as in its angular distribution. It is common to characterize this distribution by the *differential cross-section* defined as

$$\frac{d\sigma}{d\Omega} \equiv \frac{\overline{S}_r r^2}{S_{\text{incident}}}, \tag{8.49}$$

Differential cross-section

where r is the distance from the scatterer, at which the scattered wave is observed.[23] Both the definition and the notation may become clearer if we notice that according to Eq. (26), at large distances ($r >> a$), the numerator of the right-hand side of Eq. (49), and hence the differential cross-section as a whole, do not depend on r, and that its integral over the total solid angle $\Omega = 4\pi$ coincides with the total cross-section defined by Eq. (39):

$$\oint_{4\pi} \frac{d\sigma}{d\Omega} d\Omega = \frac{1}{S_{\text{incident}}} r^2 \oint_{4\pi} \overline{S}_r d\Omega = \frac{1}{S_{\text{incident}}} \oint_{r=\text{const}} \overline{S}_r d^2 r = \frac{\overline{\mathscr{P}}}{S_{\text{incident}}} \equiv \sigma . \tag{8.50}$$

For example, according to Eq. (26), the angular distribution of the radiation scattered by a single dipole is rather broad; in particular, in the quasistatic case (43), within the Born approximation,

$$\frac{d\sigma}{d\Omega} = \left(\frac{\alpha k^2}{4\pi\varepsilon_0} \right)^2 \sin^2 \Theta . \tag{8.51}$$

If the wave is scattered by a small dielectric body, with a characteristic size $a << \lambda$ (i.e., $ka << 1$), then all its parts re-radiate the incident wave coherently. Hence, we can calculate it similarly, just replacing the molecular dipole moment (43) with the total dipole moment of the object – see Eq. (3.45):

$$\mathbf{p} = \mathbf{P}V = (\kappa - 1)\varepsilon_0 \mathbf{E} V , \tag{8.52}$$

where $V \sim a^3$ is the body's volume. As a result, the differential cross-section may be obtained from Eq. (51) with the replacement $\alpha_{\text{mol}} \to (\kappa - 1)\varepsilon_0 V$:

$$\frac{d\sigma}{d\Omega} = \left(\frac{k^2 V}{4\pi} \right)^2 (\kappa - 1)^2 \sin^2 \Theta , \tag{8.53}$$

i.e. follows the same $\sin^2\Theta$ law.

The situation for extended objects, with at least one dimension of the order of (or larger than) the wavelength, is different: here we have to take into account the phase shifts between the wave's re-radiation by various parts of the body. Let us analyze this issue first for an arbitrary collection of similar point scatterers located at points $\mathbf{r}_j$. If the wave vector of the incident plane wave is $\mathbf{k}_0$, the wave's field has the phase factor $\exp\{i\mathbf{k}_0 \cdot \mathbf{r}\}$ – see Eq. (7.79). At the location $\mathbf{r}_j$ of the j^{th} scattering center, this factor equals $\exp\{i\mathbf{k}_0 \cdot \mathbf{r}_j\}$, defining the time dependence of the dipole vector $\mathbf{p}$, and hence of the scattered wave.

[23] Just as in the case of the total cross-section, this definition is also similar to that accepted at the *particle* scattering – see, e.g., CM Sec. 3.5 and QM Sec. 3.3.

According to Eq. (17), the scattered wave with a wave vector $\mathbf{k}$ (with $k = k_0$) acquires, on its way from the source point $\mathbf{r}_j$ to the observation point $\mathbf{r}$, an additional phase factor $\exp\{i\mathbf{k}\cdot(\mathbf{r} - \mathbf{r}_j)\}$, so the scattered wave field is proportional to

$$\exp\left\{i\mathbf{k}_0\cdot\mathbf{r}_j + i\mathbf{k}(\mathbf{r}-\mathbf{r}_j)\right\} \equiv e^{i\mathbf{k}\cdot\mathbf{r}}\exp\left\{-i(\mathbf{k}-\mathbf{k}_0)\cdot\mathbf{r}_j\right\}. \tag{8.54}$$

Since the first factor in the last expression does not depend on $\mathbf{r}_j$, to calculate the total scattering wave, it is sufficient to sum up the last phase factors, $\exp\{-i\mathbf{q}\cdot\mathbf{r}_j\}$, where the vector

$$\mathbf{q} \equiv \mathbf{k} - \mathbf{k}_0 \tag{8.55}$$

has the physical sense of the wave vector change at scattering.[24] It may look like the phase factor depends on our choice of the reference frame. However, according to Eq. (7.42), the average *intensity* of the scattered wave is proportional to $E_\omega E_\omega^*$, i.e. to the following real scalar function of the vector $\mathbf{q}$:

Scattering function

$$F(\mathbf{q}) = \left(\sum_j \exp\{-i\mathbf{q}\cdot\mathbf{r}_j\}\right)\left(\sum_{j'} \exp\{-i\mathbf{q}\cdot\mathbf{r}_{j'}\}\right)^* \equiv \sum_{j,j'} \exp\{i\mathbf{q}\cdot(\mathbf{r}_j - \mathbf{r}_{j'})\} = \left|I(\mathbf{q})\right|^2, \tag{8.56}$$

where the complex function

Phase sum

$$I(\mathbf{q}) \equiv \sum_j \exp\{-i\mathbf{q}\cdot\mathbf{r}_j\} \tag{8.57}$$

is called the *phase sum*, and may be calculated in any reference frame without affecting the final result given by Eq. (56).

So, besides the $\sin^2\Theta$ factor, the differential cross-section (49) of scattering by an extended object is also proportional to the scattering function (56). Its double-sum form is convenient to notice that for a system of *many* ($N >> 1$) similar but randomly located scatterers, only the terms with $j = j'$ accumulate at summation, so $F(\mathbf{q})$, and hence $d\sigma/d\Omega$, scale as N, rather than N^2 – thus justifying again our treatment of the Rayleigh scattering problem in the previous section.

Now let us apply Eq. (56) to a simple problem of just *two* similar small scatterers, separated by a fixed distance a:

$$F(\mathbf{q}) = \sum_{j,j'=1}^{2} \exp\{i\mathbf{q}\cdot(\mathbf{r}_j - \mathbf{r}_{j'})\} = 2 + \exp\{-iq_a a\} + \exp\{iq_a a\} = 2\left(1+\cos q_a a\right) = 4\cos^2\frac{q_a a}{2}, \tag{8.58}$$

where $q_a \equiv \mathbf{q}\cdot\mathbf{a}/a$ is the component of the vector $\mathbf{q}$ along the vector $\mathbf{a}$ connecting the scatterers. The apparent simplicity of this result may be a bit misleading because the mutual plane of the vectors $\mathbf{k}$ and $\mathbf{k}_0$ (and hence of the vector $\mathbf{q}$) does not necessarily coincide with the mutual plane of the vectors $\mathbf{k}_0$ and $\mathbf{E}_\omega$, so the *scattering angle* θ between $\mathbf{k}$ and $\mathbf{k}_0$ is generally different from $(\pi/2 - \Theta)$ – see Fig. 5. Moreover, the angle between the vectors $\mathbf{q}$ and $\mathbf{a}$ (within their common plane) is one more parameter independent of both θ and Θ. As a result, the angular dependence of the scattered wave's intensity (and hence $d\sigma/d\Omega$), which depends on all three angles, may be rather involved, but some of its details are irrelevant for the basic physics of interference/diffraction.

[24] In quantum mechanics, $\hbar\mathbf{q}$ has a very clear sense of the momentum transferred from the scattering object to the scattered particle (for example, a photon), and this terminology is sometimes smuggled even into classical electrodynamics texts.

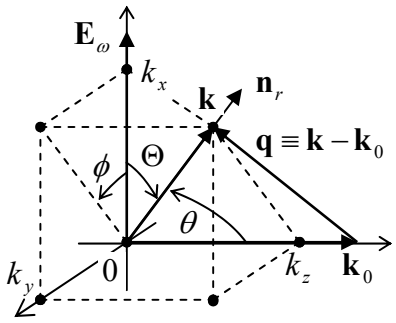


Fig. 8.5. The angles important for the general scattering problem.

This is why let me consider in detail only the simple cases when the vectors **k**, $\mathbf{k}_0$, and **a** all reside in the same plane, with $\mathbf{k}_0$ normal to **a** – see Fig. 6a.

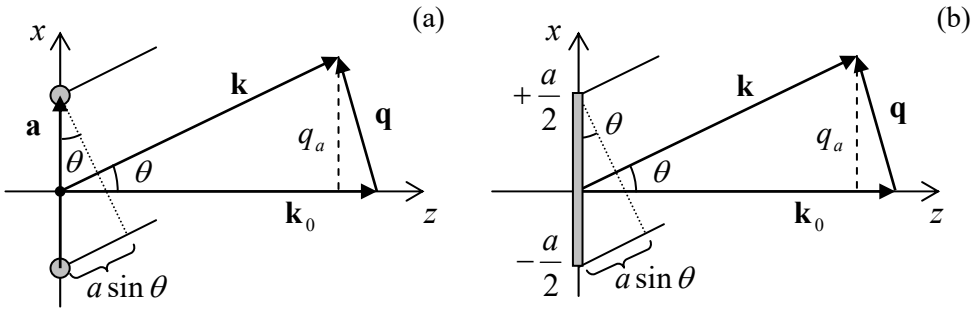


Fig. 8.6. The simplest cases of (a) interference and (b) diffraction.

In this case, $q_a = k\sin\theta$, and Eq. (58) is reduced to

$$F(\mathbf{q}) = 4\cos^2\frac{ka\sin\theta}{2}. \tag{8.59}$$

This function always has two maxima, at $\theta = 0$ and $\theta = \pi$, and, if the product ka is large enough, other maxima[25] at the special angles θ_n that satisfy the simple *Bragg condition*

$$ka\sin\theta_n = 2\pi m, \qquad \text{i.e. } a\sin\theta_n = n\lambda. \tag{8.60}$$

Bragg condition

As Fig. 6a shows, this condition may be readily understood as that of the in-phase addition (the *constructive interference*) of two coherent waves scattered from the two points, when the difference between their paths toward the observer, $a\sin\theta$, is equal to an integer number of wavelengths. At each such maximum, $F = 4$, due to the doubling of the wave amplitude and hence quadrupling its power.

If the distance between the point scatterers is large ($ka >> 1$), the first maxima (60) correspond to small scattering angles, $\theta << 1$. For this region, Eq. (59) is reduced to a simple periodic dependence of function F on the angle θ. Moreover, within the range of small θ, the wave polarization factor $\sin^2\Theta$ is virtually constant, so the angular dependence of the scattered wave's intensity, and hence of the differential cross-section, is also very simple:

$$\frac{d\sigma}{d\Omega} \propto F(\mathbf{q}) = 4\cos^2\frac{ka\theta}{2}. \tag{8.61}$$

Young's interference pattern

[25] In optics especially, such intensity maxima/minima patterns are called *interference fringes*.

This simple *interference pattern* is well known from Young's two-slit experiment.[26] (As will be discussed in the next section, the theoretical description of the two-slit experiment is more complex than that of the Born scattering, but is preferable experimentally because, at such scattering, the wave of intensity (61) has to be observed on the backdrop of a much stronger incident wave that propagates in almost the same direction, $\theta = 0$.)

A very similar analysis of scattering from $N > 2$ similar, equidistant scatterers, located along the same straight line shows that the positions (60) of the constructive interference maxima do not change (because the derivation of this condition is still applicable to each pair of adjacent scatterers), but the increase of N makes these peaks sharper and sharper. Leaving the quantitative analysis of this system for the reader's exercise, let me jump immediately to the limit $N \to 0$, in which we may ignore the scatterers' discreteness. The resulting pattern is similar to that at scattering by a continuous thin rod (see Fig. 6b), so let us first spell out the Born scattering formula for an arbitrary extended, continuous, uniform dielectric body. Transferring Eq. (56) from the sum to an integral, for the differential cross-section we get

$$\frac{d\sigma}{d\Omega} = \left(\frac{k^2}{4\pi}\right)^2 (\kappa - 1)^2 F(\mathbf{q}) \sin^2\Theta \equiv \left(\frac{k^2}{4\pi}\right)^2 (\kappa - 1)^2 \left|I(\mathbf{q})\right|^2 \sin^2\Theta, \tag{8.62}$$

where $I(\mathbf{q})$ now becomes the *phase integral*,[27]

Phase integral

$$I(\mathbf{q}) = \int_V \exp\{-i\mathbf{q}\cdot\mathbf{r}'\} d^3r', \tag{8.63}$$

with the dimensionality of volume.

Now we may return to the particular case of a thin rod (with both dimensions of the cross-section's area A much smaller than λ, but an arbitrary length a), otherwise keeping the same simple geometry as for two point scatterers – see Fig. 6b. In this case, the phase integral is just

Fraunhofer diffraction integral

$$I(\mathbf{q}) = A\int_{-a/2}^{+a/2} \exp\{-iq_a x')dx' = A\frac{\exp\{-iq_a a/2\} - \exp\{iq_a a/2\}}{-iq} \equiv V\frac{\sin\xi}{\xi}, \tag{8.64}$$

where $V = Aa$ is the volume of the rod, and ξ is the dimensionless argument defined as

$$\xi \equiv \frac{q_a a}{2} \equiv \frac{ka\sin\theta}{2}. \tag{8.65}$$

The fraction participating in the last form of Eq. (64) is met in physics so frequently that it has deserved the special name of the *sinc* (not "sync", please!) *function* (see Fig. 7):

[26] This experiment was described in 1803 by Thomas Young – one more universal genius of science, who also introduced the Young modulus in the elasticity theory (see, e.g., CM Chapter 7), besides numerous other achievements – including deciphering Egyptian hieroglyphs! It is fascinating that the first clear observation of wave interference was made as early as 1666 by another genius, Sir Isaac Newton, in the form of so-called *Newton's rings*. Unbelievably, Newton failed to give the most natural explanation of his observations – perhaps because he was vehemently opposed to the very idea of light as a wave, which was promoted in his times by others, notably by Christian Huygens. Due to Newton's enormous authority, only Young's two-slit experiments more than a century later have firmly established the wave picture of light – to be replaced by the dualistic wave/photon picture formalized by quantum electrodynamics (see, e.g., QM Ch. 9), in one more century.

[27] Since the observation point's position $\mathbf{r}$ does not participate in this formula explicitly, the prime sign in $\mathbf{r}'$ could be dropped, but I keep it as a reminder that the integral is taken over points $\mathbf{r}'$ of the *scattering object*.

$$\operatorname{sinc}\xi \equiv \frac{\sin\xi}{\xi}. \tag{8.66}$$

Sinc function

It vanishes at all points $\xi_n = \pi n$ with integer n, besides such point with $n = 0$: $\operatorname{sinc}\xi_0 \equiv \operatorname{sinc} 0 = 1$.

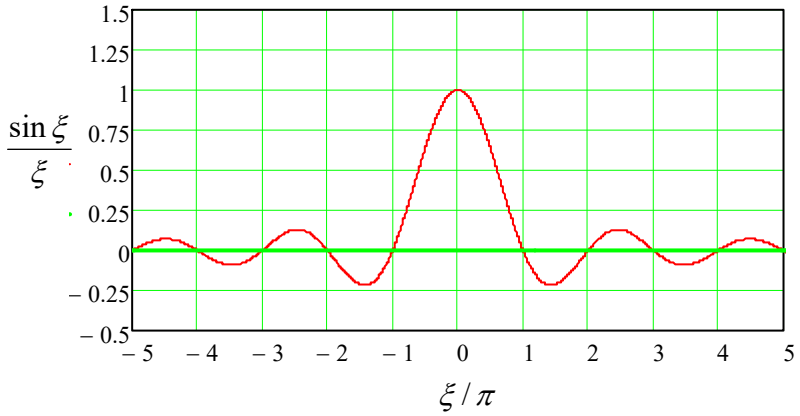


Fig. 8.7. The sinc function.

The function $F(\mathbf{q}) = V^2\operatorname{sinc}^2\xi$, given by Eq. (64) and plotted with the red line in Fig. 8, is called the *Fraunhofer diffraction pattern*.[28]

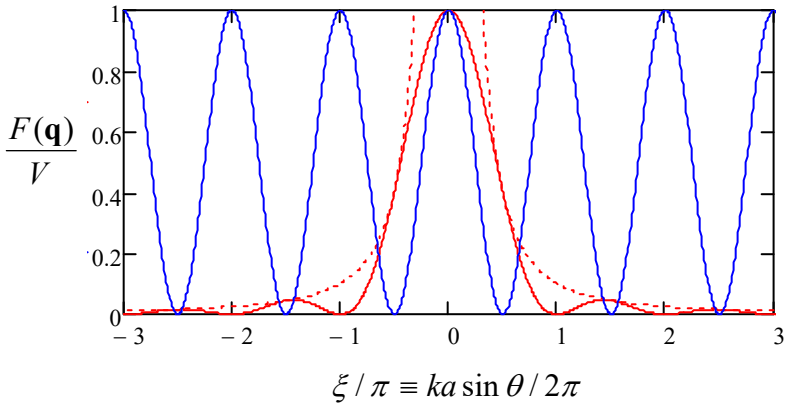


Fig. 8.8. The Fraunhofer diffraction pattern (solid red line) and its envelope $1/\xi^2$ (dashed red line). For comparison, the blue line shows the basic interference pattern $\cos^2\xi$ – cf. Eq. (61).

Note that it oscillates with the same argument's period $\Delta(ka\sin\theta) = 2\pi/ka$ as the interference pattern (61) from two-point scatterers (shown with the blue line in Fig. 8). However, at the interference, the scattered wave intensity vanishes at angles θ_n' that satisfy the condition

$$\frac{ka\sin\theta_n'}{2\pi} = n + \frac{1}{2}, \tag{8.67}$$

i.e. when the optical path difference $a\sin\alpha$ is equal to a semi-integer number of wavelengths $\lambda/2 = \pi/k$, and hence the two waves from the scatterers reach the observer in anti-phase – the so-called *destructive interference*. On the other hand, for the diffraction on a continuous rod the minima occur at a different set of scattering angles:

$$\frac{ka\sin\theta_n}{2\pi} = n, \tag{8.68}$$

[28] It is named after Joseph von Fraunhofer (1787-1826) – who invented the spectroscope, developed the diffraction grating (see below), and also discovered the dark *Fraunhofer lines* in the Sun's spectrum.

i.e. exactly where the two-point interference pattern has its maxima – please have a look at Fig. 8 again. The reason for this relation is that the wave diffraction on the rod may be considered as a simultaneous interference of waves from all its elementary fragments, and exactly at the observation angles when the rod edges give waves with phases shifted by $2\pi n$, the interior points of the rod give waves with all phases within this difference, with their algebraic sum equal to zero. As even more visible in Fig. 8, at the diffraction, the intensity oscillations are limited by a rapidly decreasing envelope function $1/\xi^2$ – while at the two-point interference, the oscillations retain the same amplitude. The reason for this fast decrease is that with each Fraunhofer diffraction period, a smaller and smaller fraction of the road gives an unbalanced contribution to the scattered wave.

If the rod's length is small ($ka << 1$, i.e. $a << \lambda$), then the sinc function's argument ξ is small at all scattering angles θ, so $I(\mathbf{q}) \approx V$, and Eq. (62) is reduced to Eq. (53). In the opposite limit, $a >> \lambda$, the first zeros of the function $I(\mathbf{q})$ correspond to very small angles θ, for which $\sin\Theta \approx 1$, so the differential cross-section is just

$$\frac{d\sigma}{d\Omega} = \left(\frac{k^2}{4\pi}\right)^2 (\kappa - 1)^2 \operatorname{sinc}^2 \frac{ka\theta}{2}, \tag{8.69}$$

i.e. Fig. 8 shows the scattering intensity as a function of the direction toward the observation point – if this point is within the plane containing the rod.

Finally, let us discuss a problem of large importance for applications: calculate the positions of the maxima of the interference pattern arising at the incidence of a plane wave on a very large 3D periodic system of point scatterers. For that, first of all, let us quantify the notion of 3D periodicity. The periodicity in one dimension is simple: the system we are considering (say, the positions of point scatterers) should be invariant with respect to the linear translation by some period a, and hence by any multiple sa of this period, where s is any integer. Anticipating the 3D generalization, we may require any of the possible *translation vectors* $\mathbf{R}$ to that the system is invariant, to be equal $s\mathbf{a}$, where the *primitive vector* $\mathbf{a}$ is directed along the (so far, the only) axis of the 1D system.

Now we are ready for the common definition of the 3D periodicity – as the invariance of the system with respect to the translation by any vector of the following set:

Bravais lattice

$$\mathbf{R} = \sum_{l=1}^{3} s_l \mathbf{a}_l, \tag{8.70}$$

where s_l are three independent integers, and $\{\mathbf{a}_l\}$ is a set of three linearly independent *primitive vectors*. The set of geometric points described by Eq. (70) is called the *Bravais lattice* (first analyzed in detail, circa 1850, by Auguste Bravais). Perhaps the most nontrivial feature of this relation is that the vectors $\mathbf{a}_l$ should not necessarily be orthogonal to each other. (That requirement would severely restrict the set of possible lattices and make it unsuitable for the description, for example, of many solid-state crystals.) For the scattering problem we are considering, we will assume that the position $\mathbf{r}_j$ of each point scatterer coincides with one of the points $\mathbf{R}$ of some Bravais lattice, with a given set of primitive vectors $\mathbf{a}_l$, so in the basic Eq. (57), the index j is coding the set of three integers $\{s_1, s_2, s_3\}$.

Now let us consider a similarly defined Bravais lattice, but in the reciprocal (wave-number) space, numbered by three independent integers $\{t_1, t_2, t_3\}$:

Reciprocal lattice

$$\mathbf{Q} = \sum_{m=1}^{3} t_m \mathbf{b}_m, \qquad \text{with } \mathbf{b}_m = 2\pi \frac{\mathbf{a}_{m''} \times \mathbf{a}_{m'}}{\mathbf{a}_m \cdot (\mathbf{a}_{m''} \times \mathbf{a}_{m'})}, \tag{8.71}$$

where in the last expression, the indices *m*, *m'*, and *m''* are all different. This is the so-called *reciprocal lattice*, which plays an important role in all physics of periodic structures, in particular in the quantum energy-band theory.[29] To reveal its most important property, and thus justify the above introduction of the primitive vectors $\mathbf{b}_m$, let us calculate the following scalar product:

$$\mathbf{R}\cdot\mathbf{Q} \equiv \sum_{l,m=1}^{3} s_l t_m \mathbf{a}_l \cdot \mathbf{b}_m \equiv 2\pi \sum_{l,m=1}^{3} s_l t_m \mathbf{a}_l \cdot \frac{\mathbf{a}_{m''}\times\mathbf{a}_{m'}}{\mathbf{a}_m\cdot\left(\mathbf{a}_{m''}\times\mathbf{a}_{m'}\right)} \equiv 2\pi \sum_{l,m=1}^{3} s_l t_k \frac{\mathbf{a}_l\cdot\left(\mathbf{a}_{m''}\times\mathbf{a}_{m'}\right)}{\mathbf{a}_m\cdot\left(\mathbf{a}_{m''}\times\mathbf{a}_{m'}\right)}. \quad (8.72)$$

Applying to the numerator of the last fraction the *operand rotation rule* of vector algebra,[30] we see that it is equal to zero if $l \neq m$, while for $l = m$ the whole fraction is evidently equal to 1. Thus the double sum (72) is reduced to a single sum:

$$\mathbf{R}\cdot\mathbf{Q} = 2\pi\sum_{l=1}^{3} s_l t_l = 2\pi\sum_{l=1}^{3} n_l\,, \quad (8.73)$$

where each of the products $n_l \equiv s_l t_l$ is an integer, and hence their sum,

$$n \equiv \sum_{l=1}^{3} n_l \equiv s_1 t_1 + s_2 t_2 + s_3 t_3, \quad (8.74)$$

is an integer as well, so the main property of the direct/reciprocal lattice couple is very simple:

$$\mathbf{R}\cdot\mathbf{Q} = 2\pi n, \qquad \text{and hence } \exp\{-i\mathbf{R}\cdot\mathbf{Q}\} = 1\,. \quad (8.75)$$

Now returning to the scattering function (56) for a Bravais lattice of point scatters, we see that if the vector $\mathbf{q} \equiv \mathbf{k} - \mathbf{k}_0$ coincides with *any* vector $\mathbf{Q}$ of the reciprocal lattice, then all terms of the phase sum (57) take their largest possible values (equal to 1), and hence the sum as the whole is largest as well, giving a constructive interference maximum. This equality, $\mathbf{q} = \mathbf{Q}$, where $\mathbf{Q}$ is given by Eq. (71), is called the *von Laue condition* (named after Max von Laue) of the constructive interference; it is, in particular, the basis of the whole field of the X-ray crystallography of solids and polymers – the main tool for revealing their atomic/molecular structure.[31]

In order to recast the von Laue condition in a more vivid geometric form, let us consider one of the vectors $\mathbf{Q}$ of the reciprocal lattice, corresponding to a certain integer *n* in Eq. (75), and notice that if that relation is satisfied for one point $\mathbf{R}$ of the direct Bravais lattice (70), i.e. for one set of the integers $\{s_1, s_2, s_3\}$, it is also satisfied for a 2D system of other integer sets, which may be parameterized, for example, by two integers S_1 and S_2:

$$s_1' = s_1 + S_1 t_3, \quad s_2' = s_2 + S_2 t_3, \quad s_3' = s_3 - S_1 t_1 - S_2 t_2\,. \quad (8.76)$$

Indeed, each of these sets has the same value of the integer *n*, defined by Eq. (74), as the original one:

$$n' \equiv s_1' t_1 + s_2' t_2 + s_3' t_3 \equiv \left(s_1 + S_1 t_3\right)t_1 + \left(s_2 + S_2 t_3\right)t_2 + \left(s_3 - S_1 t_1 - S_2 t_2\right)t_3 = n\,. \quad (8.77)$$

Since, according to Eq. (75), the vector of the distance between any pair of the corresponding points of the direct Bravais lattice (70),

[29] See, e.g., QM Sec. 3.4, where several particular Bravais lattices $\mathbf{R}$, and their reciprocals $\mathbf{Q}$, are considered.

[30] See, e.g., MA Eq. (7.6).

[31] For more reading on this important topic, I can recommend, for example, the classical monograph by B. Cullity, *Elements of X-Ray Diffraction*, 2nd ed., Addison-Wesley, 1978. (Note that its title uses the alternative name of the field, once again illustrating how blurry the boundary between the interference and diffraction is.)

$$\Delta\mathbf{R} = \Delta S_1 t_3 \mathbf{a}_1 + \Delta S_2 t_3 \mathbf{a}_2 - (\Delta S_1 t_1 + \Delta S_2 t_2)\mathbf{a}_3, \tag{8.78}$$

satisfies the condition $\Delta\mathbf{R}\cdot\mathbf{Q} = 2\pi\Delta n = 0$, this vector is normal to the (fixed) vector $\mathbf{Q}$. Hence, all the points corresponding to the 2D set (76) with arbitrary integers S_1 and S_2, are located on one geometric plane, called the *crystal* (or "lattice") *plane*. In a 3D system of $N >> 1$ scatterers (such as $N \sim 10^{20}$ atoms in a ~1-mm^3 solid crystal), with all linear dimensions comparable, such a plane contains $\sim N^{2/3} >> 1$ points. As a result, the constructive interference peaks are very sharp.

Now rewriting Eq. (75) as a relation for the vector $\mathbf{R}$'s component along the vector $\mathbf{Q}$,

$$R_Q = \frac{2\pi}{Q} n, \qquad \text{where } R_Q \equiv \mathbf{R}\cdot\mathbf{n}_Q \equiv \mathbf{R}\cdot\frac{\mathbf{Q}}{Q}, \quad \text{and } Q \equiv |\mathbf{Q}|, \tag{8.79}$$

we see that the parallel crystal planes corresponding to different numbers n (but the same $\mathbf{Q}$) are located in space periodically, with the smallest distance

$$d = \frac{2\pi}{Q}, \tag{8.80}$$

so the von Laue condition $\mathbf{q} = \mathbf{Q}$ may be rewritten as the following rule for the possible magnitudes of the scattering vector $\mathbf{q} \equiv \mathbf{k} - \mathbf{k}_0$:

$$q = \frac{2\pi n}{d}. \tag{8.81}$$

Figure 9a shows the diagram of the three wave vectors $\mathbf{k}$, $\mathbf{k}_0$, and $\mathbf{q}$, taking into account the elastic scattering condition $|\mathbf{k}| = |\mathbf{k}_0| = k \equiv 2\pi/\lambda$. From the diagram, we immediately get the famous *Bragg rule*[32] for the (equal) angles $\alpha \equiv \theta/2$ between the crystal plane and each of the vectors $\mathbf{k}$ and $\mathbf{k}_0$:

Bragg rule

$$k\sin\alpha = \frac{q}{2} = \frac{\pi n}{d}, \qquad \text{i.e. } 2d\sin\alpha = n\lambda. \tag{8.82}$$

The physical sense of this relation is very simple – see Fig. 9b drawn in the "direct" space of the radius-vectors $\mathbf{r}$, rather than in the reciprocal space of the wave vectors, as Fig. 9a. It shows that if the Bragg condition (82) is satisfied, the total difference $2d\sin\alpha$ of the optical paths of two waves, partly reflected from the adjacent crystal planes, is equal to an integer number of wavelengths, so these waves interfere constructively.

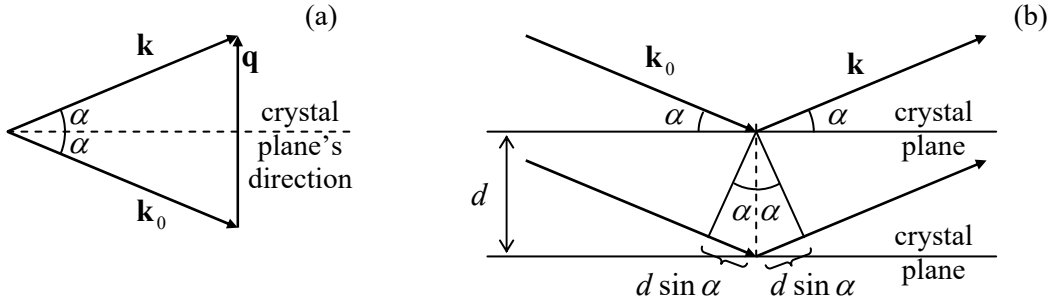


Fig. 8.9. Deriving the Bragg rule: (a) from the von Laue condition (in the reciprocal space), and (b) from a direct-space diagram. Note that the scattering angle θ equals 2α.

[32] Named after Sir William Bragg and his son, Sir William Lawrence Bragg, who were the first to demonstrate (in 1912) the X-ray diffraction by atoms in crystals. The Braggs' experiments have made the existence of atoms (before that, a somewhat hypothetical notion ignored by many physicists) indisputable.

Finally, note that the von Laue and Bragg rules, as well as the similar condition (60) for the 1D system of scatterers, are valid not only in the Born approximation but also follow from any adequate theory of scattering, because the phase sum (57) does not depend on the magnitude of the wave propagating from each elementary scatterer, provided that they are all equal.

8.5. The Huygens principle

As the reader could see, the Born approximation is very convenient for tracing the basic features of (and the difference between) the phenomena of interference and diffraction. Unfortunately, this approximation, based on the relative weakness of the scattered wave, cannot be used to describe more typical experimental implementations of these phenomena, for example, Young's two-slit experiment, or diffraction on a single slit or orifice – see, e.g. Fig. 10. Indeed, at such experiments, the orifice size a is typically much larger than the light's wavelength λ, and as a result, no clear decomposition of the fields to the "incident" and "scattered" waves is possible inside it.[33]

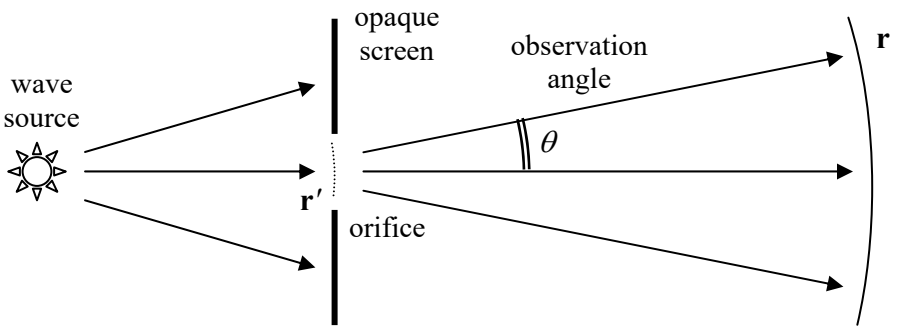


Fig. 8.10. Deriving the Huygens principle.

However, another approximation called the *Huygens* (or "Huygens-Fresnel") *principle*,[34] is very instrumental in the description of such situations. In this approach, the wave beyond the screen is represented as a linear superposition of spherical waves of the type (17), as if they were emitted by every point of the incident wave's front that has arrived at the orifice. This approximation is valid if the following strong conditions are satisfied:

$$\lambda << a << r\,, \tag{8.83}$$

where r is the distance of the observation point from the orifice. In addition, as we have seen in the last section, at small λ/a the diffraction phenomena are confined to angles $\theta \sim 1/ka \sim \lambda/a << 1$. For observation at such small angles, the mathematical expression of the Huygens principle, for the complex amplitude $f_\omega(\mathbf{r})$ of a monochromatic wave $f(\mathbf{r}, t) = \mathrm{Re}[f_\omega e^{-i\omega t}]$, is given by the following simple formula

$$f_\omega(\mathbf{r}) = C \int_{\text{orifice}} f_\omega(\mathbf{r}') \frac{e^{ikR}}{R} d^2 r'\,. \tag{8.84}$$

[33] Another complaint against the Born approximation is that it does not satisfy the so-called *optical* (or "forward scattering") *theorem* relating σ to scattering with $\mathbf{k} = \mathbf{k}_0$. This relation is especially important for the quantum-mechanical description of particle scattering, and in this series, will be discussed in its QM part (Sec. 3.3).

[34] Named after Christian Huygens (1629-1695) who had conjectured the wave nature of light (which remained controversial for more than a century, until T. Young's experiments), and Augustin-Jean Fresnel (1788-1827) who developed a quantitative theory of diffraction, and in particular gave a mathematical formulation of the Huygens principle. (Note that Eq. (91), sufficient for the purposes of this course, is not its most general form.)

Here f is any transverse component of any of the wave's fields (either **E** or **H**),[35] R is the distance between point **r**' at the orifice and the observation point **r** (i.e. the magnitude of vector **R** ≡ **r** – **r**'), and C is a complex constant.

Before describing the proof of Eq. (84), let me carry out its sanity check – which also will give us the constant C. Let us see what the Huygens principle gives for the case when the field under the integral is a plane wave with the complex amplitude $f_\omega(z)$, propagating along axis z, with an unlimited x-y front, (i.e. when there is no opaque screen at all), so in Eq. (84) we should take the whole [x, y] plane, say with $z' = 0$, as the integration area– see Fig. 11.

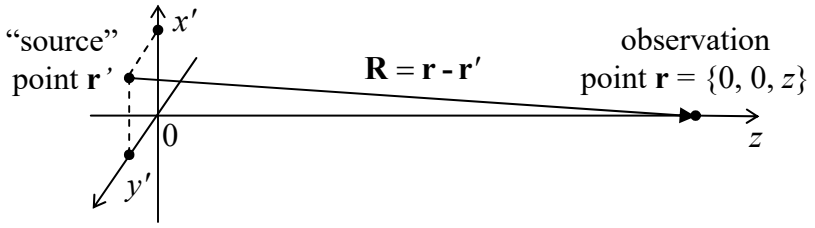


Fig. 8.11. Applying the Huygens principle to a plane incident wave.

Then, for the observation point with coordinates $x = 0$, $y = 0$, and $z > 0$, Eq. (84) yields

$$f_\omega(z) = Cf_\omega(0)\int dx'\int dy'\frac{\exp\left\{ik\left(x'^2+y'^2+z^2\right)^{1/2}\right\}}{\left(x'^2+y'^2+z^2\right)^{1/2}}. \tag{8.85}$$

Before specifying the integration limits, let us consider the range $|x'|, |y'| << z$. In this range, the square root participating in Eq. (85) twice, may be approximated as

$$\left(x'^2+y'^2+z^2\right)^{1/2} \equiv z\left(1+\frac{x'^2+y'^2}{z^2}\right)^{1/2} \approx z\left(1+\frac{x'^2+y'^2}{2z^2}\right) \equiv z+\frac{x'^2+y'^2}{2z}. \tag{8.86}$$

At $z >> \lambda$, the denominator of Eq. (85) is a much slower function of x' and y' than the exponent in the numerator, and in the former case, it is sufficient (as we will check *a posteriori*) to keep just the main, first term of expansion (86). With that, Eq. (85) becomes

$$f_\omega(z) = Cf_\omega(0)\frac{e^{ikz}}{z}\int dx'\int dy'\exp\frac{ik\left(x'^2+y'^2\right)}{2z} = Cf_\omega(0)\frac{e^{ikz}}{z}I_xI_y, \tag{8.87}$$

where I_x and I_y are two similar integrals; for example,

$$I_x = \int \exp\frac{ikx'^2}{2z}dx' = \left(\frac{2z}{k}\right)^{1/2}\int \exp\left\{i\xi^2\right\}d\xi \equiv \left(\frac{2z}{k}\right)^{1/2}\left[\int \cos\left(\xi^2\right)d\xi + i\int \sin\left(\xi^2\right)d\xi\right], \tag{8.88}$$

where $\xi \equiv (k/2z)^{1/2}x'$. These are the so-called *Fresnel integrals*. I will discuss them in more detail in the next section (see, in particular, Fig. 13), and for now, only one property[36] of these integrals is important

[35] The fact that the Huygens principle is valid for any field component should not too surprising. In the limit $a >> \lambda$, the real boundary conditions at the orifice edges are not important; it is only important for the screen that limits the orifice, to be opaque. Because of this, the Huygens principle (84) is a part of the so-called *scalar theory of diffraction*. (I will not have time to discuss the *vector theory* of these effects, which is more accurate at smaller a – see, e.g., Chapter 11 of the monograph by M. Born and E. Wolf, cited at the end of Sec. 7.1.)

for us: if taken in symmetric limits $[-\xi_0, +\xi_0]$, both of them rapidly converge to the same value, $(\pi/2)^{1/2}$, as soon as ξ_0 becomes much larger than 1. This means that even if we do not impose any exact limits on the integration area in Eq. (85), this integral converges to the value

$$f_\omega(z) = Cf_\omega(0)\frac{e^{ikz}}{z}\left\{\left(\frac{2z}{k}\right)^{1/2}\left[\left(\frac{\pi}{2}\right)^{1/2}+i\left(\frac{\pi}{2}\right)^{1/2}\right]\right\}^2 \equiv \left(C\frac{2\pi i}{k}\right)f_\omega(0)e^{ikz}, \tag{8.89}$$

due to contributions from the central area with a linear size corresponding to $\Delta\xi \sim 1$, i.e. to

$$\Delta x \sim \Delta y \sim \left(\frac{z}{k}\right)^{1/2} \sim (\lambda z)^{1/2}, \qquad . \tag{8.90}$$

so the net contribution from the front points $\mathbf{r}'$ well beyond the range (90) is negligible.[37] (Within our assumptions (83), which in particular require λ to be much less than z, the *diffraction angle* $\Delta x/z \sim \Delta y/z \sim (\lambda/z)^{1/2}$, corresponding to the important area of the front, is small.) According to Eq. (89), to sustain the unperturbed plane wave propagation, $f_\omega(z) = f_\omega(0)e^{ikz}$, the constant C has to be taken equal to $k/2\pi i$. Thus, the Huygens principle's prediction (84), in its final form, reads

$$f_\omega(\mathbf{r}) = \frac{k}{2\pi i}\int_{\text{orifice}} f_\omega(\mathbf{r}')\frac{e^{ikR}}{R}d^2r', \tag{8.91}$$

Huygens principle

and describes, in particular, the straight propagation of the plane wave (in a uniform medium).

Let me pause to emphasize how nontrivial this result is. It would be a natural corollary of Eqs. (25) (and the linear superposition principle) if all points of the orifice were filled with point scatterers that re-emit all the incident waves as spherical waves. However, as it follows from the above example, the Huygens principle also works if there is nothing in the orifice but the free space!

This is why let us discuss a proof of this principle,[38] based on Green's theorem (2.207). Let us apply it to the function $f = f_\omega$ where f_ω is the complex amplitude of a scalar component of one of the wave's fields, which satisfies the Helmholtz equation (7.204),

$$\left(\nabla^2 + k^2\right)f_\omega(\mathbf{r}) = 0, \tag{8.92}$$

and the function $g = G_\omega$ is the temporal Fourier image of the corresponding Green's function. The latter function may be defined, as usual, as the solution of the same equation with the added delta-functional right-hand side with an arbitrary coefficient, for example,

$$\left(\nabla^2 + k^2\right)G_\omega(\mathbf{r},\mathbf{r}') = -4\pi\delta(\mathbf{r}-\mathbf{r}'). \tag{8.93}$$

Using Eqs. (92) and (93) to express the Laplace operators of the functions f_ω and G_ω, we may rewrite Eq. (2.207) as

---

[36] See, e.g., MA Eq. (6.10).

[37] This result very is natural, because the function $\exp\{ikR\}$ oscillates fast with the change of $\mathbf{r}'$, so the contributions from various front points are averaged out. Indeed, the only reason why the central part of the plane $[x', y']$ gives a non-zero contribution (89) to $f_\omega(z)$ is that the phase exponents stop oscillating as $(x'^2 + y'^2)$ is reduced below $\sim z/k$ – see Eq. (86).

[38] This proof was given in 1882 by the same G. Kirchhoff whose circuit rules were discussed in Sec. 4.1 and 6.6.

$$\int_V \left\{ f_\omega \left[-k^2 G_\omega(\mathbf{r},\mathbf{r}') - 4\pi\delta(\mathbf{r}-\mathbf{r}') \right] - G_\omega(\mathbf{r},\mathbf{r}') \left[-k^2 f_\omega \right] \right\} d^3 r = \oint_S \left[f_\omega \frac{\partial G_\omega(\mathbf{r},\mathbf{r}')}{\partial n} - G_\omega(\mathbf{r},\mathbf{r}') \frac{\partial f_\omega}{\partial n} \right] d^2 r , \quad (8.94)$$

where **n** is the outward normal to the surface *S* limiting the integration volume *V*. Two terms on the left-hand side of this relation cancel, so after swapping the arguments **r** and **r**', we get

$$-4\pi f_\omega(\mathbf{r}) = \oint_S \left[f_\omega(\mathbf{r}') \frac{\partial G_\omega(\mathbf{r}',\mathbf{r})}{\partial n'} - g_\omega(\mathbf{r}',\mathbf{r}) \frac{\partial G_\omega(\mathbf{r}')}{\partial n'} \right] d^2 r' . \quad (8.95)$$

This relation is only correct if the selected volume *V* includes the point **r** (otherwise we would not get its left-hand side from the integration of the delta function), but does not include the genuine source of the wave – otherwise, Eq. (92) would have a non-zero right-hand side. Now let **r** be the field observation point, *V* be all the source-free half-space (for example, the space right of the screen in Fig. 10), so *S* is the surface of the screen, including the orifice. Then the right-hand side of Eq. (95) describes the field (at the observation point **r**) induced by the wave passing through the orifice points **r**'. Since no waves are emitted by the opaque parts of the screen, we can limit the integration by the orifice area.[39] Assuming also that the opaque parts of the screen do not re-emit the waves "radiated" by the orifice, we can take the solution of Eq. (93) to be the retarded potential for the free space:[40]

$$G_\omega(\mathbf{r},\mathbf{r}') = \frac{e^{ikR}}{R} . \quad (8.96)$$

Plugging this expression into Eq. (82), we get

Kirchhoff integral

$$-4\pi f_\omega(\mathbf{r}) = \oint_{\text{orifice}} \left[f_\omega(\mathbf{r}') \frac{\partial}{\partial n'} \left(\frac{e^{ikR}}{R} \right) - \left(\frac{e^{ikR}}{R} \right) \frac{\partial f_\omega(\mathbf{r}')}{\partial n'} \right] d^2 r' . \quad (8.97)$$

This is the so-called *Kirchhoff* (or "Fresnel-Kirchhoff") *integral*. (Again, with the integration extended over *all* boundaries of the volume *V*, this would be an exact mathematical result.) Now, let us make two additional approximations. The first of them stems from Eq. (83): at $ka >> 1$, the wave's spatial dependence in the orifice area may be represented as

$$f_\omega(\mathbf{r}') = (\text{a slow function of } \mathbf{r}') \times \exp\{i\mathbf{k}_0 \cdot \mathbf{r}'\} , \quad (8.98)$$

where "slow" means a function that changes on the scale of *a* rather than λ. If, also, $kR >> 1$, then the differentiation in Eq. (97) may be, in both instances, limited to the rapidly changing exponents, giving

$$-4\pi f_\omega(\mathbf{r}) = \oint_{\text{orifice}} i(\mathbf{k}+\mathbf{k}_0)\cdot\mathbf{n}' \frac{e^{ikR}}{R} f(\mathbf{r}') d^2 r' . \quad (8.99)$$

Second, if all observation angles are small, we may take $\mathbf{k}\cdot\mathbf{n}' \approx \mathbf{k}_0\cdot\mathbf{n}' \approx -k$. With that, Eq. (99) is reduced to the Huygens principle in its form (91).

---

[39] Actually, this is a nontrivial point of the proof. Indeed, it may be shown that the exact solution of Eq. (94) identically is equal to zero if $f(\mathbf{r}')$ and $\partial f(\mathbf{r}')/\partial n'$ vanish together at any *part* of the boundary, of a non-zero area. A more careful analysis of this issue (it is the task of the formal vector theory of diffraction, which I will not have time to pursue) confirms the validity of the described intuition-based approach at $a >> \lambda$.

[40] It follows, e.g., from Eq. (16) with a monochromatic source $q(t) = q_\omega \exp\{-i\omega t\}$, with the amplitude $q_\omega = 4\pi\varepsilon$ that fits the right-hand side of Eq. (93).

It is clear that the principle immediately gives a very simple description of the interference of waves passing through two small holes in the screen. Indeed, if the holes' sizes are negligible in comparison with the distance a between them (though still are much larger than the wavelength!), Eq. (91) yield

$$f_\omega(\mathbf{r}) = c_1 e^{ikR_1} + c_2 e^{ikR_2}, \quad \text{with } c_{1,2} \equiv k f_{1,2} A_{1,2} / 2\pi i R_{1,2}, \tag{8.100}$$

where $R_{1,2}$ are the distances between the holes and the observation point, and $A_{1,2}$ are the hole areas. For the wave intensity, Eq. (100) gives

$$\overline{S} \propto f_\omega f_\omega^* = |c_1|^2 + |c_2|^2 + 2|c_1||c_2|\cos\left[k(R_1 - R_2) + \varphi\right], \quad \text{where } \varphi \equiv \arg c_1 - \arg c_2. \tag{8.101}$$

The first two terms in the last expression clearly represent the intensities of the partial waves passed through each hole, while the last one is the result of their interference. The interference pattern's *contrast ratio*

$$\frac{\overline{S}_{\max}}{\overline{S}_{\min}} = \left(\frac{|c_1| + |c_2|}{|c_1| - |c_2|}\right)^2, \tag{8.102}$$

is the largest (infinite) when both waves have equal amplitudes.

The analysis of the interference pattern is simple if the line connecting the holes is perpendicular to wave vector $\mathbf{k} \approx \mathbf{k}_0$ – see Fig. 6a. Selecting the coordinate axes as shown in that figure, and using for the distances $R_{1,2}$ the same expansion as in Eq. (86), for the interference term in Eq. (101) we get

$$\cos\left[k(R_1 - R_2) + \varphi\right] \approx \cos\left(\frac{kxa}{z} + \varphi\right). \tag{8.103}$$

This means that the term does not depend on y, i.e. the interference pattern in the plane of constant z is a set of straight, parallel strips, perpendicular to the vector **a**, with the period given by Eq. (60), i.e. by the Bragg law.[41] This result is strictly valid only at $y^2 << z^2$; it is straightforward to use the next term in the Taylor expansion (73) to show that farther on from the interference plane $y = 0$, the strips start to diverge.

8.6. Fresnel and Fraunhofer diffraction patterns

Now let us use the Huygens principle to analyze a slightly more complex problem: plane wave's diffraction on a long, straight slit of a constant width a (Fig. 12). According to Eq. (83), to use the Huygens principle for the problem's analysis we need to have $\lambda << a << z$. Moreover, the simple version (91) of the principle is only valid for small observation angles, $|x| << z$. Note, however, that the relation between two dimensionless parameters of the problem, z/a and a/λ, which are both much less than 1, is so far arbitrary; as we will see in a minute, this relation determines the type of the observed diffraction pattern.

[41] The phase shift φ vanishes at the normal incidence of a plane wave on the holes. Note, however, that the spatial shift of the interference pattern following from Eq. (103), $\Delta x = -(z/ka)\varphi$, is extremely convenient for the experimental measurement of the phase shift between two waves, especially if it is induced by some factor (such as insertion of a transparent object into one of the interferometer's arms) that may be turned on/off at will.

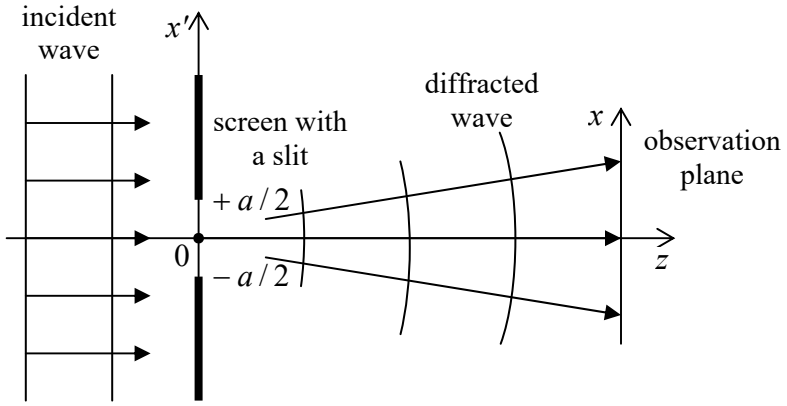


Fig. 8.12. Diffraction on a slit.

Let us apply Eq. (91) to our current problem (Fig. 12), for the sake of simplicity assuming the normal wave incidence, and taking $z' = 0$ at the screen plane:

$$f_\omega(x,z) = f_0 \frac{k}{2\pi i}\int_{-a}^{+a} dx' \int_{-\infty}^{+\infty} dy' \frac{\exp\left\{ik\left[(x-x')^2 + y'^2 + z^2\right]^{1/2}\right\}}{\left[(x-x')^2 + y'^2 + z^2\right]^{1/2}}, \tag{8.104}$$

where $f_0 \equiv f_\omega(x', 0) = \text{const}$ is the incident wave's amplitude. This is the same integral as in Eq. (85), except for the finite limits for the integration variable x', and may be simplified similarly, using the small-angle condition $(x-x')^2 + y'^2 << z^2$:

$$f_\omega(x,z) \approx f_0 \frac{k}{2\pi i}\frac{e^{ikz}}{z}\int_{-a/2}^{+a/2} dx' \int_{-\infty}^{+\infty} dy' \exp\frac{ik\left[(x-x')^2 + y'^2\right]}{2z} \equiv f_0 \frac{k}{2\pi i}\frac{e^{ikz}}{z} I_x I_y . \tag{8.105}$$

The integral over y' is the same as in the last section:

$$I_y \equiv \int_{-\infty}^{+\infty} \exp\frac{iky'^2}{2z} dy' = \left(\frac{2\pi i z}{k}\right)^{1/2}, \tag{8.106}$$

but the integral over x' is more general, because of its finite limits:

$$I_x \equiv \int_{-a/2}^{+a/2} \exp\frac{ik(x-x')^2}{2z} dx' . \tag{8.107}$$

It may be simplified in the following two (opposite) limits.

(i) *Fraunhofer diffraction* takes place when $z/a >> a/\lambda$ – the relation which may be rewritten either as $a << (z\lambda)^{1/2}$, or as $ka^2 << z$. In this limit, the ratio kx'^2/z is negligibly small for all values of x' under the integral, and we can approximate it as

$$\begin{aligned} I_x &= \int_{-a/2}^{+a/2} \exp\frac{ik\left(x^2 - 2xx' + x'^2\right)}{2z} dx' \approx \int_{-a/2}^{+a/2} \exp\frac{ik\left(x^2 - 2xx'\right)}{2z} dx' \\ &\equiv \exp\frac{ikx^2}{2z}\int_{-a/2}^{+a/2} \exp\left\{-\frac{ikxx'}{z}\right\} dx' = \frac{2z}{kx}\exp\left\{\frac{ikx^2}{2z}\right\}\sin\frac{kxa}{2z}, \end{aligned} \tag{8.108}$$

so Eq. (105) yields

$$f_\omega(x,z) \approx f_0 \frac{k}{2\pi i}\frac{e^{ikz}}{z}\frac{2z}{kx}\left(\frac{2\pi i z}{k}\right)^{1/2}\exp\left\{\frac{ikx^2}{2z}\right\}\sin\frac{kxa}{2z}, \tag{8.109}$$

and hence the relative wave intensity is

$$\frac{\bar{S}(x,z)}{S_0} = \left|\frac{f_\omega(x,z)}{f_0}\right|^2 = \frac{8z}{\pi k x^2}\sin^2\frac{kxa}{2z} \equiv \frac{2}{\pi}\frac{ka^2}{z}\operatorname{sinc}^2\left(\frac{ka\theta}{2}\right), \tag{8.110}$$

Fraunhofer diffraction pattern

where S_0 is the intensity of the incident wave, and $\theta \equiv x/z << 1$ is the observation angle. Comparing this expression with Eq. (69), we see that this diffraction pattern is exactly the same as that for a similar (uniform, 1D) object in the Born approximation – see the red line in Fig. 8. Note again that the angular width $\delta\theta$ of the Fraunhofer pattern is of the order of $1/ka$, so its linear width $\delta x = z\delta\theta$ is of the order of $z/ka \sim z\lambda/a$.[42] Hence the condition of the Fraunhofer approximation's validity may be also represented as $a << \delta x$.

(ii) *Fresnel diffraction*. In the opposite limit of a relatively wide slit, with $a >> \delta x = z\delta\theta \sim z/ka \sim z\lambda/a$, i.e. $ka^2 >> z$, the diffraction patterns at two edges of the slit are well separated. Hence, near each edge (for example, near $x' = -a/2$) we may simplify Eq. (107) as

$$I_x(x) \approx \int_{-a/2}^{+\infty} \exp\frac{ik(x-x')^2}{2z}dx' \equiv \left(\frac{2z}{k}\right)^{1/2}\int_{(k/2z)^{1/2}(x+a/2)}^{+\infty}\exp\{i\zeta^2\}d\zeta, \tag{8.111}$$

and express it via the special functions called the Fresnel integrals:[43]

$$\mathscr{C}(\xi) \equiv \left(\frac{2}{\pi}\right)^{1/2}\int_0^\xi \cos(\zeta^2)d\zeta, \qquad \mathscr{S}(\xi) \equiv \left(\frac{2}{\pi}\right)^{1/2}\int_0^\xi \sin(\zeta^2)d\zeta, \tag{8.112}$$

Fresnel integrals

whose plots are shown in Fig. 13a. As was mentioned above, at large values of their argument (ξ), both functions tend to ½.

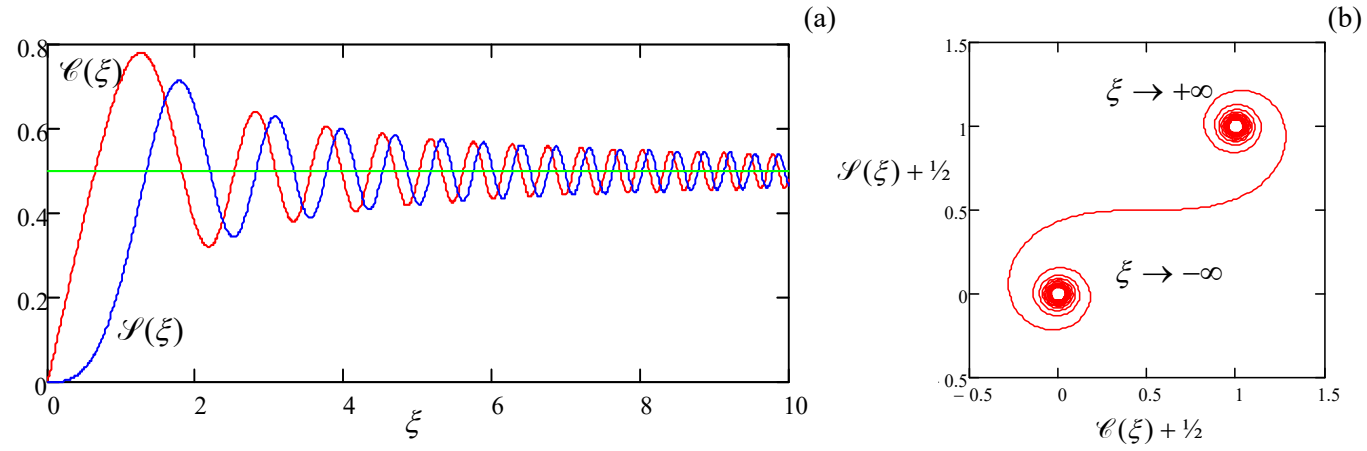


Fig. 8. 13. (a) The Fresnel integrals and (b) their parametric representation.

[42] Note also that since in this limit $ka^2 << z$, Eq. (97) shows that even the maximum value $S(0, z)$ of the diffracted wave's intensity is much lower than that (S_0) of the incident wave. This is natural because the incident power $S_0 a$ per unit length of the slit is now distributed over a much larger width $\delta x >> a$, so $S(0, z) \sim S_0 (a/\delta x) << S_0$.

[43] Slightly different definitions of these functions, affecting the constant factors, may also be met in literature.

Plugging this expression into Eqs. (105) and (111), for the diffracted wave intensity, in the Fresnel limit (i.e. at $|x + a/2| << a$), we get

Fresnel diffraction pattern

$$\frac{\overline{S}(x,z)}{S_0} = \frac{1}{2}\left\{\left[\mathscr{C}\left(\left(\frac{k}{2z}\right)^{1/2}\left(x+\frac{a}{2}\right)\right)+\frac{1}{2}\right]^2 + \left[\mathscr{S}\left(\left(\frac{k}{2z}\right)^{1/2}\left(x+\frac{a}{2}\right)\right)+\frac{1}{2}\right]^2\right\}. \tag{8.113}$$

A plot of this function (Fig. 14) shows that the diffraction pattern is very peculiar: while in the "dark" region $x < -a/2$ the wave intensity fades monotonically, the transition to the "bright" region within the gap ($x > -a/2$) is accompanied by intensity oscillations, just as at the Fraunhofer diffraction – cf. Fig. 8.

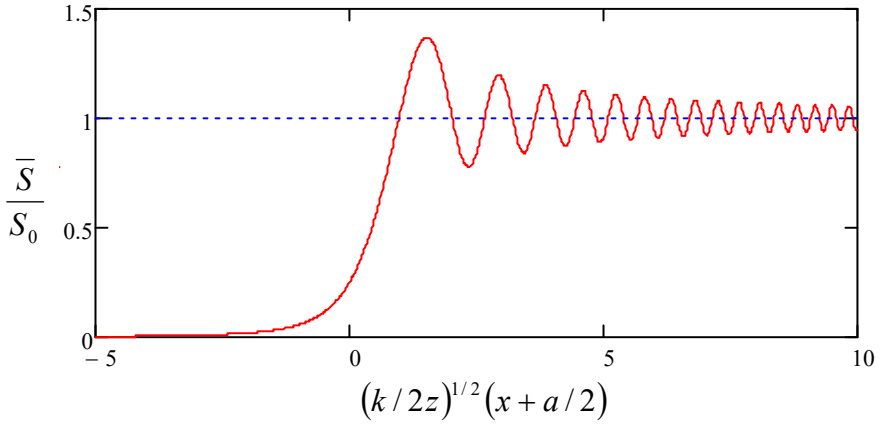


Fig. 8.14. The Fresnel diffraction pattern.

This behavior, which is described by the following asymptotes,

$$\frac{\overline{S}}{S_0} \to \begin{cases} 1 + \dfrac{1}{\sqrt{\pi}}\dfrac{\sin(\xi^2 - \pi/4)}{\xi}, & \text{for } \xi \equiv \left(\dfrac{k}{2z}\right)^{1/2}\left(x+\dfrac{a}{2}\right) \to +\infty, \\ \dfrac{1}{4\pi\xi^2}, & \text{for } \xi \to -\infty, \end{cases} \tag{8.114}$$

is essentially an artifact of "observing" just the wave intensity (i.e. its real amplitude) rather than its phase as well. Indeed, as may be seen even more clearly from the parametric representation of the Fresnel integrals, shown in Fig. 13b, these functions oscillate similarly at large positive and negative values of their argument. (This famous pattern is called either the *Euler spiral* or the *Cornu spiral*.) Physically, this means that the wave diffraction at the slit edge leads to similar oscillations of its phase at $x < -a/2$ and $x > -a/2$; however, in the latter region (i.e. inside the slit) the diffracted wave overlaps the incident wave passing through the slit directly, and their interference reveals the phase oscillations, making them visible in the observed intensity as well.

Note that according to Eq. (113), the linear scale δx of the Fresnel diffraction pattern is of the order of $(2z/k)^{1/2}$, i.e. complies with the estimate given by Eq. (90). If the slit is gradually narrowed so that its width a becomes comparable to δx,[44] the Fresnel diffraction patterns from both edges start to "collide" (interfere). The resulting wave, fully described by Eq. (107), is just a sum of two contributions of the type (111) from both edges of the slit. The resulting interference pattern is somewhat complicated, and only when a becomes substantially less than δx, it is reduced to the simple Fraunhofer pattern (110).

[44] Note that this condition may be also rewritten as $a \sim \delta x$, i.e. $z/a \sim a/\lambda$.

Of course, this crossover from the Fresnel to Fraunhofer diffraction may be also observed, at fixed wavelength λ and slit width a, by increasing z, i.e. by measuring the diffraction pattern farther and farther away from the slit.

Note also that the Fraunhofer limit is always valid if the diffraction is measured as a function of the diffraction angle θ alone. This may be done, for example, by collecting the diffracted wave with a "positive" (converging) lens and observing the diffraction pattern in its focal plane.

8.7. Geometrical optics placeholder

I would not like the reader to miss, behind all these details, the main feature of the Fresnel diffraction, which has an overwhelming practical significance. Namely, besides narrow diffraction "cones" (actually, parabolic-shaped regions) with the lateral scale $\delta x \sim (\lambda z)^{1/2}$, the wave far behind a slit of width $a >> \lambda, \delta x$, reproduces the field just behind the slit, i.e. reproduces the unperturbed incident wave inside it, and has a negligible intensity in the shade regions outside it. An evident generalization of this fact is that when a plane wave (in particular an electromagnetic wave) passes any opaque object of a large size $a >> \lambda$, it propagates around it, by distances z up to $\sim a^2/\lambda$, along straight lines, with virtually negligible diffraction effects. This fact gives the strict foundation for the notion of the *wave ray* (or *beam*), as the line perpendicular to the local front of a quasi-plane wave. In a uniform media such a ray follows a straight line,[45] but it refracts in accordance with the Snell law at the interface of two media with different values of the wave speed v, i.e. different values of the refraction index. The concept of rays enables the whole vast field of *geometric optics*, devoted mostly to ray tracing in various (sometimes very sophisticated) optical systems.

This is why, at this point, an E&M course that followed the scientific logic more faithfully than this one, would give an extended discussion of the geometric and quasi-geometric optics, including (as a minimum[46]) such vital topics as

- the so-called *lensmaker's equation* expressing the focus length f of a lens via the curvature radii of its spherical surfaces and the refraction index of the lens material,
- the *thin lens formula* relating the image distance from the lens via f and the source distance,
- the concepts of basic optical instruments such as *glasses*, *telescopes*, and *microscopes*,
- the concepts of the spherical, angular, and chromatic *aberrations* (optical image distortions).

However, since I have made a (possibly, wrong) decision to follow the common tradition in selecting the main topics for this course, I do not have time/space left for such discussion. Still, I am using this "placeholder" pseudo-section to relay my deep conviction that any educated physicist has to know the geometric optics basics. If the reader has not been exposed to this subject during their undergraduate studies, I highly recommend at least browsing one of the available textbooks.[47]

---

[45] In application to optical waves, this notion may be traced back to at least the work by Hero (a.k.a. Heron) of Alexandria (circa 170 AD). Curiously, he correctly described light reflection from one or several plane mirrors, starting from the completely wrong idea of light propagation *from* the eye of the observer *to* the observed object.

[46] Admittedly, even this list leaves aside several spectacular effects, including such a beauty as *conical refraction* in biaxial crystals – see, e.g., Chapter 15 of the textbook by M. Born and E. Wolf, cited in the end of Sec. 7.1.

[47] My top recommendation for that purpose would be Chapters 3-6 and Sec. 8.6 in Born and Wolf. A simpler alternative is Chapter 10 in G. Fowles, *Introduction to Modern Optics*, 2nd ed., Dover, 1989. Note also that the venerable field of optical microscopy is currently revitalized by holographic/tomographic methods, using the

8.8. Fraunhofer diffraction from more complex scatterers

So far, our quantitative analysis of diffraction has been limited to a very simple geometry – a single slit in an otherwise opaque screen (Fig. 12). However, in the most important Fraunhofer limit, $z >> ka^2$, it is easy to get a very simple expression for the plane wave diffraction/interference by a plane orifice (with a linear size scale a) of arbitrary shape. Indeed, the evident 2D generalization of the approximation (106)-(107) is

$$\begin{aligned} I_x I_y &= \int_{\text{orifice}} \exp\frac{ik\left[(x-x')^2+(y-y')^2\right]}{2z}dx'dy' \\ &\approx \exp\left\{\frac{ik(x^2+y^2)}{2z}\right\}\int_{\text{orifice}} \exp\left\{-i\frac{kxx'}{z}-i\frac{kyy'}{z}\right\}dx'dy', \end{aligned} \tag{8.115}$$

so besides the inconsequential total phase factor, Eq. (105) is reduced to

General Fraunhofer diffraction pattern

$$f(\boldsymbol{\rho}) \propto f_0 \int_{\text{orifice}} \exp\{-i\boldsymbol{\kappa}\cdot\boldsymbol{\rho}'\}d^2\rho' \equiv f_0 \int_{\text{all screen}} T(\boldsymbol{\rho}')\exp\{-i\boldsymbol{\kappa}\cdot\boldsymbol{\rho}'\}d^2\rho' . \tag{8.116}$$

Here the 2D vector $\boldsymbol{\kappa}$ (not to be confused with wave vector $\mathbf{k}$, which is virtually perpendicular to $\boldsymbol{\kappa}$!) is defined as

$$\boldsymbol{\kappa} \equiv k\frac{\boldsymbol{\rho}}{z} \approx \mathbf{q} \equiv \mathbf{k} - \mathbf{k}_0 , \tag{8.117}$$

and $\boldsymbol{\rho} = \{x, y\}$ and $\boldsymbol{\rho}' = \{x', y'\}$ are 2D radius vectors in the, respectively, observation and orifice planes – both nearly normal to the vectors $\mathbf{k}$ and $\mathbf{k}_0$.[48] In the last form of Eq. (116), the function $T(\boldsymbol{\rho}')$ describes the screen's transparency at point $\boldsymbol{\rho}'$, and the integral is over the whole screen plane $z' = 0$. (Though the two forms of Eq. (116) are strictly equivalent only if $T(\boldsymbol{\rho}')$ is equal to either 1 or 0, its last form may be readily obtained from Eq. (91) with $f(\mathbf{r}') = T(\boldsymbol{\rho}')f_0$ for any transparency profile, provided that $T(\boldsymbol{\rho}')$ is any function that changes substantially only at distances much larger than $\lambda \equiv 2\pi/k$.)

From the mathematical point of view, the last form of Eq. (116) is just the 2D spatial Fourier transform of the function $T(\boldsymbol{\rho}')$, with the variable $\boldsymbol{\kappa}$ defined by the observation point's position: $\boldsymbol{\rho} \equiv (z/k)\,\boldsymbol{\kappa} \equiv (z\lambda/2\pi)\,\boldsymbol{\kappa}$. This interpretation is useful because of the experience we all have with the Fourier transform, if only in the context of its time/frequency applications. For example, if the orifice is a single small hole, $T(\boldsymbol{\rho}')$ may be approximated by a delta function, so Eq. (116) yields $|f(\boldsymbol{\rho})| \approx$ const. This result corresponds (at least for the small diffraction angles $\theta \equiv \rho/z$, for which the Huygens approximation is valid) to a spherical wave spreading from the point-like orifice. Next, for two small holes, Eq. (116) immediately gives the interference pattern (103). Let me now use Eq. (116) to analyze other simplest (and most important) 1D transparency profiles, leaving a few 2D cases for the reader's exercise.

(i) A single slit of width a (Fig. 12) may be described by transparency

---

scattered wave's phase information. These methods are especially productive in biology and medicine – see, e.g., M. Brezinski, *Optical Coherence Tomography*, Academic Press, 2006, and G. Popescu, *Quantitative Phase Imaging of Cells and Tissues*, McGraw-Hill (2011).

[48] Note that for a thin uniform plate of the same shape as the orifice we are discussing now, the Born phase integral (63) with $q << k$ gives a result functionally similar to Eq. (116).

$$T(\boldsymbol{\rho}') = \begin{cases} 1, & \text{for } |x'| < a/2, \\ 0, & \text{otherwise.} \end{cases} \tag{8.118}$$

Its substitution into Eq. (116) yields

$$f(\boldsymbol{\rho}) \propto f_0 \int_{-a/2}^{+a/2} \exp\{-i\kappa_x x'\}\, dx' = f_0 \frac{\exp\{-i\kappa_x a/2\} - \exp\{i\kappa_x a/2\}}{-i\kappa_x} \propto \operatorname{sinc}\left(\frac{\kappa_x a}{2}\right) = \operatorname{sinc}\left(\frac{kxa}{2z}\right), \tag{8.119}$$

naturally returning us to Eqs. (64) and (110), and hence to the red lines in Fig. 8 for the wave intensity. (Please note again that Eq. (116) describes only the Fraunhofer, but not the Fresnel diffraction!)

(ii) Two infinitely narrow, similar, parallel slits with a larger distance a between them (i.e. the simplest model of Young's two-slit experiment) may be described by taking

$$T(\boldsymbol{\rho}') \propto \delta\left(x' - \frac{a}{2}\right) + \delta\left(x' + \frac{a}{2}\right), \tag{8.120}$$

so Eq. (116) yields the generic 1D interference pattern,

$$f(\boldsymbol{\rho}) \propto f_0 \left[\exp\left\{-\frac{i\kappa_x a}{2}\right\} + \exp\left\{\frac{i\kappa_x a}{2}\right\}\right] \propto \cos\frac{\kappa_x a}{2} = \cos\frac{kxa}{2z}, \tag{8.121}$$

whose intensity is shown with the blue line in Fig. 8.

(iii) In a more realistic model of Young's experiment, each slit has a width (say, w) that is much larger than the light wavelength λ, but still much smaller than the slit spacing a. This situation may be described by the following transparency function

$$T(\boldsymbol{\rho}') = \sum_{\pm} \begin{cases} 1, & \text{for } |x' \pm a/2| < w/2, \\ 0, & \text{otherwise,} \end{cases} \tag{8.122}$$

for which Eq. (116) yields a natural combination of the results (119) (with a replaced with w) and (121):

$$f(\mathbf{r}) \propto \operatorname{sinc}\left(\frac{kxw}{2z}\right)\cos\left(\frac{kxa}{2z}\right). \tag{8.123}$$

This is the usual interference pattern, but modulated with a Fraunhofer-diffraction envelope – shown in Fig. 15 with the dashed blue line. Since the function $\operatorname{sinc}^2\xi$ decreases very fast beyond its first zeros at $\xi = \pm\pi$, the practical number of observable interference fringes is close to $2a/w$.

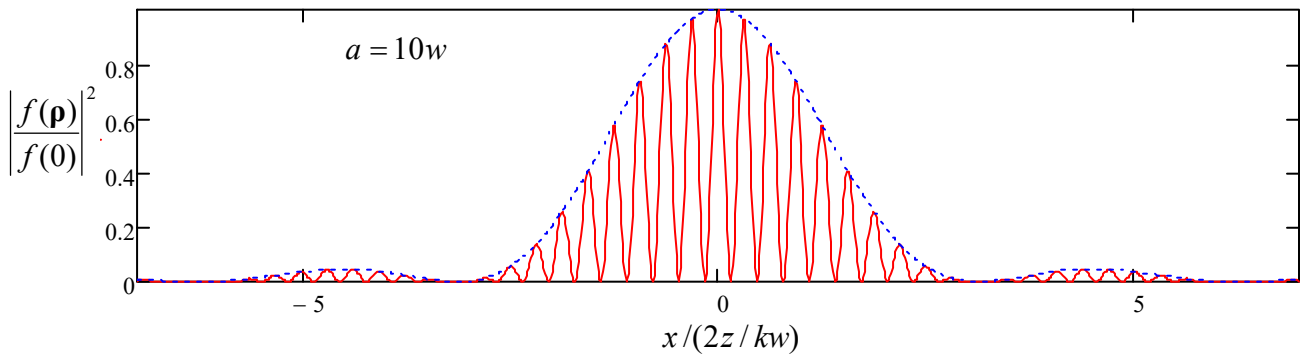


Fig. 8.15. Young's double-slit interference pattern for a finite-width slit.

(iv) A structure very useful for experimental and engineering practice is a set of many parallel, similar slits, called the *diffraction grating*.[49] If the slit's width is much smaller than period d of the grating, its transparency function may be approximated as

$$T(\boldsymbol{\rho}') \propto \sum_{n=-\infty}^{+\infty} \delta(x' - nd), \tag{8.124}$$

and Eq. (116) yields

$$f(\boldsymbol{\rho}) \propto \sum_{n=-\infty}^{n=+\infty} \exp\{-in\kappa_x d\} = \sum_{n=-\infty}^{n=+\infty} \exp\left\{-i\frac{nkxd}{z}\right\}. \tag{8.125}$$

This sum vanishes for all values of $\kappa_x d$ that are not multiples of 2π, so the result describes sharp intensity peaks at the following diffraction angles:

$$\theta_m \equiv \left(\frac{x}{z}\right)_m = \left(\frac{\kappa_x}{k}\right)_m = \frac{2\pi}{kd}m = \frac{\lambda}{d}m. \tag{8.126}$$

Taking into account that this result is only valid for small angles $|\theta_m| \ll 1$, it may be interpreted exactly as Eq. (59) – see Fig. 6a. However, in contrast with the interference (121) from two slits, the destructive interference from many slits kills the net wave as soon as the angle is even slightly different from each value (60). This is very convenient for spectroscopic purposes because the diffraction lines produced by multi-frequency waves do not overlap even if the frequencies of their adjacent components are very close.

Two unavoidable features of practical diffraction gratings make their properties different from this simple, ideal picture. First, the finite number N of slits, which may be described by limiting the sum (125) to the interval $n = [-N/2, +N/2]$, results in a non-zero spread, $\delta\theta/\theta \sim 1/N$, of each diffraction peak, and hence in the reduction of the grating's spectral resolution. (Unintentional variations of the inter-slit distance d have a similar effect, so before the advent of high-resolution photolithography, special high-precision mechanical tools had been used for grating fabrication.)

Second, a finite slit width w leads to the diffraction peak pattern modulation by the $\mathrm{sinc}^2(kw\theta/2)$ envelope, similar to the pattern shown in Fig. 15. Actually, for spectroscopic purposes, such modulation is sometimes a plus, because only one diffraction peak (say, with $m = \pm 1$) is practically used, and if the frequency spectrum of the analyzed wave is very broad (covers more than one octave), the higher peaks produce undesirable hindrance. Because of this reason, w is frequently selected to be equal exactly to $d/2$, thus suppressing each other diffraction maximum. Moreover, sometimes semi-transparent films are used to make the transparency function $T(\mathbf{r}')$ continuous and close to a sinusoidal one:

$$T(\boldsymbol{\rho}') \approx T_0 + T_1 \cos\frac{2\pi x'}{d} \equiv T_0 + \frac{T_1}{2}\left(\exp\left\{i\frac{2\pi x'}{d}\right\} + \exp\left\{-i\frac{2\pi x'}{d}\right\}\right). \tag{8.127}$$

Plugging the last expression into Eq. (116) and integrating, we see that the output wave consists of just 3 components: the direct-passing wave (proportional to T_0) and two diffracted waves (proportional to T_1) propagating in the directions of the two lowest Bragg angles, $\theta_{\pm 1} = \pm\lambda/d$.[50]

[49] The rudimentary diffraction grating effect, produced by the parallel fibers of a bird's feather, was discovered as early as 1673 by James Gregory (who also invented the "Gregorian" telescope – one of the basic designs for reflecting telescopes).

The same Eq. (116) may be also used to obtain one more general (and rather curious) result, called the *Babinet principle*.[51] Consider two experiments with the diffraction of similar plane waves on two "complementary" screens – such that together they would cover the whole plane, without a hole or an overlap. (Think, for example, about an opaque disk of radius R and a large opaque screen with a round orifice of the same radius.) Then, according to the Babinet principle, the diffracted wave patterns produced by these two screens in all directions with $\theta \neq 0$ are *identical*.

The proof of this principle is straightforward: since the transparency functions produced by the screens are complementary:

$$T(\boldsymbol{\rho}') \equiv T_1(\boldsymbol{\rho}') + T_2(\boldsymbol{\rho}') = 1, \tag{8.128}$$

and the diffracted wave is (in the Fraunhofer approximation only!) a Fourier transform of $T(\boldsymbol{\rho}')$, which is a linear operation, we get

$$f_1(\boldsymbol{\rho}) + f_2(\boldsymbol{\rho}) = f_0(\boldsymbol{\rho}), \tag{8.129}$$

where f_0 is the wave "scattered" by the composite screen with $T_0(\boldsymbol{\rho}') \equiv 1$, i.e. the unperturbed initial wave propagating in the initial direction ($\theta = 0$). In all other directions, $f_1 = -f_2$, i.e. the diffracted waves are indeed similar besides the difference in sign – which is equivalent to a phase shift by $\pm\pi$. However, it is important to remember that the Babinet principle notwithstanding, in real experiments, with screens at finite distances, the diffracted waves may interfere with the unperturbed plane wave $f_0(\boldsymbol{\rho})$, leading to different diffraction patterns in cases 1 and 2 – see, e.g., Fig. 14 and its discussion.

8.9. Magnetic dipole and electric quadrupole radiation

Throughout this chapter, we have seen how many important results may be obtained from Eq. (26) for the electric dipole radiation by a small-size source (Fig. 1). Only in rare cases when this radiation is absent, for example, if the dipole moment $\mathbf{p}$ of the source equals zero (or does not change in time – either at all or at the frequency of our interest), higher-order effects may be important. I will now discuss the main two of them, *quadrupole electric radiation* and *dipole magnetic radiation*.

In Sec. 2 above, the electric dipole radiation was calculated by plugging the expansion (19) into the exact formula (17b) for the retarded vector potential $\mathbf{A}(\mathbf{r}, t)$. Let us make a more exact calculation, by keeping the second term of that expansion as well:

$$\mathbf{j}\left(\mathbf{r}', t - \frac{R}{v}\right) \approx \mathbf{j}\left(\mathbf{r}', t - \frac{r}{v} + \frac{\mathbf{r}'\cdot\mathbf{n}}{v}\right) \equiv \mathbf{j}\left(\mathbf{r}', t' + \frac{\mathbf{r}'\cdot\mathbf{n}}{v}\right), \qquad \text{where } t' \equiv t - \frac{r}{v}. \tag{8.130}$$

Since the expansion is only valid if the last term in the time argument of $\mathbf{j}$ is relatively small, in the Taylor expansion of $\mathbf{j}$ with respect to that argument we may keep just two leading terms:

$$\mathbf{j}\left(\mathbf{r}', t' + \frac{\mathbf{r}'\cdot\mathbf{n}}{v}\right) \approx \mathbf{j}(\mathbf{r}', t') + \frac{\partial \mathbf{j}(\mathbf{r}', t')}{\partial t'}\frac{(\mathbf{r}'\cdot\mathbf{n})}{v}, \tag{8.131}$$

[50] Similar tricks are used in the so-called *phased-array antennas*, broadly used in radar systems and radioastronomy, in which electronically controlled mutual phase shifts of microwave signals feeding many similar component antennas are used to steer the direction of the resulting narrow beam. For more on this important technology, see, e.g. T. Milligan, *Modern Antenna Design,* 2nd ed., Wiley (2005).

[51] Named after Jacques Babinet (1784-1874) who made several important contributions to optics.

so Eq. (17b) yields $\mathbf{A} = \mathbf{A}_d + \mathbf{A}'$, where $\mathbf{A}_d$ is the electric dipole contribution as given by Eq. (23), and $\mathbf{A}'$ is the new term of the next order in the small parameter $r' << r$:

$$\mathbf{A}'(\mathbf{r},t) = \frac{\mu}{4\pi r v}\frac{\partial}{\partial t'}\int \mathbf{j}(\mathbf{r}',t')(\mathbf{r}'\cdot\mathbf{n})d^3r'. \tag{8.132}$$

Just as it was done in Sec. 2, let us evaluate this term for a system of non-relativistic particles with electric charges q_k and radius vectors $\mathbf{r}_k(t)$:

$$\mathbf{A}'(\mathbf{r},t) = \frac{\mu}{4\pi r v}\left[\frac{d}{dt}\sum_k q_k \dot{\mathbf{r}}_k(\mathbf{r}_k\cdot\mathbf{n})\right]_{t=t'}. \tag{8.133}$$

Using the "bac minus cab" identity of the vector algebra again,[52] the vector operand of Eq. (133) may be rewritten as

$$\begin{aligned}\dot{\mathbf{r}}_k(\mathbf{r}_k\cdot\mathbf{n}) &\equiv \frac{1}{2}\dot{\mathbf{r}}_k(\mathbf{r}_k\cdot\mathbf{n}) + \frac{1}{2}\dot{\mathbf{r}}_k(\mathbf{n}\cdot\mathbf{r}_k) = \frac{1}{2}(\mathbf{r}_k\times\dot{\mathbf{r}}_k)\times\mathbf{n} + \frac{1}{2}\mathbf{r}_k(\mathbf{n}\cdot\dot{\mathbf{r}}_k) + \frac{1}{2}\dot{\mathbf{r}}_k(\mathbf{n}\cdot\mathbf{r}_k)\\ &\equiv \frac{1}{2}(\mathbf{r}_k\times\dot{\mathbf{r}}_k)\times\mathbf{n} + \frac{1}{2}\frac{d}{dt}[\mathbf{r}_k(\mathbf{n}\cdot\mathbf{r}_k)],\end{aligned} \tag{8.134}$$

so the right-hand side of Eq. (133) may be represented as a sum of two terms, $\mathbf{A}' = \mathbf{A}_m + \mathbf{A}_q$, where

$$\mathbf{A}_m(\mathbf{r},t) = \frac{\mu}{4\pi r v}\dot{\mathbf{m}}(t')\times\mathbf{n} \equiv \frac{\mu}{4\pi r v}\dot{\mathbf{m}}\left(t-\frac{r}{v}\right)\times\mathbf{n}, \qquad \text{with } \mathbf{m}(t) \equiv \frac{1}{2}\sum_k \mathbf{r}_k(t)\times q_k\dot{\mathbf{r}}_k(t); \tag{8.135}$$

$$\mathbf{A}_q(\mathbf{r},t) = \frac{\mu}{8\pi r v}\left[\frac{d^2}{dt^2}\sum_k q_k\mathbf{r}_k(\mathbf{n}\cdot\mathbf{r}_k)\right]_{t=t'}. \tag{8.136}$$

Comparing the second of Eqs. (135) with Eq. (5.91), we see that $\mathbf{m}$ is just the total magnetic moment of the source. On the other hand, the first of Eqs. (135) is absolutely similar in structure to Eq. (23), with $\mathbf{p}$ replaced with $(\mathbf{m}\times\mathbf{n})/v$, so for the corresponding component of the magnetic field it gives (in the same approximation $r >> \lambda$) a result similar to Eq. (24):

Magnetic dipole radiation: field

$$\mathbf{B}_m(\mathbf{r},t) = \frac{\mu}{4\pi r v}\nabla\times\left[\dot{\mathbf{m}}\left(t-\frac{r}{v}\right)\times\mathbf{n}\right] = -\frac{\mu}{4\pi r v^2}\mathbf{n}\times\left[\ddot{\mathbf{m}}\left(t-\frac{r}{v}\right)\times\mathbf{n}\right]. \tag{8.137}$$

According to this expression, just as at the electric dipole radiation, the vector $\mathbf{B}$ is perpendicular to the vector $\mathbf{n}$, and its magnitude is also proportional to $\sin\Theta$, where Θ is now the angle between the direction toward the observation point and the second time derivative of the vector $\mathbf{m}$ – rather than $\mathbf{p}$:

$$B_m = \frac{\mu}{4\pi r v^2}\ddot{m}\left(t-\frac{r}{v}\right)\sin\Theta. \tag{8.138}$$

As a result, the intensity of this *magnetic dipole radiation* has a similar angular distribution:

Magnetic dipole radiation: power

$$S_r = ZH^2 = \frac{Z}{(4\pi v^2 r)^2}\left[\ddot{m}\left(t-\frac{r}{v}\right)\right]^2\sin^2\Theta \tag{8.139}$$

[52] If you still need it, see MA Eq. (7.5).

- cf. Eq. (26), besides the (generally) different meaning of the angle Θ.

Note, however, that this radiation is usually much weaker than its electric-dipole counterpart. For example, for a non-relativistic particle with electric charge q, moving on a trajectory of linear size ~a, the electric dipole moment is of the order of qa, while its magnetic moment scales as $qa^2\omega$, where ω is the motion frequency. As a result, the ratio of the magnetic and electric dipole radiation intensities is of the order of $(a\omega/v)^2$, i.e. the squared ratio of the particle's speed to the speed of the emitted waves – that has to be much smaller than 1 for our non-relativistic calculation to be valid.

The angular distribution of the *electric quadrupole radiation* described by Eq. (136) is more involved. To show this, let us add to $\mathbf{A}_q$ a vector parallel to $\mathbf{n}$ (i.e. directed along the wave's propagation), getting

$$\mathbf{A}_q(\mathbf{r},t) \to \frac{\mu}{24\pi r v}\ddot{\mathscr{Q}}\left(t-\frac{r}{v}\right), \quad \text{where } \mathscr{Q} \equiv \sum_k q_k\left\{3\mathbf{r}_k(\mathbf{n}\cdot\mathbf{r}_k)-\mathbf{n}r_k^2\right\}, \tag{8.140}$$

since this addition does not contribute to the transverse components of the electric and magnetic fields, i.e. to the radiated wave. According to the above definition of the vector $\mathscr{Q}$, its Cartesian components may be represented as

$$\mathscr{Q}_j = \sum_{j'=1}^{3}\mathscr{Q}_{jj'}n_{j'}, \tag{8.141}$$

where $\mathscr{Q}_{jj'}$ are the elements of the electric quadrupole tensor of the system – see the last of Eqs. (3.4):[53]

$$\mathscr{Q}_{jj'} = \sum_k q_k\left(3r_j r_{j'}-r^2\delta_{jj'}\right)_k . \tag{8.142}$$

Now taking the curl of the first of Eqs. (140) at $r >> \lambda$, we get

$$\mathbf{B}_q(\mathbf{r},t) = -\frac{\mu}{24\pi r v^2}\mathbf{n}\times\dddot{\mathscr{Q}}\left(t-\frac{r}{v}\right). \tag{8.143}$$

Electric quadrupole radiation: field

This expression is similar to Eqs. (24) and (137), but according to Eqs. (140) and (142), components of the vector $\mathscr{Q}$ do depend on the direction of the vector $\mathbf{n}$, leading to a different angular dependence of S_r.

As the simplest example, let us consider the system of two equal point electric charges moving symmetrically, at equal distances $d(t) << \lambda$ from a stationary center – see Fig. 16.

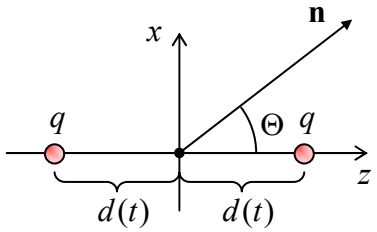


Fig. 8.16. The simplest system emitting electric quadrupole radiation.

Due to the symmetry of the system, its dipole moments $\mathbf{p}$ and $\mathbf{m}$ (and hence its electric and magnetic dipole radiation) vanish, but the quadrupole tensor (142) still has non-zero elements. With the coordinate choice shown in Fig. 16, these elements are diagonal:

[53] Let me hope that the reader has already acquired some experience in the calculation of this tensor's elements – e.g., for the simple systems specified in Problems 3.2-3.4.

$$\mathcal{Q}_{xx} = \mathcal{Q}_{yy} = -2qd^2, \qquad \mathcal{Q}_{zz} = 4qd^2. \tag{8.144}$$

With the x-axis selected within the common plane of the z-axis and the direction $\mathbf{n}$ toward the observation point (Fig. 16), so $n_x = \sin\Theta$, $n_y = 0$, and $n_z = \cos\Theta$, Eq. (141) yields

$$\mathcal{Q}_x = -2qd^2 \sin\Theta, \qquad \mathcal{Q}_y = 0, \qquad \mathcal{Q}_z = 4qd^2 \cos\Theta, \tag{8.145}$$

and the vector product in Eq. (143) has only one non-vanishing Cartesian component:

$$\left(\mathbf{n} \times \dddot{\mathcal{Q}}\right)_y = n_z \dddot{\mathcal{Q}}_x - n_x \dddot{\mathcal{Q}}_z = -6q \sin\Theta \cos\Theta \frac{d^3}{dt^3}\left[d^2(t)\right]. \tag{8.146}$$

As a result, the quadrupole radiation intensity, $S \propto B_q^2$, is proportional to $\sin^2\Theta\cos^2\Theta$, i.e. vanishes not only along the symmetry axis of the system (as the electric-dipole and the magnetic-dipole radiations would), but also in all directions perpendicular to this axis, reaching its maxima at $\Theta = \pm\pi/4$.

For more complex systems, the angular distribution of the electric quadrupole radiation may be different, but it may be proved that its total (instant) power always obeys the following simple formula:

Electric quadrupole radiation: power

$$\mathcal{P}_q = \frac{Z}{720\pi v^4} \sum_{j,j'=1}^{3} \left(\dddot{\mathcal{Q}}_{jj'}\right)^2. \tag{8.147}$$

Let me finish this section by giving, also without proof, one more fact important for some applications: due to their different spatial structure, the magnetic-dipole and electric-quadrupole radiation fields do not interfere, i.e. the total power of radiation (neglecting the electric-dipole and higher multipole terms) may be found as the sum of these components, calculated independently. On the contrary, the electric-dipole and magnetic-dipole radiations of the same system typically interfere coherently, so their radiation fields (rather than powers) should be summed up.

8.10. Exercise problems

8.1. Equation (8) obviously has standing-wave solutions $\chi(r, t) = \mathrm{Re}\,[C\sin kr \exp\{-i\omega t\}]$, turning the scalar potential $\phi = \chi/r$ into a finite constant at $r = 0$ and into zero at $kr = \pi n$, with $n = 0, 1, 2, \ldots$ This fact seems to imply that a cavity of radius R, carved inside a good conductor, has resonant modes with a purely radial electric field $\mathbf{E}(\mathbf{r}) = \mathbf{n}_r E(r)$ and that the lowest nonvanishing of them, with $k = \pi/R$, gives the lowest (fundamental) frequency $\omega \equiv vk = \pi(v/R)$ of the cavity. Is this conclusion correct?

8.2. Simplify the Lorentz reciprocity theorem (6.121) for space-localized field sources. Then find out what it says about the fields of two compact, well-separated sources of electric-dipole radiation.

8.3. In the electric-dipole approximation, calculate the angular distribution and the total power of electromagnetic radiation by the hydrogen atom within the following classical model: an electron rotates, at a constant distance R, about a much heavier proton. Use this result to calculate the law of a gradual reduction of R in time. Finally, evaluate the classical lifetime of the atom by borrowing the initial value of R from quantum mechanics: $R(0) = r_B \approx 0.53\times10^{-10}$ m.

8.4. A non-relativistic particle of mass m, with electric charge q, is placed into a time-independent uniform magnetic field $\mathbf{B}$. Derive the law of decrease of the particle's kinetic energy due to

its electromagnetic radiation at the *cyclotron frequency* $\omega_c = qB/m$. Evaluate the rate of such *radiation cooling* of electrons in a magnetic field of 1 T, and estimate the energy interval in which this result is quantitatively correct.

Hint: The cyclotron motion will be discussed in detail (for arbitrary particle velocities) in Sec. 9.6 below, but I hope that the reader already knows that in the non-relativistic case ($v << c$), the above formula for ω_c may be readily obtained by combining the 2nd Newton law $mv_\perp^2/R = qv_\perp B$ for the particle's circular rotation under the effect of the magnetic component of the Lorentz force (5.10), and the geometric relation $v_\perp = R\omega_c$. (Here $\mathbf{v}_\perp$ is the particle's velocity in the plane normal to the vector **B**.)

8.5. A particle with mass m, electric charge q, and an initial kinetic energy $T << mc^2$ collides head-on with a much more massive particle of charge $\mathcal{F}q$, in free space. Calculate the total energy of electromagnetic radiation during this collision, assuming it to be much lower than T.

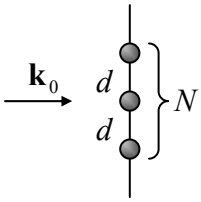
8.6. Solve the dipole antenna radiation problem discussed in Sec. 2 (see Fig. 3) for the optimal value $l = \lambda/2$ of its length, assuming that the current distribution in each of its arms is sinusoidal: $I(z, t) = I_0\cos(\pi z/l)\cos\omega t$. [54]

8.7. A plane wave is scattered by a localized object in free space. Relate the differential cross-section of the wave's scattering to the average force it exerts on the object. Use this general relation to calculate the force exerted by a plane monochromatic wave on a free non-relativistic particle and compare the result with those obtained in Problems 7.4 and 7.5.

8.8. Use the Lorentz oscillator model of a bound charge, given by Eq. (7.30), to explore the transition between the two scattering limits discussed in Sec. 3 and, in particular, the *resonant scattering* taking place at $\omega \approx \omega_0$. In the last context, discuss the contribution of scattering to the oscillator's damping.

8.9.* A sphere of radius R, made of a material with a uniform permanent electric polarization $\mathbf{P}_0$ and a constant mass density ρ, is free to rotate about its center. Calculate its average total cross-section for scattering of a linearly polarized plane electromagnetic wave of frequency $\omega << R/c$, incident from free space, in the weak-wave limit, assuming that the initial orientation of the polarization vector $\mathbf{P}_0$ is random.

8.10. Use Eq. (56) to analyze the interference/diffraction pattern produced by a plane wave's scattering on a set of N similar equidistant small objects on a straight line normal to the direction of the incident wave's propagation – see the figure on the right. Discuss the trend(s) of the pattern in the limit $N \to \infty$.



8.11. Use the Born approximation to calculate the differential cross-section of a plane wave's scattering by a uniform dielectric sphere of an arbitrary radius R. In the limits $kR << 1$ and $1 << kR$ (where k is the wave number), analyze the angular dependence of the differential cross-section and calculate the total cross-section of scattering.

[54] As was emphasized in Sec. 2, this is a reasonable guess rather than a controllable approximation. The exact (rather involved!) theory shows that this assumption gives errors ~5%, depending on the wire's diameter.

8.12. A sphere of radius R is made of a uniform dielectric material, with an arbitrary dielectric constant. Calculate its total cross-section of scattering a linearly-polarized low-frequency ($k << 1/R$) wave and compare the result with the solution of the previous problem.

8.13. Use the Born approximation to calculate the differential cross-section of a plane wave's scattering on a right circular cylinder of length l and radius R, for an arbitrary angle of incidence.

8.14. Formulate the quantitative condition of the Born approximation's validity for a uniform dielectric scatterer, with all linear dimensions of the order of the same scale a.

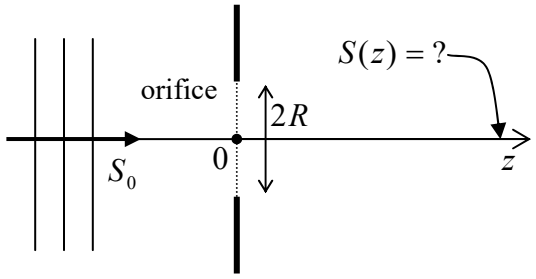
8.15. If a scatterer absorbs some part of the incident wave's power, it may be characterized by an *absorption cross-section* σ_a defined similarly to Eq. (39) for the scattering cross-section:

$$\sigma_a \equiv \frac{\overline{\mathscr{P}_a}}{|E_\omega|^2 / 2Z_0},$$

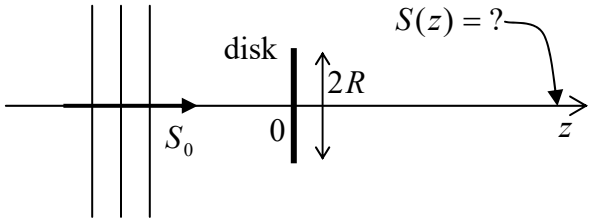
where the numerator is the time-averaged absorbed power. Use two different approaches to calculate σ_a of a very small sphere of radius $R << k^{-1}, \delta_s$, made of a nonmagnetic material with an Ohmic conductivity σ and the high-frequency permittivity $\varepsilon_{opt} = \varepsilon_0$. Can σ_a of such a sphere be larger than its geometric cross-section πR^2?

8.16. Use the Huygens principle to calculate the wave's intensity on the symmetry plane of the slit diffraction experiment (i.e. at $x = 0$ in Fig. 12), for arbitrary ratio z/ka^2.

8.17. A plane wave with wavelength λ is normally incident on an opaque planar screen with a round orifice of radius $R >> \lambda$. Use the Huygens principle to calculate the passing wave's intensity on the system's symmetry axis, at distances $z >> R$ from the screen (see the figure on the right), and analyze the result.



8.18. A plane monochromatic wave is now normally incident on an opaque circular disk of radius $R >> \lambda$. Use the Huygens principle to calculate the wave's intensity at a distance $z >> R$ behind the disk's center – see the figure on the right. Discuss the result.

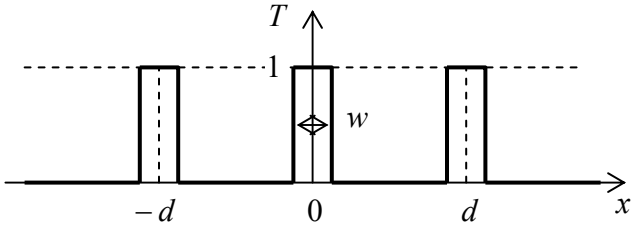


8.19. Use the Huygens principle to analyze the Fraunhofer diffraction of a plane wave normally incident on a square-shaped hole, of size $a \times a$, in an opaque screen. Sketch the diffraction pattern you would observe at a sufficiently large distance, and quantify the expression "sufficiently large" for this case.

8.20. Use the Huygens principle to analyze the propagation of a monochromatic Gaussian beam described by Eq. (7.181), with the initial characteristic width $a_0 >> \lambda$, in a uniform isotropic medium. Use the result for a semi-quantitative derivation of the so-called *Abbe limit* for the spatial resolution of

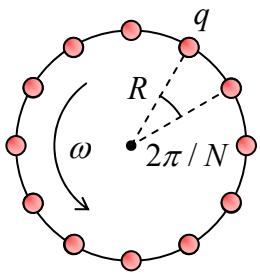
an optical system: $w_{min} = \lambda/2\sin\theta$, where θ is the half-angle of the wave cone propagating from the object and captured by the system.

8.21. Within the Fraunhofer approximation, analyze the pattern produced by a diffraction grating with the 1D-periodic transparency profile shown in the figure on the right, for the normal incidence of a monochromatic plane wave.



8.22. N equal point charges are attached, at equal intervals, to a circle rotating with a constant angular velocity about its center – see the figure on the right. For what values of N does the system emit:

(i) the electric dipole radiation?
(ii) the magnetic dipole radiation?
(iii) the electric quadrupole radiation?



8.23. What general statements can you make about:

(i) the electric dipole radiation, and
(ii) the magnetic dipole radiation,

due to a collision of an arbitrary number of similar non-relativistic classical particles?

8.24. Calculate the angular distribution and the total power radiated by a small planar loop antenna of radius R, fed with ac current with frequency ω and amplitude I_0, into free space.

8.25. The orientation of a magnetic dipole, with a constant magnitude m of its moment, is rotating about a certain axis with an angular velocity ω, with the angle α between them staying constant. Calculate the angular distribution and the average power of its radiation into the free space.

8.26. Solve Problem 12 (also in the low-frequency limit $kR << 1$), for the case when the sphere's material has a frequency-independent Ohmic conductivity σ, and $\varepsilon_{opt} = \varepsilon_0$, in two limits:

(i) of a very large skin depth ($\delta_s >> R$), and
(ii) of a very small skin depth ($\delta_s << R$).

8.27. Complete the solution of the problem started in Sec. 9, by calculating the full power of radiation of the system of two charges oscillating in antiphase along the same straight line – see Fig. 16. Also, calculate the average radiation power for the case of harmonic oscillations, $d(t) = a\cos\omega t$, compare it with the case of a single charge performing similar oscillations, and interpret the difference.

8.28. The system of four alternating charges located at the angles of a square, considered in Problem 3.3(i), is now being rotated around the axis normal to their plane and passing through the square's center, with a constant angular frequency $\omega << v/a$. Calculate the time-averaged angular distribution and the total power of the resulting radiation.