Section 3: Electromagnetic Waves 1

EM waves in vacuum

In regions of space where there are no charges and currents, Maxwell equations read

$$\nabla \cdot \mathbf{E} = 0 \tag{3.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{3.3}$$

$$\nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{3.4}$$

They are a set of coupled, first order, partial differential equations for **E** and **B**. They can be decoupled by applying curl to eqs. (3.3) and (3.4):

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\nabla \times \frac{\partial \mathbf{B}}{\partial t} = -\frac{\partial}{\partial t}(\nabla \times \mathbf{B}) = -\mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{3.5}$$

$$\nabla \times (\nabla \times \mathbf{B}) = \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} = \nabla \times \left(\mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) = \mu_0 \varepsilon_0 \frac{\partial}{\partial t}(\nabla \times \mathbf{E}) = -\mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} \tag{3.6}$$

Since $\nabla \cdot \mathbf{E} = 0$ and $\nabla \cdot \mathbf{B} = 0$ we have

$$\nabla^2 \mathbf{E} = \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{3.7}$$

$$\nabla^2 \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} \tag{3.8}$$

We now have *separate* equations for **E** and **B**, but they are of *second* order; that's the price you pay for decoupling them. In vacuum, then, each Cartesian component of **E** and **B** satisfies the three-dimensional wave equation

$$\nabla^2 f = \mu_0 \varepsilon_0 \frac{\partial^2 f}{\partial t^2} \tag{3.9}$$

The solution of this equation is a wave. So Maxwell's equations imply that empty space supports the propagation of electromagnetic waves, traveling at a speed

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} = 3.00 \cdot 10^8 \, m/s \tag{3.10}$$

which happens to be precisely the velocity of light, *c*. The implication is astounding: light is an electromagnetic wave. Of course, this conclusion does not surprise anyone today, but imagine what a revelation it was in Maxwell's time! Remember how ε_0 and μ_0 came into the theory in the first place: they were constants in Coulomb's law and the Biot-Savart law, respectively. You measure them in experiments involving charged pith balls, batteries, and wires—experiments having nothing whatever to do with light. And yet, according to Maxwell's theory you can calculate *c* from these two numbers. Notice the crucial role played by Maxwell's contribution to Ampere's law; without it, the wave equation would not emerge, and there would be no electromagnetic theory of light.

Monochromatic plane waves

Since different frequencies in the visible range correspond to different *colors,* such waves are called *monochromatic*. Suppose that the waves are traveling in the *z* direction and have *no x or y dependence*; these are called *plane* waves, because the fields are uniform over every plane perpendicular to the direction of propagation. We are interested, then, in fields of the form

$$\mathbf{E}(z,t) = \mathbf{E}_0 e^{i(kz-\omega t)} \tag{3.11}$$

$$\mathbf{B}(z,t) = \mathbf{B}_0 e^{i(kz-\omega t)}\,, \tag{3.12}$$

where $\mathbf{E}_0$ and $\mathbf{B}_0$ are the (complex) amplitudes (the *physical* fields, of course, are the real parts of **E** and **B**). Substituting eqs.(3.11) and (3.12) to eqs. (3.7) and (3.8) respectively we find that

$$c = \frac{\omega}{k}\,. \tag{3.13}$$

Here, k is the wave number, which is related to the wavelength of the wave by the equation

$$\lambda = \frac{2\pi}{k}\,, \tag{3.14}$$

and ω is the angular frequency of EM wave.

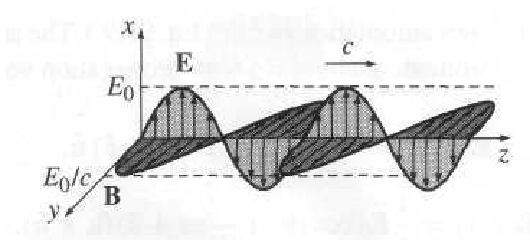

Fig. 1 This is the paradigm for a monochromatic plane wave. The wave is polarized in the *x* direction (by convention, we use the direction of **E** to specify the polarization of an electromagnetic wave).

Now, the wave equations for **E** and **B** were derived from Maxwell's equations. However, whereas every solution to Maxwell's equations (in empty space) must obey the wave equation, the converse is *not* true; Maxwell's equations impose extra constraints on $\mathbf{E}_0$ and $\mathbf{B}_0$. In particular, since $\nabla \cdot \mathbf{E} = 0$ and $\nabla \cdot \mathbf{B} = 0$, it follows that

$$E_{0z} = B_{0z} = 0 \tag{3.15}$$

That is, *electromagnetic waves are transverse:* the electric and magnetic fields are perpendicular to the direction of propagation. Moreover, Faraday's law, $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$, implies a relation between the electric and magnetic amplitudes:

$$\nabla \times \mathbf{E} = \left(\nabla \times \mathbf{E}_0\right) e^{i(kz-\omega t)} - \mathbf{E}_0 \times \nabla e^{i(kz-\omega t)} = -\mathbf{E}_0 \times \nabla e^{i(kz-\omega t)} = -\begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ E_{0x} & E_{0y} & 0 \\ 0 & 0 & ik \end{vmatrix} e^{i(kz-\omega t)} =$$
$$\left(-\hat{\mathbf{x}} ikE_{0y} + \hat{\mathbf{y}} ikE_{0x}\right) e^{i(kz-\omega t)} = \left(\hat{\mathbf{x}} i\omega B_{0x} + \hat{\mathbf{y}} i\omega B_{0y}\right) e^{i(kz-\omega t)} \tag{3.16}$$

which results in

$$-kE_{0y} = \omega B_{0x}, \quad kE_{0x} = \omega B_{0y}, \tag{3.17}$$

or, more compactly:

$$\mathbf{B}_0 = \frac{k}{\omega}(\hat{\mathbf{z}} \times \mathbf{E}_0) = \frac{1}{c}(\hat{\mathbf{z}} \times \mathbf{E}_0) \tag{3.18}$$

Evidently, **E** and **B** are *in phase* and *mutually perpendicular;* their (real) amplitudes are related by

$$B_0 = \frac{k}{\omega} E_0 = \frac{1}{c} E_0 \tag{3.19}$$

The fourth of Maxwell's equations, $\nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}$, does not yield an independent condition; it simply reproduces Eq. (3.16).

There is nothing special about the *z* direction, of course—we can easily generalize to monochromatic plane waves traveling in an arbitrary direction. The notation is facilitated by the introduction of the *wave vector*, **k**, pointing in the direction of propagation, whose magnitude is the wave number *k.* The scalar product $\mathbf{k} \cdot \mathbf{r}$ is the appropriate generalization *of kz*, so

$$\mathbf{E}(\mathbf{r}, t) = E_0 \mathbf{e} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}, \tag{3.20}$$

$$\mathbf{B}(\mathbf{r}, t) = \frac{1}{c} E_0 (\mathbf{n} \times \mathbf{e}) e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} = \frac{1}{c}(\mathbf{n} \times \mathbf{E}). \tag{3.21}$$

Where vector $\mathbf{n} = \frac{\mathbf{k}}{k}$ is the unit vector in the direction of propagation of the EM wave and **e** is the polarization vector. Because **E** is transverse,

$$\mathbf{n} \cdot \mathbf{e} = 0. \tag{3.22}$$

Linear and circular polarizations

The plane wave (3.20) and (3.21) is a wave with its electric field vector always in the direction **e**. Such a wave is said to be *linearly polarized* with polarization vector $\mathbf{e}_1 = \mathbf{e}$. Evidently there exists another wave which is linearly polarized with polarization vector $\mathbf{e}_2 \neq \mathbf{e}_1$ and is linearly independent of the first. Thus the two waves are

$$\begin{aligned} \mathbf{E}_1(\mathbf{r}, t) &= E_1 \mathbf{e}_1 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \\ \mathbf{E}_2(\mathbf{r}, t) &= E_2 \mathbf{e}_2 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \end{aligned} \tag{3.23}$$

with $\mathbf{B}_{1,2} = \frac{1}{c}(\mathbf{n} \times \mathbf{E}_{1,2})$.

They can be combined to give the most general homogeneous plane wave propagating in the direction $\mathbf{k} = k\mathbf{n}$:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_1(\mathbf{r}, t) + \mathbf{E}_2(\mathbf{r}, t) = (E_1 \mathbf{e}_1 + E_2 \mathbf{e}_2) e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}. \tag{3.24}$$

The amplitudes E_1 and E_2 are complex numbers, to allow the possibility of a phase difference between waves of different linear polarization.

If E_1 and E_2 have the *same phase* wave (3.24) represents a *linearly polarized wave,* with its polarization vector making an angle $\tan\theta = E_2 / E_1$ with respect to $\mathbf{e}_1$ and a magnitude $E = \sqrt{E_2^{\ 2} + E_1^{\ 2}}$ as shown in Fig.2a.

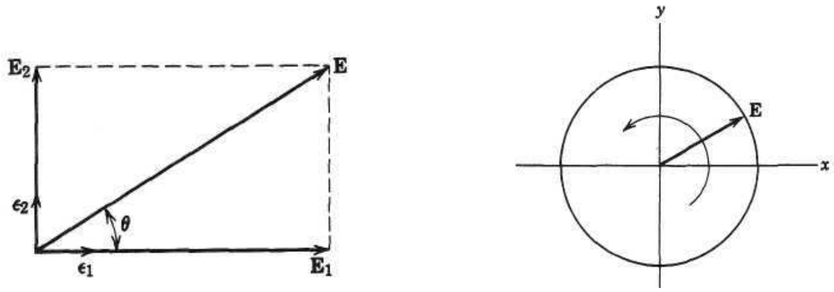

Fig. 2 **(a)** Electric field of a linearly polarized wave. **(b)** Electric field of a circularly polarized wave.

If E_1 and E_2 have *different phases,* the wave (3.24) is *elliptically polarized.* To understand what this means let us consider the simplest case, *circular polarization.* Then E_1 and E_2 have the same magnitude, but differ in phase by 90°. The wave (3.24) becomes:

$$\mathbf{E}(\mathbf{r},t) = E_0\left(\mathbf{e}_1 \pm i\mathbf{e}_2\right)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \tag{3.25}$$

with E_0 the common real amplitude. We imagine axes chosen so that the wave is propagating in the positive z direction, while $\mathbf{e}_1$ and $\mathbf{e}_2$ are in the x and y directions, respectively. Then the components of the actual electric field, obtained by taking the real part of (3.25), are

$$\begin{aligned} E_x(\mathbf{r},t) &= E_0\cos(kz-\omega t) \\ E_y(\mathbf{r},t) &= \mp E_0\sin(kz-\omega t) \end{aligned} \tag{3.26}$$

At a *fixed point in space,* the fields (3.26) are such that the electric vector is constant in magnitude, but sweeps around in a circle at a frequency ω, as shown in Fig.2b. For the upper sign $\left(\mathbf{e}_1 + i\mathbf{e}_2\right)$, the rotation is counterclockwise when the observer is facing into the oncoming wave. This wave is called *left circularly polarized* in optics. In the terminology of modern physics, however, such a wave is said to have *positive helicity.* The latter description seems more appropriate because such a wave has a positive projection of angular momentum on the z axis. For the lower sign $\left(\mathbf{e}_1 - i\mathbf{e}_2\right)$, the rotation of **E** is clockwise when looking into the wave; the wave is *right circularly polarized* (optics): it has *negative helicity.*

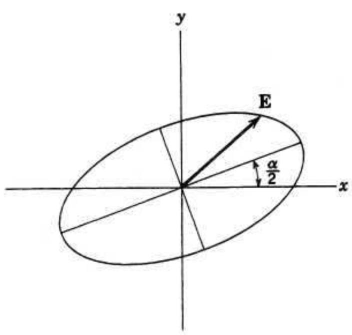

Figure 3. Electric field and magnetic induction for an elliptically polarized wave.

The two circularly polarized waves (3.26) form an equally acceptable set of basic fields for description of a general state of polarization. If we introduce the complex orthogonal unit vectors:

$$\mathbf{e}_\pm = \frac{1}{\sqrt{2}}\left(\mathbf{e}_1 \pm i\mathbf{e}_2\right), \tag{3.27}$$

then a general representation, equivalent to (3.24), is

$$\mathbf{E}(\mathbf{r},t) = \left(E_+\mathbf{e}_+ + E_-\mathbf{e}_-\right)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \tag{3.28}$$

where E_+ and E_- are complex amplitudes. If E_+ and E_- have different magnitudes, but the same phase, (3.28) represents an elliptically polarized wave with principal axes of the ellipse in the directions of $\mathbf{e}_1$ and $\mathbf{e}_2$. The ratio of semimajor to semiminor axis is $\left|(1+r)/(1-r)\right|$, where $r = E_-/E_+$. If the amplitudes have a phase difference between them, $E_-/E_+ = re^{i\alpha}$, then the ellipse traced out by the $\mathbf{E}$ vector has its axes rotated by an angle $\alpha/2$. Fig.3 shows the general case of elliptical polarization and the ellipse traced out by **E** at a given point in space. For $r = \pm 1$ we get back a linearly polarized wave.

Energy and momentum of electromagnetic waves

Energy per unit volume stored in electromagnetic fields is

$$u = \frac{1}{2}\left(\varepsilon_0 E^2 + \frac{1}{\mu_0}B^2\right). \tag{3.29}$$

Note that here the E and B are real quantities (real pars of complex quantities used in the previous paragraph). In case of monochromatic plane wave

$$B_0{}^2 = \frac{1}{c^2}E_0{}^2 = \mu_0\varepsilon_0 E_0{}^2, \tag{3.30}$$

so the magnetic and electric contribution are equal

$$u = \varepsilon_0 E^2 = \varepsilon_0 E_0{}^2\cos^2(kz-\omega t+\delta). \tag{3.31}$$

As the wave travels, it carriers this energy alone with it. The energy flux density (energy per unit area, per unit time) transported by the fields is given by the Poynting vector:

$$\mathbf{S} = \frac{1}{\mu_0}\left(\mathbf{E}\times\mathbf{B}\right). \tag{3.32}$$

For monochromatic plane waves propagating in the z direction,

$$\mathbf{S} = \frac{1}{\mu_0}\frac{E^2}{c}\hat{\mathbf{z}} = c\varepsilon_0 E^2\hat{\mathbf{z}} = c\varepsilon_0 E_0{}^2\cos^2(kz-\omega t+\delta)\hat{\mathbf{z}} = cu\hat{\mathbf{z}}. \tag{3.33}$$

Notice that **S** is the energy density (u) times the velocity of the waves (c) – as it *should be*. For in a time Δt, a length $c\Delta t$ passes through area A (Fig.4), carrying with it an energy $uAc\Delta t$. The energy per unit time, per unit area, transported by the wave is therefore uc.

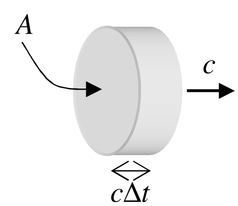

Fig.4

Electromagnetic fields not only carry *energy,* they also carry *momentum.* The momentum density stored in the fields is

$$\mathbf{p}_{em} = \frac{1}{c^2}\mathbf{S}. \tag{3.34}$$

For monochromatic plane waves, then,

$$\mathbf{p}_{em} = \frac{1}{c}\varepsilon_0 E_0{}^2\cos^2(kz-\omega t+\delta)\hat{\mathbf{z}} = \frac{1}{c}u\hat{\mathbf{z}}. \tag{3.35}$$

In the case of *light,* the wavelength is so short (~ 5 x 10^{-7} m), and the period so brief (~ 10^{-15} s), that any macroscopic measurement will encompass many cycles. Typically, therefore, we're not interested in the fluctuating cosine-squared term in the energy and momentum densities; all we want is the *average* value. Now, the average of cosine-squared over a complete cycle is ½ so

$$\langle u \rangle = \frac{1}{2}\varepsilon_0 E_0^2 . \tag{3.36}$$

$$\langle \mathbf{S} \rangle = \frac{1}{2}c\varepsilon_0 E_0{}^2 \hat{\mathbf{z}} . \tag{3.37}$$

$$\langle \mathbf{p}_{em} \rangle = \frac{1}{2c}\varepsilon_0 E_0{}^2 \hat{\mathbf{z}} . \tag{3.38}$$

We use brackets, <>, to denote the time average over a complete cycle. The average power per unit area transported by an electromagnetic wave is called the *intensity*:

$$I \equiv \langle S \rangle = \frac{1}{2}c\varepsilon_0 E_0{}^2 . \tag{3.39}$$

Electromagnetic waves in matter

In regions of matter where there are no *free* charges and *free* currents, Maxwell equations are

$$\nabla \cdot \mathbf{D} = 0 \tag{3.40}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3.41}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{3.42}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \tag{3.43}$$

If the matter is *linear*

$$\mathbf{D} = \varepsilon \mathbf{E}, \quad \mathbf{H} = \frac{1}{\mu}\mathbf{B} \tag{3.44}$$

and *homogeneous* (ε and μ are constants), Maxwell equations reduce to

$$\nabla \cdot \mathbf{E} = 0 \tag{3.45}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3.46}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{3.47}$$

$$\nabla \times \mathbf{B} = \varepsilon\mu \frac{\partial \mathbf{E}}{\partial t} , \tag{3.48}$$

which are different from Maxwell equations in vacuum only by the replacement of $\varepsilon_0\mu_0$ to $\varepsilon\mu$. Evidently electromagnetic waves propagate trough a linear homogeneous medium at a speed

$$\mathrm{v} = \frac{1}{\sqrt{\varepsilon\mu}} = \frac{c}{n} , \tag{3.49}$$

where

$$n \equiv \sqrt{\frac{\varepsilon\mu}{\varepsilon_0\mu_0}} , \tag{3.50}$$

is the *index of refraction*. For most (non-ferromagnetic) materials μ is very close to μ_0 so

$$n \cong \sqrt{\varepsilon_r}\ , \tag{3.51}$$

where $\varepsilon_r \equiv \varepsilon / \varepsilon_0$ is the dielectric constant. Since ε_r is always greater than 1, light travels more slowly through matter. In our consideration of EM waves below we assume for simplicity that $\mu = \mu_0$. In that case all the results we obtained above to EM waves in vacuum are valid with replacement ε_0 by ε.

Reflection and transmission of EM waves at normal incidence

If a wave passes from one transparent medium into another, there is a reflected wave and a transmitted wave. The details depend on the exact nature of the electrodynamic boundary conditions which are

$$\varepsilon_1 E_1^{\perp} = \varepsilon_2 E_2^{\perp} \tag{3.52}$$

$$B_1^{\perp} = B_2^{\perp} \tag{3.53}$$

$$\mathbf{E}_1^{\parallel} = \mathbf{E}_2^{\parallel} \tag{3.54}$$

$$\mathbf{B}_1^{\parallel} = \mathbf{B}_2^{\parallel} \tag{3.55}$$

Here we assumed that the two media which are characterized by indices 1 and 2 have different electric permeabilities $\varepsilon_1 \neq \varepsilon_2$ but the same magnetic permeabilities $\mu_1 = \mu_2 = \mu_0$. These equations relate the electric and magnetic fields just to the left and just to the right of the interface between two linear media. Now use them to deduce the laws governing reflection and refraction of electromagnetic waves at normal incidence

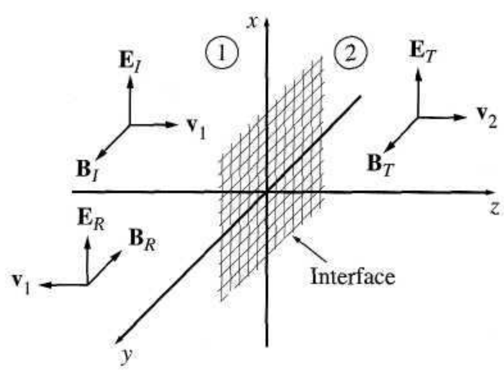

Fig.5

Suppose the *xy* plane forms the boundary between two linear media. A plane wave of frequency ω, traveling in the *z* direction and polarized in the *x* direction, approaches the interface from the left as is shown in Fig.5:

$$\mathbf{E}_i(z,t) = E_{0i} e^{i(k_1 z - \omega t)} \hat{\mathbf{x}} \tag{3.56}$$

$$\mathbf{B}_i(z,t) = \frac{1}{\mathrm{v}_1} E_{0i} e^{i(k_1 z - \omega t)} \hat{\mathbf{y}}\ . \tag{3.57}$$

It gives rise to the reflected wave

$$\mathbf{E}_r(z,t) = E_{0r} e^{i(-k_1 z - \omega t)} \hat{\mathbf{x}} \tag{3.58}$$

$$\mathbf{B}_r(z,t) = -\frac{1}{\mathrm{v}_1} E_{0r} e^{i(-k_1 z - \omega t)} \hat{\mathbf{y}}\ , \tag{3.59}$$

which travels back to the left in medium (1), and a transmitted wave

$$\mathbf{E}_t(z,t) = E_{0t} e^{i(k_2 z - \omega t)} \hat{\mathbf{x}} \tag{3.60}$$

$$\mathbf{B}_t(z,t) = \frac{1}{v_2} E_{0i} e^{i(k_2 z - \omega t)} \hat{\mathbf{y}} , \tag{3.61}$$

which continues on the right in medium (2). Note that the minus sign in $\mathbf{B}_r$ is required by Eq.(3.21).

At $z = 0$, the combined fields on the left, $\mathbf{E}_i + \mathbf{E}_r$ and $\mathbf{B}_i + \mathbf{B}_r$, must join the fields on the right, $\mathbf{E}_t$ and $\mathbf{B}_t$, in accordance with the boundary conditions. In this case there are no components perpendicular to the surface, so (3.52) and (3.53) are trivial. However, (3.54) and (3.55) require that

$$E_{0i} + E_{0r} = E_{0t} \tag{3.62}$$

$$\frac{1}{v_1}\left(E_{0i} - E_{0r}\right) = \frac{1}{v_2} E_{0t} \tag{3.63}$$

Eq. (3.63) can be rewritten as follows

$$E_{0i} - E_{0r} = \beta E_{0t} , \tag{3.64}$$

where

$$\beta = \frac{v_1}{v_2} = \frac{n_2}{n_1} . \tag{3.65}$$

Equations (3.62) and (3.64) are easily solved for the outgoing amplitudes, in terms of the incident amplitude:

$$E_{0r} = \frac{1-\beta}{1+\beta} E_{0i}, \quad E_{0t} = \frac{2}{1+\beta} E_{0i} \tag{3.66}$$

In terms of the indices of refraction this result is as follows

$$E_{0r} = \frac{n_1 - n_2}{n_1 + n_2} E_{0i}, \quad E_{0t} = \frac{2n_1}{n_1 + n_2} E_{0i} . \tag{3.67}$$

In order to calculate the fraction of energy which is transmitted and reflected we need to find the intensity (average power per unit area) which is given by eq.(3.39) with ε_0 replaced by ε :

$$I = \frac{1}{2} \varepsilon v E_0^{\ 2} . \tag{3.68}$$

The ratio of the reflected intensity to the incident intensity is therefore

$$R \equiv \frac{I_r}{I_i} = \left| \frac{E_{0r}}{E_{0i}} \right|^2 = \left| \frac{n_1 - n_2}{n_1 + n_2} \right|^2 , \tag{3.69}$$

Whereas the ration of the transmitted intensity to the incident intensity is

$$T \equiv \frac{I_t}{I_i} = \frac{\varepsilon_2 v_2}{\varepsilon_1 v_1} \left| \frac{E_{0r}}{E_{0i}} \right|^2 = \frac{4 n_1 n_2}{\left(n_1 + n_2\right)^2} , \tag{3.70}$$

R is called the *reflection coefficient* and T the *transmission coefficient*; they measure the fraction of the incident energy that is reflected and transmitted, respectively. Notice that

$$R + T = 1 , \tag{3.71}$$

as conservation of energy, of course, requires. For instance, when light passes from air ($n_1 = 1$) into glass ($n_2 = 1.5$), $R = 0.04$ and $T = 0.96$. Not surprisingly, most of the light is transmitted.

Reflection and Transmission at Oblique Incidence

Now we consider the more general case of *oblique* incidence, in which the incoming wave meets the boundary at an arbitrary angle θ_i (Fig.6). The normal incidence considered in the previous section is really just a special case of oblique incidence, with $\theta_i = 0$.

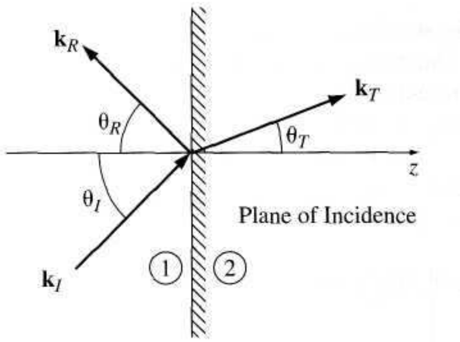

Fig.6

Suppose, then, that a monochromatic plane wave

$$\mathbf{E}_i(\mathbf{r},t) = \mathbf{E}_{0i} e^{i(\mathbf{k}_i\cdot\mathbf{r}-\omega t)}; \quad \mathbf{B}_i(\mathbf{r},t) = \frac{1}{\mathrm{v}_1}\left(\hat{\mathbf{k}}_i \times \mathbf{E}_i\right) \tag{3.72}$$

approaches from the left, giving rise to a reflected wave,

$$\mathbf{E}_r(\mathbf{r},t) = \mathbf{E}_{0r} e^{i(\mathbf{k}_r\cdot\mathbf{r}-\omega t)}; \quad \mathbf{B}_r(\mathbf{r},t) = \frac{1}{\mathrm{v}_1}\left(\hat{\mathbf{k}}_r \times \mathbf{E}_r\right) \tag{3.73}$$

and a transmitted wave

$$\mathbf{E}_t(\mathbf{r},t) = \mathbf{E}_{0t} e^{i(\mathbf{k}_t\cdot\mathbf{r}-\omega t)}; \quad \mathbf{B}_t(\mathbf{r},t) = \frac{1}{\mathrm{v}_2}\left(\hat{\mathbf{k}}_t \times \mathbf{E}_t\right). \tag{3.74}$$

All three waves have the same *frequency* ω that is determined once and for all at the source. The three wave numbers are related so that

$$k_i \mathrm{v}_1 = k_r \mathrm{v}_1 = k_t \mathrm{v}_2 = \omega \quad or \quad k_i = k_r = \frac{\mathrm{v}_2}{\mathrm{v}_1} k_t = \frac{n_1}{n_2} k_t\ . \tag{3.75}$$

The combined fields in medium (1), $\mathbf{E}_i + \mathbf{E}_r$ and $\mathbf{B}_i + \mathbf{B}_r$, must now be joined to the fields $\mathbf{E}_t$ and $\mathbf{B}_t$ in medium (2), using the boundary conditions (3.52) - (3.55). These all share the generic structure

$$(\ \ )e^{i(\mathbf{k}_i\cdot\mathbf{r}-\omega t)} + (\ \ )e^{i(\mathbf{k}_r\cdot\mathbf{r}-\omega t)} = (\ \ )e^{i(\mathbf{k}_t\cdot\mathbf{r}-\omega t)} \tag{3.76}$$

The important thing to notice is that the *x, y,* and *t* dependence is confined to the exponents. Because the boundary conditions must hold at all points on the plane, and for all times, these exponential factors must be equal at $z = 0$. The time factors are already equal. As for the spatial terms, evidently when $z = 0$

$$\mathbf{k}_i \cdot \mathbf{r} = \mathbf{k}_r \cdot \mathbf{r} = \mathbf{k}_t \cdot \mathbf{r} \tag{3.77}$$

or, more explicitly,

$$(k_i)_x x + (k_i)_y y = (k_r)_x x + (k_r)_y y = (k_t)_x x + (k_t)_y y \tag{3.78}$$

for all *x* and all *y*. Eq.(3.78) can *only* hold if the components are separately equal, for if $x = 0$, we get

$$(k_i)_y = (k_r)_y = (k_t)_y \tag{3.79}$$

while $y = 0$ gives

$$(k_i)_x = (k_r)_x = (k_t)_x \tag{3.80}$$

We can always choose axes in such a way that $\mathbf{k}_i$ lies in the xz plane $(k_i)_y = (k_r)_y = (k_t)_y = 0$. Eq. (3.79) leads to

First Law: The incident, reflected, and transmitted wave vectors form a plane (called the *plane of incidence*), which also includes the normal to the surface (here, the *z* axis).

Meanwhile, Eq.(3.80) implies that

$$k_i \sin\theta_i = k_r \sin\theta_r = k_t \sin\theta_t \tag{3.81}$$

where θ_i is the *angle of incidence*, θ_r is the *angle of reflection*, and θ_t is the angle of transmission, more commonly known as the *angle of refraction***,** all of them measured with respect to the normal (Fig. 6). In view of Eq. (3.81), then,

Second Law: The angle of incidence is equal to the angle of reflection,

$$\theta_i = \theta_r \tag{3.82}$$

This is the *law of reflection***.**

As for the transmitted angle, there is *Third Law*:

$$\frac{\sin\theta_t}{\sin\theta_i} = \frac{n_1}{n_2} \tag{3.83}$$

This is the *law of refraction* or *Snell's law*.

These are the three fundamental laws of geometrical optics.

Now that we have taken care of the exponential factors – they cancel, given Eq. (3.77) – the boundary conditions (3.52) – (3.55) become:

$$\varepsilon_1 \left(\mathbf{E}_{0i} + \mathbf{E}_{0r}\right)_z = \varepsilon_2 \left(\mathbf{E}_{0t}\right)_z \tag{3.84}$$

$$\left(\mathbf{B}_{0i} + \mathbf{B}_{0r}\right)_z = \left(\mathbf{B}_{0t}\right)_z \tag{3.85}$$

$$\left(\mathbf{E}_{0i} + \mathbf{E}_{0r}\right)_{x,y} = \left(\mathbf{E}_{0t}\right)_{x,y} \tag{3.86}$$

$$\left(\mathbf{B}_{0i} + \mathbf{B}_{0r}\right)_{x,y} = \left(\mathbf{B}_{0t}\right)_{x,y} \tag{3.87}$$

where $\mathbf{B}_0(\mathbf{r},t) = \frac{1}{\mathrm{v}}\left(\hat{\mathbf{k}} \times \mathbf{E}_0\right)$ in each case. The last two represent *pairs* of equations, one for the *x*-component and one for the *y*-component.

Suppose that the polarization of the incident wave is *parallel* to the plane of incidence (the *xz* plane in Fig. 7); then the reflected and transmitted waves are also polarized in this plane. Then Eq. (3.84) reads

$$\varepsilon_1 \left(-E_{0i} \sin\theta_i + E_{0r} \sin\theta_r\right) = \varepsilon_2 \left(-E_{0t} \sin\theta_t\right) \tag{3.88}$$

Eq. (3.85) adds nothing (0 = 0), since the magnetic fields have no z components; Eq. (3.86) becomes

$$E_{0i} \cos\theta_i + E_{0r} \cos\theta_r = E_{0t} \cos\theta_t ; \tag{3.89}$$

and Eq. (3.87) gives

$$\frac{1}{\mathrm{v}_1}\left(E_{0i} - E_{0r}\right) = \frac{1}{\mathrm{v}_2} E_{0t} ; \tag{3.90}$$

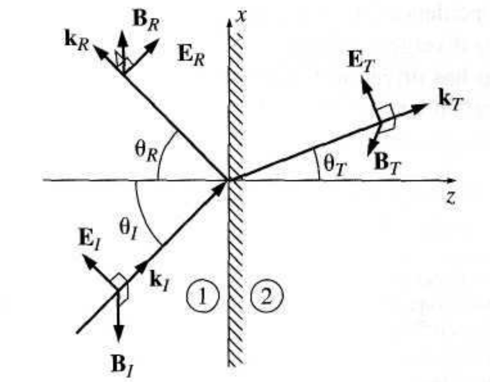

Fig. 7

Given the laws of reflection and refraction [(3.82) and (3.83)], Eqs. (3.88) and (3.90) reduce to

$$E_{0i} - E_{0r} = \beta E_{0t}\,, \tag{3.91}$$

where as before

$$\beta = \frac{v_1}{v_2} = \frac{n_2}{n_1}\,. \tag{3.92}$$

Eq. (3.89) says

$$E_{0i} + E_{0r} = \alpha E_{0t}\,, \tag{3.93}$$

where

$$\alpha = \frac{\cos\theta_t}{\cos\theta_i}\,. \tag{3.94}$$

Solving Eqs. (3.91) and (3.93) for the reflected and transmitted amplitudes, we obtain

$$E_{0r} = \frac{\alpha-\beta}{\alpha+\beta}E_{0i}; \quad E_{0t} = \frac{2}{\alpha+\beta}E_{0i}. \tag{3.95}$$

These are known as *Fresnel's equations*, for the case of polarization in the plane of incidence. There are two other Fresnel equations, giving the reflected and transmitted amplitudes when the polarization is *perpendicular* to the plane of incidence. These equations you are asked to derive at home. Notice that the transmitted wave is always *in phase* with the incident one; the reflected wave is either in phase ("right side up"), if $\alpha > \beta$, or 180° out of phase ("upside down"), if $\alpha < \beta$.

The amplitudes of the transmitted and reflected waves depend on the angle of incidence, because α is a function of θ_i:

$$\alpha = \frac{\sqrt{1-\sin^2\theta_t}}{\cos\theta_i} = \frac{\sqrt{1-\left[(n_1/n_2)\sin\theta_i\right]^2}}{\cos\theta_i}\,; \tag{3.96}$$

In the case of normal incidence (θ_i = 0), α = 1, and we recover Eq. (3.67). At grazing incidence (θ_i = 90°), α diverges, and the wave is totally reflected. Interestingly, there is an intermediate angle, θ_B (called *Brewster's angle*), at which the reflected wave is completely extinguished. According to Eq. (3.95), this occurs when $\alpha = \beta$ or

$$\sin^2\theta_B = \frac{\beta^2}{1+\beta^2}\,; \tag{3.97}$$

or equivalently

$$\tan\theta_B = \beta = \frac{n_2}{n_1}; \tag{3.98}$$

Figure 8 shows a plot of the transmitted and reflected amplitudes as functions of θ_i, for light incident on glass (n_2 = 1.5) from air (n_1 = 1). On the graph, a *negative* number indicates that the wave is 180° out of phase with the incident beam – the amplitude itself is the absolute value.

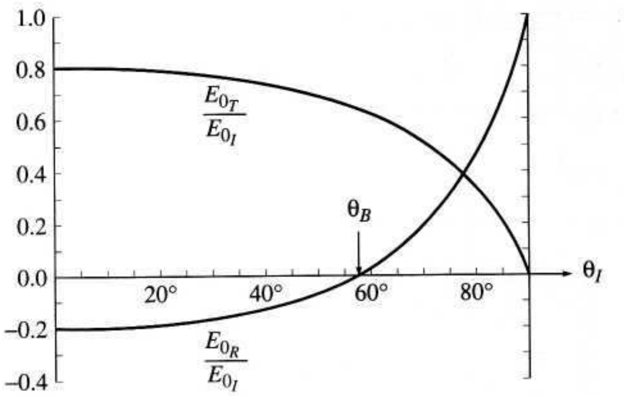

Fig.8

If the wave is polarized *perpendicular* to the plane of incidence there is no Brewster's angle for *any n*. Therefore, if a plane wave of mixed polarization is incident on a plane interface at the Brewster angle, the reflected radiation is *completely plane-polarized* with polarization vector *perpendicular* to the plane of incidence. This behavior can be utilized to produce beams of plane-polarized light but is not as efficient as other means employing anisotropic properties of some dielectric media. Even if the unpolarized wave is reflected at angles other than the Brewster angle, there is a tendency for the reflected wave to be predominantly polarized perpendicular to the plane of incidence.

The power per unit area striking the interface is $\mathbf{S}\cdot\hat{\mathbf{z}}$. Thus the incident intensity is

$$I_i = \frac{1}{2}\varepsilon_1 v_1 E_{0i}^2 \cos\theta_i, \tag{3.99}$$

while the reflected and transmitted intensities are

$$I_r = \frac{1}{2}\varepsilon_1 v_1 E_{0r}^2 \cos\theta_r, \tag{3.100}$$

$$I_t = \frac{1}{2}\varepsilon_2 v_2 E_{0t}^2 \cos\theta_t. \tag{3.101}$$

The cosines are there because the intensities represent the average power per unit area of *interface,* and the interface is at an angle to the wave front. The reflection and transmission coefficients for waves polarized parallel to the plane of incidence are

$$R \equiv \frac{I_r}{I_i} = \left|\frac{E_{0r}}{E_{0i}}\right|^2 = \left(\frac{\alpha-\beta}{\alpha+\beta}\right)^2 \tag{3.102}$$

$$T \equiv \frac{I_t}{I_i} = \frac{\varepsilon_2 v_2}{\varepsilon_1 v_1}\left|\frac{E_{0r}}{E_{0i}}\right|^2 \frac{\cos\theta_t}{\cos\theta_i} = \alpha\beta\left(\frac{2}{\alpha+\beta}\right)^2 \tag{3.103}$$

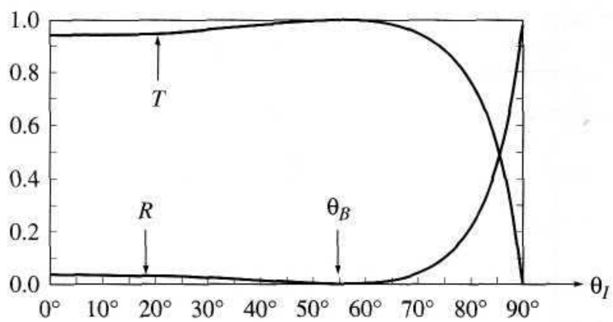

Fig.9

They are plotted as functions of the angle of incidence in Fig. 9 (for the air/glass interface). *R* is the fraction of the incident energy that is reflected – naturally, it goes to zero at Brewster's angle; *T* is the fraction transmitted – it goes to 1 at θ_B. Note that $R + T = 1$, as required by conservation of energy: the energy per unit time *reaching* a particular patch of area on the surface is equal to the energy per unit time *leaving* the patch.

There is another important phenomenon which is called *total internal reflection.* The word "internal" implies that the incident and reflected waves are in a medium of larger index of refraction than the refracted wave ($n_1 > n_2$). Snell's law (7.36) shows that, if $n_1 > n_2$, then $\theta_t > \theta_i$. Consequently, *there is a critical angle* when $\theta_i = \theta_i^0$ at which $\theta_t = \pi/2$,

$$\sin\theta_i^0 = \frac{n_2}{n_1}, \tag{3.104}$$

i.e., the refracted wave is propagated parallel to the surface. There can be no energy flow across the surface. Hence at that angle of incidence there must be total reflection. What happens if $\theta_i > \theta_i^0$? To answer this we first note that, for $\theta_i > \theta_i^0$,

$$\sin\theta_t = \frac{n_1}{n_2}\sin\theta_i = \frac{\sin\theta_i}{\sin\theta_i^0} > 1\,. \tag{3.105}$$

This means that θ_t is a complex angle with a purely imaginary cosine.

$$\cos\theta_t = i\sqrt{\left(\frac{\sin\theta_i}{\sin\theta_i^0}\right)^2 - 1} \tag{3.106}$$

The meaning of these complex quantities becomes clear when we consider the propagation factor for the refracted wave:

$$e^{i\mathbf{k}_t\cdot\mathbf{r}} = e^{ik_2(x\sin\theta_t + z\cos\theta_t)} = e^{ik_2 x\frac{\sin\theta_i}{\sin\theta_i^0}} e^{-k_2 z\sqrt{\left(\frac{\sin\theta_i}{\sin\theta_i^0}\right)^2 - 1}} \tag{3.107}$$

This shows that, for $\theta_i > \theta_i^0$, the refracted wave is propagated only parallel to the surface and is attenuated exponentially beyond the interface. The attenuation occurs within a very few wavelengths of the boundary, except for $\theta_i \cong \theta_i^0$.