Reflection from metals

We've talked about reflectance and transmittance from dielectrics. We also need to take a long look in the mirror and consider how light waves interact with conductors, which obey Ohm's law of $I = (1/R)V$ or $J = \sigma E$ in terms of a current density J, electric field E, and conductance σ. With this, Maxwell's equations in 1D become

$$\frac{\partial^2}{\partial x^2}E = \mu\epsilon\frac{\partial^2 E}{\partial t^2} + \mu\sigma\frac{\partial E}{\partial t} \quad (1)$$

$$\frac{\partial^2}{\partial x^2}B = \mu\epsilon\frac{\partial^2 B}{\partial t^2} + \mu\sigma\frac{\partial B}{\partial t} \quad (2)$$

Since E and B have the same form, we'll consider E only in what follows. We'll now do our usual thing of considering solutions of the form $E = E_0 e^{-i(kx-\omega t)}$.
Some conductivities σ in $1/(\Omega \cdot \text{meters})$: undoped silicon 1.2×10^3, aluminum 3.77×10^7, copper 5.96×10^7, silver 6.30×10^7.
Aluminum: magnetic susceptibility is $\chi_m = (\mu/\mu_0 - 1) = 2.1 \times 10^{-5}$.

Metals II

To consider solutions of the form $E = E_0 e^{-i(kx-\omega t)}$, we need to take some derivatives:

$$\begin{aligned}
\frac{\partial}{\partial t} E &= (i\omega) E_0 e^{-i(kx-\omega t)} \\
\frac{\partial^2}{\partial t^2} E &= \frac{\partial}{\partial t} (i\omega) E_0 e^{-i(kx-\omega t)} = (-\omega^2) E_0 e^{-i(kx-\omega t)} \\
\frac{\partial}{\partial x} E &= (-ik) E_0 e^{-i(kx-\omega t)} \\
\frac{\partial^2}{\partial x^2} E &= \frac{\partial}{\partial x} (-ik) E_0 e^{-i(kx-\omega t)} = (-k^2) E_0 e^{-i(kx-\omega t)}
\end{aligned} \tag{3}$$

We can use these results for expanding Eq. 1:

$$\begin{aligned}
\frac{\partial^2}{\partial x^2} E &= \mu\epsilon \frac{\partial^2 E}{\partial t^2} + \mu\sigma \frac{\partial E}{\partial t} \\
(-k^2) E_0 e^{-i(kx-\omega t)} &= (-\omega^2)\mu\epsilon E_0 e^{-i(kx-\omega t)} + \mu\sigma (i\omega) E_0 e^{-i(kx-\omega t)}
\end{aligned} \tag{4}$$

Metals III

From the differential equation of Eq. 4, we have

$$\begin{aligned} k^2 &= \omega^2 \mu\epsilon - i\omega\mu\sigma \\ k &= \omega\sqrt{\mu\epsilon}\sqrt{1 - i\frac{\sigma}{\omega\epsilon}} = k_r - ik_i \end{aligned} \tag{5}$$

so that wave propagation becomes

$$E_0 e^{-i(kx-\omega t)} = E_0 e^{-i(k_r x-\omega t)}\, e^{-k_i x} \tag{6}$$

with a phase velocity of $v_p = \omega/k_r$ and an amplitude attenuation of $e^{-k_i x}$.

Metals IV

Again, we had from Eq. 5 the form

$$k = \omega\sqrt{\mu\epsilon}\sqrt{1 - i\frac{\sigma}{\omega\epsilon}} = k_r - ik_i$$

How should we handle the factor $\sqrt{1 - i\sigma/(\omega\epsilon)}$? At optical frequencies ($\lambda = 500$ nm corresponds to $\omega = 2\pi c/\lambda = 3.8 \times 10^{15}$ rad/sec), let's consider values for two different materials:

$$\begin{aligned} \text{Aluminum:} \quad & \frac{\sigma}{\omega\epsilon_0} = \frac{3.77 \times 10^{7}}{3.8 \times 10^{15} \cdot 8.85 \times 10^{-12}} = 1.1 \times 10^{3} \\ \text{Undoped silicon:} \quad & \frac{\sigma}{\omega\epsilon} = \frac{1.2 \times 10^{3}}{3.8 \times 10^{15} \cdot 16 \cdot 8.85 \times 10^{-12}} = 2.2 \times 10^{-3} \end{aligned}$$

where for silicon we have a dielectric constant of $k = \epsilon/\epsilon_0 = 16$. As you can see, we have two entirely different limits for $\sqrt{1 - i\sigma/(\omega\epsilon)}$.

Metals: low conductivity limit

Let's first consider the case of low conductivity $\sigma \ll \omega\epsilon$, such as the case of undoped silicon. In that case, we can approximate k from Eq. 5 as

$$k = \omega\sqrt{\mu\epsilon}\sqrt{1 - i\frac{\sigma}{\omega\epsilon}} \simeq \omega\sqrt{\mu\epsilon}\left(1 - i\frac{\sigma}{2\omega\epsilon}\right) \tag{7}$$

where we've used the binomial expansion. The phase velocity then becomes

$$v = \frac{\omega}{k_r} = \frac{\omega}{\omega\sqrt{\mu\epsilon}} = \frac{c}{n} \tag{8}$$

but the light intensity is attenuated according to $(e^{-k_i x})^2 = e^{-2k_i x}$ with a $1/e$ attenuation distance of

$$\frac{1}{2k_i} = \frac{2\omega\epsilon}{\omega\sigma\sqrt{\mu\epsilon}} = \frac{2}{\sigma}\sqrt{\frac{\epsilon}{\mu}}. \tag{9}$$

For undoped silicon, $1/(2k_i) = 4.4 \times 10^{-6}$ m or 4.4 μm.

Metals: high conductivity limit

If instead we consider the high conductivity limit of $\sigma \gg \omega\epsilon$ (which we saw earlier was the case for aluminum), we can approximate k from Eq. 5 as

$$k = \omega\sqrt{\mu\epsilon}\sqrt{1 - i\frac{\sigma}{\omega\epsilon}} \simeq \omega\sqrt{\mu\epsilon}\sqrt{-i\frac{\sigma}{\omega\epsilon}} = \sqrt{-i\sigma\mu\omega}. \tag{10}$$

Now $\sqrt{-i} = \sqrt{e^{-i\pi/2}} = (e^{-i\pi/2})^{1/2} = e^{-i\pi/4} = (1-i)/\sqrt{2}$, so we have $k = k_r - ik_i$ with $|k_r| = |k_i|$, and

$$k_r = -k_i = \sqrt{\frac{\sigma\mu\omega}{2}}. \tag{11}$$

The wave intensity is attenuated according to $(e^{-i(k_r - ik_i)x})^2 = (e^{-ik_r x}e^{-k_i x})^2 \Rightarrow e^{-2k_i x}$ with a $1/e$ attenuation distance of

$$\frac{1}{2k_i} = \frac{\sqrt{2}}{2\sqrt{\sigma\mu\omega}} = \frac{1}{\sqrt{2\sigma\mu\omega}} \tag{12}$$

which works out to be only 1.67×10^{-9} meters or 16.7 Å, corresponding to the distance of a few atoms! Mirrors need only a very thin metal coating!

Metals: high conductivity limit II

The index of refraction in this case is

$$n = \frac{c}{v_p} = \frac{k}{\omega\sqrt{\mu_0\epsilon_0}} = \frac{1}{\omega\sqrt{\mu_0\epsilon_0}}\sqrt{\frac{\sigma\mu\omega}{2}}\frac{(1-i)}{\sqrt{2}} = \sqrt{\frac{\sigma}{\omega\epsilon_0}}\frac{(1-i)}{2} \quad (13)$$

where we have assumed $\mu = \mu_0$ or $\chi_m \ll 1$. For aluminum with green light, we have $n = 16.7(1-i)$.

Reflection coefficients for metals

The Fresnel equations gave reflection coefficients of

$$r_{\mathrm{TE}} = \frac{\cos\theta - \sqrt{n^2 - \sin^2\theta}}{\cos\theta + \sqrt{n^2 - \sin^2\theta}} \tag{14}$$

$$r_{\mathrm{TM}} = \frac{n^2\cos\theta - \sqrt{n^2 - \sin^2\theta}}{n^2\cos\theta + \sqrt{n^2 - \sin^2\theta}} \tag{15}$$

If we have $n \gg 1$, the $\sin\theta$ and $\cos\theta$ terms will be small compared to n and the reflectivity will approach $R \to 1$.

Waves from point sources

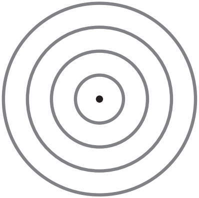
- The wave equation for a plane wave is

$$\psi = \psi_0 \exp\left[-i(\vec{k}\cdot\vec{z} - \omega t)\right]. \quad (16)$$

- For a spherical wave, the wave equation is

$$\psi = \psi_0 \frac{\lambda}{r} \exp\left[-i(kr - \omega t)\right]. \quad (17)$$

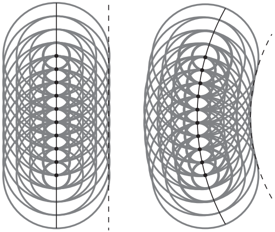
  Notice that k is no longer a vector because the spherical wave goes every which way; no direction is defined.



- How can we connect the two?

The Huygens construction

- Christian Huygens (1629–1695): consider many point sources placed next to each other and in phase. Looks like a plane wave!
- Consider many point sources but with a phase delay/advance, and also magnitude modulation, that depends on position. Refer to this magnitude and phase modulation (which is complex) as $\tilde{g}(x_0, y_0)$.
- In other words, we treat each point in the input plane $(x_0, y_0, 0)$ as a Huygens point source with amplitude $\tilde{g}(x_0, y_0)$.



Adding Huygens sources

To get the amplitude $\psi(x,y,z)$ at a downstream point (x,y,z), we want to add up all the Huygens point sources (ignoring the time dependance term $\exp[i\omega t]$):

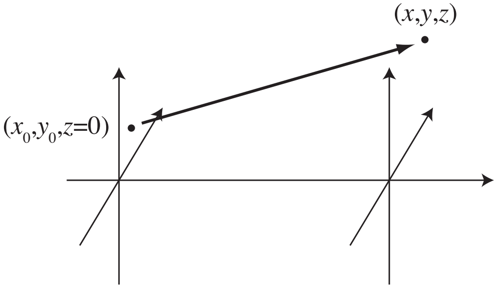
$$\psi(x,y,z) = \psi_0 \frac{\lambda}{A} \int_{x_0} \int_{y_0} \tilde{g}(x_0,y_0) \frac{\exp[-ikr]}{r} \cos\theta. \tag{18}$$

The $1/A$ term is for an integration area to cancel out $\int\int dx_0\, dy_0$, and the $\cos\theta$ term is an obliquity factor which we can normally ignore. Since ψ is an amplitude (amplitude=magnitude$\cdot$ exp[$i$phase]), the irradiance E is given by $E = \psi\,\psi^\dagger$.

Adding Huygens sources II

The radius r from the spherical wave source is given by

$$r = \sqrt{z^2 + (x - x_0)^2 + (y - y_0)^2} = z\sqrt{1 + \frac{(x - x_0)^2}{z^2} + \frac{(y - y_0)^2}{z^2}}. \tag{19}$$



Propagation distances

Again, from Eq. 19 we have

$$r = \sqrt{z^2 + (x - x_0)^2 + (y - y_0)^2} = z\sqrt{1 + \frac{(x - x_0)^2}{z^2} + \frac{(y - y_0)^2}{z^2}}$$

In the limit of $[(x - x_0)^2 + (y - y_0)^2 \ll z^2$, we can write this as

$$\begin{aligned} r &\simeq z\left[1 + \frac{(x - x_0)^2}{2z^2} + \frac{(y - y_0)^2}{2z^2} - \frac{(x - x_0)^4}{8z^4} - \frac{(y - y_0)^4}{8z^4} + \ldots\right] \\ &\simeq z\left[1 + \frac{x^2 + y^2}{2z^2} + \frac{x_0^2 + y_0^2}{2z^2} - \frac{xx_0 + yy_0}{z^2}\right] \end{aligned} \quad (20)$$

The second version of the above assumes the *Fresnel approximation*, where we ignore the x^4/z^4-type terms. Let $\rho = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ represent transverse distances; the Fresnel approximation assumes

$$\frac{2\pi z}{\lambda}\frac{\rho^4}{8z^4} \ll \frac{\pi}{2} \qquad \text{or} \qquad z^3 \gg \frac{\rho^4}{2\lambda} \quad (21)$$

Numerical example

- Again, the Fresnel approximation assumes $z^3 \gg \rho^4/(2\lambda)$. If we say that ρ=0.5 mm and $\lambda = 500$ nm, we require $z^3 \gg (0.4\text{mm})^3$.
- Since the numerical aperture of a lens is defined as N.A.$= n \sin\theta$, where θ is the angle subtended by the lens relative to the optical axis, in this example we effectively have N.A.$\ll 1 \sin(\arctan \rho/z)$ or N.A.< 0.4 if we pick $z > 1$ mm.
  - This means that high resolution microscope objectives which have large N.A. must be designed without the use of the Fresnel approximation.
  - Note: one can relate the numerical aperture and the "f number" or $f^{\#}$—which is the focal length divided by diameter—of a lens with N.A.$= 2/f^{\#}$.

Fresnel approximation II

Again, we want to write the expression of Eq. 18 of

$$\psi(x,y,z) = \psi_0 \frac{\lambda}{A} \int_{x_0} \int_{y_0} \tilde{g}(x_0,y_0) \frac{\exp[-ikr]}{r} \cos\theta$$

using the Fresnel approximation version (Eq. 20) of expanding the source-to-measurement distance r giving

$$kr \simeq \frac{2\pi z}{\lambda} \left[1 + \frac{x^2+y^2}{2z^2} + \frac{x_0^2+y_0^2}{2z^2} - \frac{xx_0+yy_0}{z^2} \right]$$

We'll use this expansion for r in the $\exp[ikr]$ phase term, while in the $1/r$ magnitude terms we'll simply use $1/z$. We then have

$$\begin{aligned} \psi(x,y,z) &= \psi_0 \frac{\lambda}{z} \frac{1}{A} \int_{x_0} \int_{y_0} \tilde{g}(x_0,y_0) \qquad (22) \\ & \exp\left[-i\frac{2\pi z}{\lambda} - i\pi \frac{x^2+y^2}{\lambda z} - i\pi \frac{x_0^2+y_0^2}{\lambda z} + i2\pi \frac{xx_0+yy_0}{\lambda z} \right] \end{aligned}$$

Fresnel approximation III

Again, we have from Eq. 22 the expression

$$\begin{aligned}\psi(x,y,z) &= \psi_0 \frac{\lambda}{z}\frac{1}{A}\int_{x_0}\int_{y_0}\tilde{g}(x_0,y_0) \\ &\quad \exp\left[-i\frac{2\pi z}{\lambda} - i\pi\frac{x^2+y^2}{\lambda z} - i\pi\frac{x_0^2+y_0^2}{\lambda z} + i2\pi\frac{xx_0+yy_0}{\lambda z}\right]\end{aligned}$$

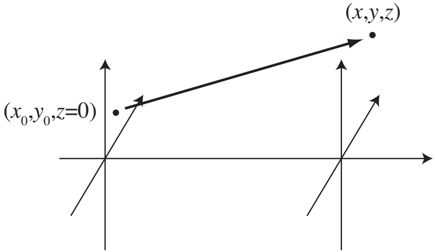
Look at the terms that depend on only x and y: they do not depend on the variables of integration, so we can pull them out of the integral, giving

$$\begin{aligned}\psi(x,y,z) &= \psi_0 \frac{\lambda}{z}\frac{1}{A}\exp\left[-i\frac{2\pi z}{\lambda}\right]\exp\left[-i\pi\frac{x^2+y^2}{\lambda z}\right] \\ &\quad \int_{x_0}\int_{y_0}\tilde{g}(x_0,y_0)\exp\left[-i\pi\frac{x_0^2+y_0^2}{\lambda z}\right]\exp\left[i2\pi\frac{xx_0+yy_0}{\lambda z}\right]\end{aligned} \tag{23}$$

This is the Fresnel-Kirchoff diffraction equation. The prefactor phase terms only matter if we talk about light amplitudes; if we're just using ψ to calculate $I = \psi^*\psi$, we can forget about those phase factors outside the integral.

The Fraunhofer approximation

- What if the object at the input plane is a relatively small pinhole, such that $(x_0^2 + y_0^2)/(\lambda z)$ never reaches a very large value?



- This condition can be written as

$$z \gg 4\frac{x_0^2 + y_0^2}{\lambda} \tag{24}$$

If we again use $x_0 = 0.4$ mm and $\lambda = 500$ nm, we require $z \gg 2$ mm. We might satisfy this condition in many cases! This condition is known as the *Fraunhofer approximation*.

- We'll also ignore the phase factor $\exp[-i2\pi z/\lambda]$, since it just tells us that the wavefield changes phase by 2π every time we travel a wavelength in distance.

Fraunhofer approximation II

In the Fraunhofer approximation of Eq. 24, and ignoring phase factors outside the integral because they will disappear when calculating $I = \psi^* \psi$, we can write the Fresnel-Kirchoff diffraction integral of Eq. 23 as

$$\psi(x, y, z) \simeq \psi_0 \frac{\lambda}{z} \frac{1}{A} \int_{x_0} \int_{y_0} \tilde{g}(x_0, y_0) \exp\left[i2\pi \left(\frac{xx_0}{\lambda z} + \frac{yy_0}{\lambda z} \right) \right] dx_0 \, dy_0 \quad (25)$$

Remember that this just results from adding up all the Huygens point sources with their magnitude and phase variations $\tilde{g}(x_0, y_0)$. The form of this equation should ring a bell, however!