

Excerpt from Frisvad, J. R. *Light, Matter, and Geometry: The Cornerstones of Appearance Modelling*. PhD thesis, Technical University of Denmark, May 2008.

## CHAPTER 4

# Electromagnetic Radiation

*Whether it be molecules, the waves of the sea or myriads of stars, elements of nature form overall structures.*

Peter Haarby describing Inge Lise Westman's paintings

---

In the previous chapter we found a number of correspondences between quantum fields and Maxwell's equations. In particular, we found that the electromagnetic field vectors capture the mean effects of quantum particles. The Maxwell equations that we found correspondences to are sometimes referred to as the microscopic Maxwell equations. These equations only involve the electric and magnetic field vectors ( $\mathbf{E}$  and  $\mathbf{B}$ ), and the current and charge densities ( $\mathbf{j}$  and  $\rho$ ). The charge and current densities are material specific quantities referring to the behaviour of the electrons in the material. The field vectors are a more macroscopic way of describing light.

What we are really interested in is the progress of energy in a field. Therefore we introduce the Poynting vector in this chapter (Sec. 4.1). The Poynting vector is a quantity describing the energy flow in an electrodynamic field. Based on this quantity, we try to say something about the propagation of electromagnetic energy. This leads to a justification for the inverse square law of radiation and a formal solution for the microscopic equations. Afterwards we move to a more macroscopic description of charges and currents (Sec. 4.2). This is necessary since it is difficult to model every atom in a graphics scene. More macroscopic material properties are introduced, and, with those, we obtain the so-called macroscopic Maxwell equations. The next step is to investigate the dif-

ferent wave functions that we can use as solutions for the macroscopic equations (Sec. 4.3). Using some simplifying assumptions, we arrive at plane waves as a simple solution. Finally, we describe some wave theory which is used extensively in graphics. In particular, we derive the law of reflection, the law of refraction, and the Fresnel equations for reflection and transmission (Sec. 4.4). As an additional feature we show that variants of these laws and formulae are also valid for the important case of inhomogeneous waves (almost any wave propagating in an absorbing material is inhomogeneous).

## 4.1 Microscopic Maxwell Equations

From the theory of quantum electrodynamics we understand that the interaction of photons and electrons to a certain extent agrees with Maxwell's equations:

$$c^2 \nabla \times \mathbf{B} = \frac{\mathbf{j}}{\varepsilon_0} + \frac{\partial \mathbf{E}}{\partial t} \quad (4.1)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (4.2)$$

$$\nabla \cdot \mathbf{E} = \rho / \varepsilon_0 \quad (4.3)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (4.4)$$

where  $\mathbf{E}$  and  $\mathbf{B}$  are the electric and magnetic field vectors,  $\mathbf{j}$  and  $\rho$  are the current and charge densities,  $c$  is the speed of light in vacuum, and  $\varepsilon_0$  is the vacuum permittivity. The original version of Maxwell's equations included two additional vectors ( $\mathbf{H}$  and  $\mathbf{D}$ ) because Maxwell (and other early workers) were not aware of the internal structure of atoms. They did not know that charges are bound to atoms and that atomic magnetism is caused by circulating currents [Feynman et al. 1964, Sec. 32-2]. This means that they had to work with a more general set of equations than what is really needed to describe nature. Maxwell's old magnetic vector  $\mathbf{H}$  proves to be useful in a more macroscopic description of the electromagnetic field. Therefore the field  $\mathbf{H}$  is still used extensively and we will introduce it later in this chapter.

In our thought experiment of the previous chapter concerning a quantum field simulator for rendering realistic images, we reasoned that the intensity of RGB colour values is determined by the number of photons arriving per second. The relative amount of red, green, and blue is determined by the energies of the photons. When we reduce the quantum particles to waves and use only Maxwell's equations, we have to think of it a little differently. The number of photons arriving per second is rather the magnitude of the energy flux in the field and the energies of the photons are the wavelengths present in the field. When we

look at Maxwell's equations there is no quantity describing the energy of the field. We need such a quantity in order to relate the electromagnetic field to the RGB colour values of our image.

Simple arguments show that the loss of energy per unit time and per unit volume due to work done by the electromagnetic field is the quantity  $\mathbf{E} \cdot \mathbf{j}$  [Feynman et al. 1964]. This was also known to Maxwell [1873, Vol. II, Chapter VI] and using his equation (4.1) involving the curl of the magnetic field, one can also write this quantity as

$$\mathbf{E} \cdot \mathbf{j} = \varepsilon_0 c^2 \mathbf{E} \cdot (\nabla \times \mathbf{B}) - \varepsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} .$$

Poynting [1884] essentially showed that another way to write it is

$$\mathbf{E} \cdot \mathbf{j} = -\nabla \cdot (\varepsilon_0 c^2 \mathbf{E} \times \mathbf{B}) - \frac{\partial}{\partial t} \left( \frac{\varepsilon_0}{2} \mathbf{E} \cdot \mathbf{E} + \frac{\varepsilon_0 c^2}{2} \mathbf{B} \cdot \mathbf{B} \right) .$$

Looking back at the equation (3.20) of charge conservation, this equation is remarkably similar. Only there is a loss of electromagnetic energy ( $-\mathbf{E} \cdot \mathbf{j}$ ) whereas there is no loss of charge in the field. Considering this analogy, one defines the energy flux  $\mathbf{S}$  of the field, also called Poynting's vector, and the energy density  $u$  of the field as follows:

$$\mathbf{S} = \varepsilon_0 c^2 \mathbf{E} \times \mathbf{B} \quad (4.5)$$

$$u = \frac{\varepsilon_0}{2} (|\mathbf{E}|^2 + c^2 |\mathbf{B}|^2) , \quad (4.6)$$

such that

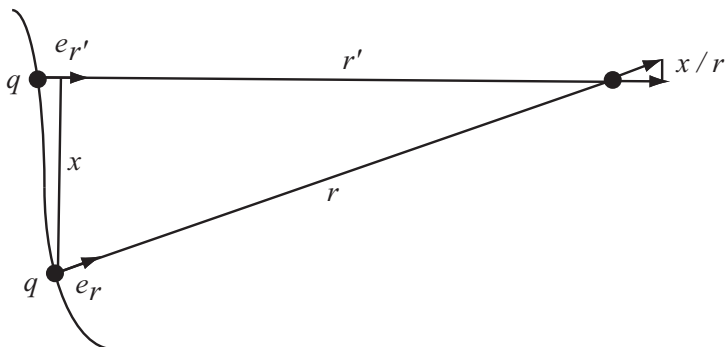
$$\nabla \cdot \mathbf{S} + \frac{\partial u}{\partial t} = -\mathbf{E} \cdot \mathbf{j} . \quad (4.7)$$

In the previous chapter we saw how the expression for the energy density (4.6) agrees well with the quantized description of the energy in a photon field (cf. Equations 3.7 and 3.17). Thus we can use the intensity of Poynting's vector  $|\mathbf{S}|$  to represent the intensity of the colour values in our renderer.

Knowing how to find the energy flux of the field, the next thing we need to know, is how to follow the propagation of the waves through a scene. How are waves of light emitted, how are they absorbed, how do they interact with matter? To start with emission, the formula for electromagnetic radiation by one individual point charge in free space is as follows [Feynman et al. 1963, Sec. 28-1]:

$$\mathbf{E} = \frac{q}{4\pi\varepsilon_0} \left( \frac{\vec{e}_{r'}}{r'^2} + \frac{r'}{c} \frac{d}{dt} \left( \frac{\vec{e}_{r'}}{r'^2} \right) + \frac{1}{c^2} \frac{d^2}{dt^2} \vec{e}_{r'} \right) \quad (4.8)$$

$$\mathbf{B} = \vec{e}_{r'} \times \mathbf{E}/c , \quad (4.9)$$



**Figure 4.1:** A moving point charge  $q$  at a large distance  $r \approx r'$  from an observer. This figure illustrates why the displacement of the vector  $\vec{e}_r(t)$  on the unit sphere is approximately equal to the displacement  $x(t)/r$ .

where  $q$  is the charge and  $\mathbf{r}(t) = r(t)\vec{e}_r(t)$  is the vector at time  $t$  from the charge toward the position we are considering in the field. The distance is denoted  $r(t)$  and the unit vector describing the direction of  $\mathbf{r}$  is denoted  $\vec{e}_r(t)$ . The distance and direction appearing in the formulae (4.8–4.9) are retarded such that  $r' = r(t - r'/c)$  and  $\vec{e}_{r'} = \vec{e}_r(t - r'/c)$ . Note that these expressions are recursive. We cannot determine the electromagnetic field at an instance in time without knowing where the charges were at some time in the past.

If we look closer at the equation for radiation from a single charge (4.8), it reveals that the first two terms will vanish quickly as the distance to the charge increases. So let us assume that  $r$  is large and concentrate on the third term only:

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0 c^2} \frac{d^2 \vec{e}_{r'}}{dt^2}.$$

Let us assume that the charge moves slowly compared to the speed of light. Then the charge will move only a short distance from  $\mathbf{r}'$  to  $\mathbf{r}$ . If we let  $x$  denote the length of  $\mathbf{r} - \mathbf{r}'$  projected on a plane normal to  $\mathbf{r}'$ , the sine of the angle between  $\mathbf{r}$  and  $\mathbf{r}'$ , that is,  $x/r$ , will approximately be the change in  $\vec{e}_{r'}$ . This is illustrated in Figure 4.1. Since the distance  $r \approx r'$  is large, it is almost constant. Therefore the acceleration of  $\vec{e}_{r'}$  is approximately  $a'_\perp/r'$ , where  $a'_\perp$  is the retarded, projected acceleration of the charge itself. The prime denotes that it is retarded, which means that it is at the time  $t - r'/c$ , and the symbol  $\perp$  denotes that it is projected on a plane normal to the viewing direction  $\mathbf{r}'$ . With this approximation, we have an expression for the magnitude of the electric field vector

$$|\mathbf{E}| = \frac{q}{4\pi\epsilon_0 c^2 r'} a'_\perp.$$

From the relation (4.9) between  $\mathbf{E}$  and  $\mathbf{B}$  we find the following (which is valid for energy radiated through a vacuum)

$$|\mathbf{B}| = |\mathbf{E}|/c , \quad (4.10)$$

Inserting in the expression for Poynting's vector (4.5), we find that the intensity of the energy flux is

$$|\mathbf{S}| = \varepsilon_0 c |\mathbf{E}|^2 . \quad (4.11)$$

If we assume that the charge is accelerating in a direction forming the angle  $\theta$  with  $\mathbf{r}'$ , we get the following intensity for a single charge far away from the point of observation:

$$|\mathbf{S}| = \frac{q^2}{16\pi^2 \varepsilon_0 c^3 r'^2} a'^2 \sin^2 \theta , \quad (4.12)$$

where  $a'$  is the retarded acceleration of the charge. Note that the intensity of the radiation falls off with the square of the distance  $r'$  to the source. This is the justification for the inverse square law of radiation and now we are aware of the simplifying assumptions involved in its derivation. If we get close enough or if the charge is moving fast enough, the inverse square law is no longer valid. In that case we should use the general result given by Equations 4.11 and 4.8 (except for  $r' = 0$ ).

The equation for a single point charge is interesting, but we need more than one charge to describe the scenes that we want to render in graphics. If we return to the vector potential  $\mathbf{A}$  and scalar potential  $\phi$  defined indirectly in the previous chapter (Equation 3.8), and choose the Lorentz gauge, we have

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} , \quad \mathbf{B} = \nabla \times \mathbf{A} , \quad \nabla \cdot \mathbf{A} = -\frac{1}{c^2} \frac{\partial\phi}{\partial t} . \quad (4.13)$$

Then Maxwell's equations (4.1–4.4) become

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\frac{\mathbf{j}}{\varepsilon_0 c^2} \quad (4.14)$$

$$\nabla^2 \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\varepsilon_0} . \quad (4.15)$$

These two equations are actually four differential equations of identical structure ( $\mathbf{A}$  has three components). Formally they have the same solution. This general solution for Maxwell's equations at the time  $t$  is [Feynman et al. 1964, Sec. 21-3]

$$\mathbf{A}(\mathbf{x}, t) = \int \frac{\mathbf{j}(\mathbf{y}, t - r_{\mathbf{x}\mathbf{y}}/c)}{4\pi\varepsilon_0 c^2 r_{\mathbf{x}\mathbf{y}}} d\mathbf{y} \quad (4.16)$$

$$\phi(\mathbf{x}, t) = \int \frac{\rho(\mathbf{y}, t - r_{\mathbf{x}\mathbf{y}}/c)}{4\pi\varepsilon_0 r_{\mathbf{x}\mathbf{y}}} d\mathbf{y} , \quad (4.17)$$

where  $\mathbf{x}$  and  $\mathbf{y}$  are positions in space. The field vectors are obtained by insertion in the equations (4.13) which define the potentials. It is possible to derive the field of a single point charge (4.8–4.9) using this solution [Feynman et al. 1964, Chapter 21].

## 4.2 Macroscopic Maxwell Equations

Since we do not want to simulate the interaction of every charge with the field, we need some more macroscopic measures. Considering the general solution (4.16–4.17), macroscopic expressions for the charge and current densities ( $\rho$  and  $\mathbf{j}$ ) seem to be the right way to go. Suppose we want to model a material with  $N$  atoms per unit volume. We model each atom as having just one general dipole moment  $q\mathbf{d}$ , where  $q$  is the magnitude of the charges in the atom and  $\mathbf{d}$  is a vector denoting their separation. A dipole is two charges separated by a very short distance, but under a few assumptions any assembly of point charges approximately has a dipole potential. The assumptions are that the charges should be (a) located in a small limited region, (b) neutral as a whole, and (c) observed at a large distance [Feynman et al. 1964, Sec. 6-5]. The dipole moment per unit volume is called the *polarisation vector* and is given by

$$\mathbf{P} = Nq\mathbf{d} .$$

In this dipole approximation the polarisation vector is proportional to the electric field vector  $\mathbf{E}$ , we write

$$\mathbf{P} = \varepsilon_0\chi_e\mathbf{E} ,$$

and refer to  $\chi_e$  as the *electric susceptibility*. The charge and current densities due to the polarisation of a material are [Feynman et al. 1964, Sec. 10-3 and Sec. 32-2]

$$\rho_{\text{pol}} = -\nabla \cdot \mathbf{P} \quad , \quad \mathbf{j}_{\text{pol}} = \frac{d\mathbf{P}}{dt} . \quad (4.18)$$

If there are no charges or currents in free space and no magnetisation currents in the material, these charge and current densities are the only ones present. The polarisation vector is therefore a more macroscopic or phenomenological way of describing the charges and currents in a dielectric.

Magnetisation is not related to the charge density, “the magnetisation of materials comes from circulating currents within the atoms” [Feynman et al. 1964, Sec. 36-1]. To describe this, one introduces another macroscopic quantity: The *magnetisation vector*  $\mathbf{M}$ . It is defined indirectly by

$$\mathbf{j}_{\text{mag}} = \nabla \times \mathbf{M} . \quad (4.19)$$

Summing up the different terms in the charge and current densities, we have

$$\begin{aligned}\rho &= \rho_{\text{free}} + \rho_{\text{pol}} \\ \mathbf{j} &= \mathbf{j}_{\text{free}} + \mathbf{j}_{\text{pol}} + \mathbf{j}_{\text{mag}} .\end{aligned}$$

Inserting these in Maxwell's equations and moving the terms around, we get the following result:

$$\nabla \times (\varepsilon_0 c^2 \mathbf{B} - \mathbf{M}) = \mathbf{j}_{\text{free}} + \frac{\partial}{\partial t} (\varepsilon_0 \mathbf{E} + \mathbf{P}) \quad (4.20)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (4.21)$$

$$\nabla \cdot (\varepsilon_0 \mathbf{E} + \mathbf{P}) = \rho_{\text{free}} \quad (4.22)$$

$$\nabla \cdot \mathbf{B} = 0 . \quad (4.23)$$

If we introduce two additional field vectors:

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \quad (4.24)$$

$$\mathbf{H} = \varepsilon_0 c^2 \mathbf{B} - \mathbf{M} , \quad (4.25)$$

we get the original Maxwell equations. Sometimes this version of the equations is referred to as the macroscopic Maxwell equations because they involve the phenomenological polarisation and magnetisation vectors ( $\mathbf{P}$  and  $\mathbf{M}$ ).

The magnetisation vector is often assumed to be proportional to  $\mathbf{H}$  such that

$$\mathbf{M} = \chi_m \mathbf{H} , \quad (4.26)$$

where  $\chi_m$  is the *magnetic susceptibility*. Then a rearrangement of Equation 4.25 gives

$$\mathbf{H} = \frac{\varepsilon_0 c^2}{1 + \chi_m} \mathbf{B} .$$

This assumption is, however, only valid for a very limited set of magnetic materials. Another assumption we can make is that free charges only appear as conduction in a material. It has been found experimentally that metals produce a current with a density  $\mathbf{j}$  proportional to  $\mathbf{E}$  [Feynman et al. 1964, Sec. 32-6]. To model this relationship, we introduce another phenomenological quantity called the *conductivity*  $\sigma$ , such that

$$\mathbf{j}_{\text{free}} = \sigma \mathbf{E} .$$

To summarise these macroscopic or phenomenological material properties, we have

$$\mathbf{D} = \varepsilon_0 (1 + \chi_e) \mathbf{E} = \varepsilon \mathbf{E} \quad (4.27)$$

$$\mathbf{B} = (1 + \chi_m) / (\varepsilon_0 c^2) \mathbf{H} = \mu \mathbf{H} \quad (4.28)$$

$$\mathbf{j}_{\text{free}} = \sigma \mathbf{E} . \quad (4.29)$$

To shorten notation, these equations also introduce the permittivity  $\varepsilon$  and the permeability  $\mu$ . Note that all these macroscopic material properties are independent of the direction of the field. For this reason they are often referred to as the *isotropic* material equations. When using these material equations, we cannot model as general a case as if we use Equations 4.20–4.23. If we use the indirect definitions of  $\mathbf{P}$  and  $\mathbf{M}$  (Equations 4.18 and 4.19), we do not have to make any simplifying assumptions about the material. Unfortunately the polarisation and magnetisation vectors are not easy to model, so in the following we will use the isotropic material equations.

Let us briefly get an idea about how the isotropic material properties relate to real-world materials [Born and Wolf 1999]: If  $\sigma$  is not negligibly small, the material is a *conductor* (which roughly means that it has some electrons that are not bound to any particular atom such that they are able to produce a “free” current). As an example metals are good conductors. If the material is not a conductor, it is called a *dielectric*. The electric properties are then determined solely by the permittivity  $\varepsilon$ . If  $\mu$  differs appreciably from unity, the material is *magnetic*. In particular, if  $\mu > 1$ , the material is *paramagnetic*, while if  $\mu < 1$  it is *diamagnetic*. The material properties are all wavelength dependent.

### 4.3 Time-Harmonic Solution and Plane Waves

If we look at the differential equations (4.14–4.15) which give rise to the general solution for the electromagnetic field, they reveal that the vector and scalar potentials have wave solutions at locations in space where there is no charge or current density (where  $\rho$  and  $\mathbf{j}$  are zero). Let us represent the solution as time-harmonic plane waves. Then the potentials have the form

$$\begin{aligned}\mathbf{A}(\mathbf{x}, t) &= \operatorname{Re} \left( \mathbf{A}_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{x})} \right) \\ \phi(\mathbf{x}, t) &= \operatorname{Re} \left( \phi_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{x})} \right) ,\end{aligned}$$

where  $\mathbf{k}$  is the *wave vector* and  $\operatorname{Re}$  takes the real part of a complex quantity. If we insert these solutions in the expressions (4.13) for  $\mathbf{E}$  and  $\mathbf{B}$ , we again get time-harmonic plane waves. This plane wave solution is valid for radiation in free space, that is, when  $\rho$  and  $\mathbf{j}$  are zero. Let us try to figure out if it is also valid in a more general case.

Having stepped away from the quantum theories, we can safely express the field vectors in terms of Fourier transforms (the assumption is that the waves have a



continuous range of frequencies):

$$\mathbf{F}(\mathbf{x}, t) = \frac{1}{\pi} \int_0^\infty \mathcal{F}(\mathbf{x}, \omega) e^{-i\omega t} d\omega .$$

This is the argument why we are allowed to represent the field vectors as a superposition of time-harmonic functions:

$$\mathbf{F}(\mathbf{x}, t) = \text{Re}(\mathbf{F}_0(\mathbf{x})e^{-i\omega t}) .$$

In this representation both  $\mathbf{D}$  and  $\mathbf{E}$  are functions of  $e^{-i\omega t}$ . From the material equation (4.27) we then conclude that the permittivity  $\varepsilon$  does not depend on time. Similarly the permeability  $\mu$  does not depend on time.

The time-harmonic representation of the electromagnetic field is very convenient. Therefore let us write Maxwell's equations using complex time-harmonic vector functions. To make it clear that the field vectors are complex, we follow the notation of Bohren and Huffman [1983] and use the subscript  $c$ . As an example the time-harmonic representation of the electric field vector is

$$\mathbf{E}_c = \mathbf{E}_0(\mathbf{x})e^{-i\omega t} ,$$

where it is understood that we obtain the physical electric field vector by taking the real part  $\mathbf{E} = \text{Re}(\mathbf{E}_c)$ . With a loss of generality that is of no significance in a graphics context, we neglect charges moving freely through empty space, that is, we set  $\rho_{\text{free}} = 0$ . Using the isotropic material equations (4.27–4.29) and the time independence of  $\varepsilon$  and  $\mu$ , we get the following time-harmonic version of the macroscopic Maxwell equations:

$$\nabla \times \mathbf{H}_c = (\sigma - i\omega\varepsilon)\mathbf{E}_c \quad (4.30)$$

$$\nabla \times \mathbf{E}_c = i\omega\mu\mathbf{H}_c \quad (4.31)$$

$$\nabla \cdot (\varepsilon\mathbf{E}_c) = 0 \quad (4.32)$$

$$\nabla \cdot (\mu\mathbf{H}_c) = 0 . \quad (4.33)$$

Note that we have packed most of the important charges and currents into the  $\mathbf{H}_c$  vector and the material properties. By insertion of the plane wave expressions

$$\mathbf{E}_c(\mathbf{x}, t) = \mathbf{E}_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{x})} , \quad \mathbf{H}_c(\mathbf{x}, t) = \mathbf{H}_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{x})} , \quad (4.34)$$

we observe that plane waves do not in general satisfy the conditions. But they do satisfy them if we assume that the material properties are independent of position, that is, if the material is *homogeneous*. Thus plane waves are not only a solution in the free electromagnetic field, but also when we use the isotropic

material equations and assume that the material is homogeneous. Inserting the plane wave solution, we get the following Maxwell equations:

$$\mathbf{k} \times \mathbf{H}_0 = -\omega(\varepsilon + i\sigma/\omega)\mathbf{E}_0 \quad (4.35)$$

$$\mathbf{k} \times \mathbf{E}_0 = \omega\mu\mathbf{H}_0 \quad (4.36)$$

$$\mathbf{k} \cdot \mathbf{E}_0 = 0 \quad (4.37)$$

$$\mathbf{k} \cdot \mathbf{H}_0 = 0 \quad (4.38)$$

This reveals that Maxwell's equations require plane waves to satisfy the following conditions:

$$\mathbf{k} \cdot \mathbf{E}_0 = \mathbf{k} \cdot \mathbf{H}_0 = \mathbf{E}_0 \cdot \mathbf{H}_0 = 0 \quad (4.39)$$

$$\mathbf{k} \cdot \mathbf{k} = \omega^2\mu(\varepsilon + i\sigma/\omega) \quad (4.40)$$

where all the vectors may be complex and  $\varepsilon_c = \varepsilon + i\sigma/\omega$  is sometimes called the complex permittivity (or the complex dielectric constant). The latter equation is particularly interesting, it denotes the relation between material and wave propagation.

Let us take a look at the real and imaginary parts of the wave vector  $\mathbf{k}$ . We write

$$\mathbf{k} = \mathbf{k}' + i\mathbf{k}'' = k'\bar{\mathbf{s}}' + ik''\bar{\mathbf{s}}'' \quad ,$$

where  $k' = |\mathbf{k}'|$  and  $k'' = |\mathbf{k}''|$  such that  $\bar{\mathbf{s}}'$  and  $\bar{\mathbf{s}}''$  are unit vectors in the direction of real and imaginary part of the wave vector respectively. If the real part of the wave vector  $\mathbf{k}'$  is parallel to the imaginary part  $\mathbf{k}''$ , the wave is said to be *homogeneous*. Otherwise it is *inhomogeneous*. Of course,  $\mathbf{k}'' = \mathbf{0}$  is parallel to any vector, why a wave is homogeneous if  $\mathbf{k}$  is real-valued. If  $\mathbf{k}$  is complex, the exponential term of the plane wave expressions (4.34) is as follows

$$e^{i\mathbf{k} \cdot \mathbf{x}} = e^{i\mathbf{k}' \cdot \mathbf{x}} e^{-\mathbf{k}'' \cdot \mathbf{x}} \quad .$$

Here we may observe that  $\mathbf{k}'$  is the vector normal to the surface of constant phase and  $\mathbf{k}''$  is normal to the surface of constant amplitude. The phase velocity is then  $v = \omega/k'$  and the amplitude is damped (or decays) in the direction  $\bar{\mathbf{s}}''$  at the rate  $k''$ .

If we consider the relation (4.40) describing the rule for propagation of plane waves in homogeneous matter, it is obvious that a phenomenological quantity with the following definition is convenient:

$$n = n' + in'' = c\sqrt{\mu(\varepsilon + i\sigma/\omega)} \quad (4.41)$$

It is called the (complex) index of refraction, or refractive index. If we insert it in Equation 4.40, we get

$$\mathbf{k} \cdot \mathbf{k} = \mathbf{k}' \cdot \mathbf{k}' - \mathbf{k}'' \cdot \mathbf{k}'' + i2\mathbf{k}' \cdot \mathbf{k}'' = \frac{\omega^2}{c^2}n^2 \quad (4.42)$$

For materials that are not strong absorbers,  $\mathbf{k}'' \cdot \mathbf{k}'$  will be so small that we can neglect it. Then if we equate the real parts (and assume that the index of refraction is positive), we get

$$k' \approx \frac{\omega}{c} n' .$$

Hence, the real part of the refractive index is nearly the ratio of the speed of light in vacuum to the phase velocity  $n' \approx k' c / \omega = c/v$ . Equating the imaginary parts, has the result

$$2k' k'' \cos \theta = \frac{\omega^2}{c^2} 2n' n'' ,$$

where  $\theta$  is the angle between the real and imaginary parts of the wave vector ( $\mathbf{k}'$  and  $\mathbf{k}''$ ). Using the approximate expression for  $k'$ , we have

$$k'' \approx \frac{\omega}{c} \frac{n''}{\cos \theta} = \frac{2\pi}{\lambda} \frac{n''}{\cos \theta} .$$

where  $\lambda$  is the wavelength *in vacuum*. This means that the imaginary part  $n''$  is an expression related to the absorption of light in the material.

The index of refraction is a nice way to sum up all the material properties, and indeed it is a quantity which is measured as one of the key optical properties of materials. We will return to the optical properties of materials in Part II.

Considering the energy of the field, we know that it is the magnitude of the Poynting vector:

$$|\mathbf{S}| = |\varepsilon_0 c^2 \mathbf{E} \times \mathbf{B}| = \varepsilon_0 c^2 |\mathbf{E} \times \mathbf{B}| .$$

For the plane wave solution, we have

$$|\mathbf{S}| = \varepsilon_0 c^2 \mu |\operatorname{Re}(\mathbf{E}_c) \times \operatorname{Re}(\mathbf{H}_c)| ,$$

where we have used one of the isotropic material equations (4.28) and the plane wave expressions (4.34). When the plane wave expressions are inserted, we get

$$|\mathbf{S}| = \varepsilon_0 c^2 \mu \left| \operatorname{Re}(\mathbf{E}_0 e^{-i(\omega t - \mathbf{k}' \cdot \mathbf{x})}) \times \operatorname{Re}(\mathbf{H}_0 e^{-i(\omega t - \mathbf{k}' \cdot \mathbf{x})}) \right| e^{-2k'' \mathbf{s}'' \cdot \mathbf{x}} .$$

Since the exponential terms which involve  $\omega t$  and  $\mathbf{k}' \cdot \mathbf{x}$  are only oscillations, it follows that  $2k''$  is the exponential attenuation of the energy flux as the wave propagates through the material. This attenuation is called the *absorption coefficient* and in graphics we use the symbol  $\sigma_a$  (which should not be confused with the conductivity  $\sigma$ ). The relationship between the imaginary part of the refractive index and the absorption coefficient is

$$\sigma_a = 2k'' \approx \frac{4\pi n''}{\lambda \cos \theta} , \quad (4.43)$$

where  $\cos \theta = 1$  for homogeneous plane waves. After explaining all these quantities, we have come from a description of radiation from point charges at a microscopic level to a description of absorption of plane waves at a macroscopic level. We have not yet discussed how the wave changes when it meets a surface. This is the subject of the following section.

## 4.4 Reflection and Refraction

Let us consider a plane wave incident on a smooth surface. Due to the photon spin discussed in Chapter 3, it is convenient to resolve all the waves we deal with into two independent plane wave components. The wave components we choose are the wave with the electric vector perpendicular to the plane of incidence,  $\perp$ -polarised light, and the wave with the electric vector parallel to the plane of incidence,  $\parallel$ -polarised light. From experience we know that light incident on a smooth surface gives rise to two waves: A reflected and a transmitted wave. In the following we denote the incident wave by the subscript  $i$ , the reflected by the subscript  $r$ , and the transmitted by the subscript  $t$ . The boundary conditions given by Maxwell's equations require that the tangential component of the electric vector is continuous across the boundary of the surface. The  $\perp$ -polarised component of the electric vector is clearly tangent to the surface at the point of incidence, therefore at the boundary:

$$\mathbf{E}_{\perp i} + \mathbf{E}_{\perp r} = \mathbf{E}_{\perp t} .$$

This must hold at all times and no matter where we place the point of incidence in space. Suppose we place the point of incidence at the origin of our coordinate system (where  $\mathbf{x} = \mathbf{0}$ ), then

$$\mathbf{E}_{0i}^{\perp} e^{-i\omega_i t} + \mathbf{E}_{0r}^{\perp} e^{-i\omega_r t} = \mathbf{E}_{0t}^{\perp} e^{-i\omega_t t} .$$

This is true only if

$$\omega_i = \omega_r = \omega_t . \quad (4.44)$$

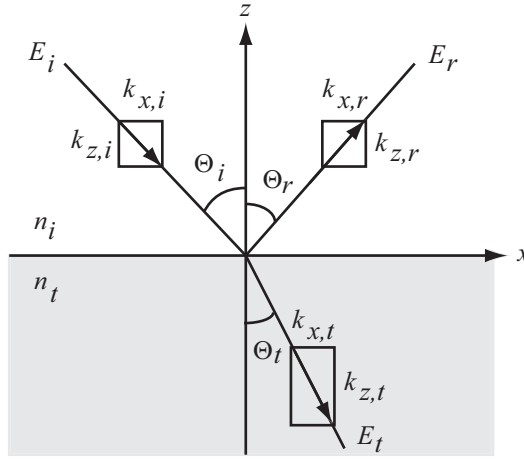
Then the exponential factors cancel out, and we have:

$$\mathbf{E}_{0i}^{\perp} + \mathbf{E}_{0r}^{\perp} = \mathbf{E}_{0t}^{\perp} . \quad (4.45)$$

In addition, since the frequency of the reflected and transmitted waves is the same as that of the incident wave (4.44), Equation 4.42 gives

$$\frac{\mathbf{k}_i \cdot \mathbf{k}_i}{n_i^2} = \frac{\mathbf{k}_r \cdot \mathbf{k}_r}{n_r^2} = \frac{\mathbf{k}_t \cdot \mathbf{k}_t}{n_t^2} . \quad (4.46)$$

In a sense this shows how the relation  $\mathbf{k} \cdot \mathbf{k} = k_0^2 n^2$  governs the propagation of a plane wave in homogeneous matter ( $k_0 = \omega/c$  is the wave number in vacuum).



**Figure 4.2:** A plane wave reflected and refracted on a surface  $z = 0$  with normal in the direction of the  $z$ -axis. The  $xz$ -plane is the plane of incidence.

Let us orient our coordinate system such that the tangent plane is the  $xy$ -plane and the plane of incidence is the  $xz$ -plane. Confer Figure 4.2. Then at the boundary, where  $z = 0$ , we have at the time  $t = 0$

$$\mathbf{E}_{0i}^\perp e^{i(k_{x,i}x + k_{y,i}y)} + \mathbf{E}_{0r}^\perp e^{i(k_{x,r}x + k_{y,r}y)} = \mathbf{E}_{0t}^\perp e^{i(k_{x,t}x + k_{y,t}y)} ,$$

This must hold for all  $x$  and  $y$  on the boundary, thus

$$k_{x,i} = k_{x,r} = k_{x,t} \quad \text{and} \quad k_{y,i} = k_{y,r} = k_{y,t} . \quad (4.47)$$

Since  $\mathbf{k}_i$  and the normal to the surface at the point of incidence span the plane of incidence,  $\mathbf{k}_i$  is parallel to it and therefore has no  $y$  component, that is,  $k_{y,i} = 0$ . Then according to our new result (4.47), we have

$$k_{y,i} = k_{y,r} = k_{y,t} = 0 . \quad (4.48)$$

In other words the reflected and transmitted waves lie in the plane of incidence. Using Equation 4.46, we get

$$k_{x,i}^2 + k_{z,i}^2 = k_{x,r}^2 + k_{z,r}^2 ,$$

where the  $x$  components cancel out (cf. Equation 4.47), and the mathematical solution is  $k_{z,r} = k_{z,i}$  or  $k_{z,r} = -k_{z,i}$ . The reflected wave was, however, defined to be propagating in the same medium as the incident wave, and if  $k_{z,r}$  has the same sign as  $k_{z,i}$ , the reflected wave is moving across the boundary. Therefore the only solution that makes physical sense is

$$k_{z,r} = -k_{z,i} . \quad (4.49)$$

Equations 4.47, 4.48, and 4.49 summarise the law of reflection for plane waves: *The reflected wave lies in the plane of incidence, the (complex) angle of reflection is equal to the (complex) angle of incidence.*

The law of refraction is found in a similar way. From Equations 4.46 and 4.48 we have

$$\frac{\mathbf{k}_t \cdot \mathbf{k}_t}{n_t^2} = \frac{\mathbf{k}_i \cdot \mathbf{k}_i}{n_i^2} .$$

Dividing by  $k_{x,t}^2 = k_{x,i}^2$  (the equality is from Equation 4.47) and juggling the terms around, the result is

$$n_i \sin \Theta_i = n_t \sin \Theta_t , \quad (4.50)$$

where  $\sin \Theta_i = k_{x,i}/\sqrt{\mathbf{k}_i \cdot \mathbf{k}_i}$  is sine of the complex angle of incidence and  $\sin \Theta_t = k_{x,t}/\sqrt{\mathbf{k}_t \cdot \mathbf{k}_t}$  is sine of the complex angle of refraction. We refer to this result as the generalised Snell's law . The law of refraction for plane waves is thus: *The refracted wave lies in the plane of incidence, the (complex) angle of refraction follows the generalised Snell's law (4.50).*

From the plane wave Maxwell equations (4.36) we have:

$$\mathbf{H}_0 = \frac{1}{\omega \mu} \mathbf{k} \times \mathbf{E}_0 .$$

Considering the  $x$  component of  $\mathbf{H}_0$ , and seeing that  $\mathbf{H}$  is also required to be continuous across the boundary, we have

$$-\frac{k_{z,i}}{\omega_i \mu_i} \mathbf{E}_{0i}^\perp - \frac{k_{z,r}}{\omega_r \mu_i} \mathbf{E}_{0r}^\perp = -\frac{k_{z,t}}{\omega_t \mu_t} \mathbf{E}_{0t}^\perp .$$

Recalling that the frequencies are equal (cf. Equation 4.44) and using the law of reflection  $k_{z,r} = -k_{z,i}$ , another way to write this is

$$k_{z,i} \mathbf{E}_{0i}^\perp - k_{z,i} \mathbf{E}_{0r}^\perp = k_{z,t} \frac{\mu_i}{\mu_t} \mathbf{E}_{0t}^\perp .$$

If we neglect the relative difference in permeability between the materials and insert our earlier result (4.45), this equation leads to

$$\mathbf{E}_{0r}^\perp = \frac{k_{z,i} - k_{z,t}}{k_{z,i} + k_{z,t}} \mathbf{E}_{0i}^\perp .$$

Dividing through by  $\sqrt{\mathbf{k}_i \cdot \mathbf{k}_i} \mathbf{E}_{0i}^\perp$  and again using Equation 4.46, we obtain the Fresnel equation for the reflection of the  $\perp$ -polarised component of the light:

$$\tilde{r}_\perp = \frac{\mathbf{E}_{0r}^\perp}{\mathbf{E}_{0i}^\perp} = \frac{n_i \cos \Theta_i - n_t \cos \Theta_t}{n_i \cos \Theta_i + n_t \cos \Theta_t} , \quad (4.51)$$

where  $\cos \Theta_i = k_{z,i}/\sqrt{\mathbf{k}_i \cdot \mathbf{k}_i}$  is cosine of the complex angle of incidence and  $\cos \Theta_t = k_{z,t}/\sqrt{\mathbf{k}_t \cdot \mathbf{k}_t}$  is cosine of the complex angle of refraction. These complex angles that we have now mentioned a few times, do not have the same simple geometrical interpretation as real angles. Nevertheless, they are still useful because cosine of a complex angle is the dot product of two normalised complex vectors.

The  $\parallel$ -polarised component is obtained in a similar fashion, but in this case the electric field vector is no longer parallel to the tangential plane. To use the condition which says that the tangential component of the electric vector is continuous across the boundary, we have to project the  $\parallel$ -polarised component to the tangential plane using cosine of the complex angles, such that

$$\mathbf{E}_{\parallel i} \cos \Theta_i + \mathbf{E}_{\parallel r} \cos \Theta_r = \mathbf{E}_{\parallel t} \cos \Theta_t .$$

With this relation, and again requiring that  $\mathbf{H}_0$  is continuous across the boundary, we obtain

$$\tilde{r}_{\parallel} = \frac{\mathbf{E}_{0r}^{\parallel}}{\mathbf{E}_{0i}^{\parallel}} = \frac{n_t \cos \Theta_i - n_i \cos \Theta_t}{n_t \cos \Theta_i + n_i \cos \Theta_t} . \quad (4.52)$$

The Fresnel equations (4.51,4.52) describe amplitude ratios, but often we are only interested in the flow of energy. To translate the amplitude ratios into energy ratios (reflectances), we square the absolute values [Born and Wolf 1999] such that

$$R_{\perp} = |\tilde{r}_{\perp}|^2 \quad \text{and} \quad R_{\parallel} = |\tilde{r}_{\parallel}|^2 .$$

The reflected  $\perp$ -polarised light is then the amount of incident  $\perp$ -polarised light times the reflectance  $R_{\perp}$ , and the amount of reflected  $\parallel$ -polarised light is the amount of incident  $\parallel$ -polarised light times the reflectance  $R_{\parallel}$ . For unpolarised light, the total reflectance is

$$R = \frac{1}{2}(R_{\perp} + R_{\parallel}) .$$

The transmittances are one minus the reflectances:

$$T_{\perp} = 1 - R_{\perp} \quad , \quad T_{\parallel} = 1 - R_{\parallel} \quad , \quad T = 1 - R .$$

The Fresnel equations illustrate that light may become polarised upon reflection. Polarisation was, however, used mostly as a mathematical convenience in the derivation of the Fresnel equations. Maxwell's equations do not give us any reason why polarisation is needed to model light. We can represent it in the wave theory, but we cannot explain why it is of any physical consequence. The photon spin is the reason why polarisation changes the properties of light. In

Chapter 3 we saw that photons are spin one particles which cannot exist in the rest state. Therefore only two angular momenta ( $-\hbar$  and  $\hbar$ ) are possible for photons. This means that two polarisation components (like the  $\perp$  and  $\parallel$  components chosen here) are appropriate.

This chapter has introduced several different ways to model electromagnetic radiation. The ways have been of diminishing exactitude. We have used more and more simplifications in order to describe a more and more macroscopic case. Let us briefly explore at what levels the material in this chapter allows us to construct rendering algorithms. As the most exact way of rendering using Maxwell's equations, we should use the general solution (4.16–4.17) for the microscopic equations (4.1–4.4). This solution involves only the simplifications that we discussed in the previous chapter as compared to the quantum theory. To succeed with such a renderer, we would have to model materials at an atomic level. It might be possible to model materials at a slightly more macroscopic level using the dipole approximation. We can derive current and charge densities from such dipoles and use them for the integrations that find the vector and scalar potentials of the field. One way to accomplish this integration is using molecular dynamics [Rapaport 2004]. If we compute the vector and scalar potentials for every patch on the image plane (every pixel), we can find the field vectors and with those we can find the Poynting vector leading to the colour values that we need. Of course we would need to evaluate the field at an appropriate number of wavelengths distributed throughout the visible part of the spectrum.

Taking one more step up the ladder towards a more feasible way of rendering realistic images, we arrived at the macroscopic Maxwell equations (4.20–4.23). The simplification was that we introduced two phenomenological vectors (the polarisation and magnetisation vectors) in order to represent materials at a more macroscopic level. Unfortunately we did not find an easier solution for the general version of the macroscopic Maxwell equations. Mostly we used them to move on to the time-harmonic Maxwell equations (4.30–4.33). The simplification at this point was that we introduced the isotropic material equations (4.27–4.29). Beside assuming that the materials are isotropic these equations also involve the simplifying assumptions that the permittivity, permeability, and conductivity of the materials are proportional to the field vectors  $\mathbf{E}$  and  $\mathbf{H}$ . This is certainly not true in general, but the class of materials for which it is true is happily rather large. Unfortunately we did not find a simple general solution for the time-harmonic Maxwell equations either. To find that, we would need something more general than plane waves. When we move to geometrical optics in the next chapter, we will be able to model a more general type of wave at the cost of making assumptions about the wavelength. In this chapter we had to assume that the materials are also homogeneous in order to fit a plane wave solution to the time-harmonic equations.



It is certainly possible to construct a rendering algorithm based on the plane wave solution (4.34) of Maxwell's equations. With the additional simplification of using scalar waves this was indeed the solution employed by Moravec [1981] in his wave-theoretical rendering scheme. I do not know if an implementation of Moravec's algorithm has been attempted on modern hardware, but I suppose reasonable results would be obtained today in comparison to the less successful results in the original paper. Moravec's approach is to model waves by sweeping a 2D array of complex values over an entire scene and storing new sources at points where light is reflected. The reflected light is then propagated in the following sweep going in the opposite direction. The disadvantage of this approach is that the scene must be modelled by finely sliced "object planes" parallel to the image plane. Preferably there should be a plane for every half wavelength. This is a highly impractical and very storage intensive way of representing objects. Extraction of the planes from an implicit surface representation might be the way to go.

Even if we do not want to construct a wave-based rendering algorithm, the electromagnetic field theory is very useful for describing the interaction of light and matter in a more detailed way than would be possible if we did not know about it. The Fresnel equations are a good example of results that we cannot derive without employing some wave theory. So this is the crux of the matter: we need the wave theory (and in some cases even the quantum theory) to describe the interaction of light and matter when the simpler rendering methods fail to do so in sufficient detail. In this chapter we introduced the complex index of refraction at a phenomenological level. To predict theoretical values for the index of refraction, we would have to resort to a quantum description of matter. Similarly, we have to resort to the electromagnetic field theory to predict the phenomenological quantities which we use to describe the scattering of light in conventional rendering algorithms.

In the following chapters, we will move to the theories used for rendering in graphics today. You will see where the phenomenological description of scattering comes in, and a scheme for theoretical prediction of phenomenological scattering properties is presented (using Maxwell's equations) in Part II. First let us explore the most advanced ray theory available. It is often referred to as *geometrical optics*, and it is derived from the theory described in this chapter.

# Bibliography

---

- BOHREN, C. F., AND HUFFMAN, D. R. 1983. *Absorption and Scattering of Light by Small Particles*. John Wiley & Sons, Inc.
- BORN, M., AND WOLF, E. 1999. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, seventh (expanded) ed. Cambridge University Press.
- FEYNMAN, R. P., LEIGHTON, R. B., AND SANDS, M. 1963. *The Feynman Lectures on Physics: Mainly Mechanics, Radiation, and Heat*. Addison-Wesley Publishing Company, Reading, Massachusetts. The Definitive Edition published by Pearson Addison Wesley in 2006.
- FEYNMAN, R. P., LEIGHTON, R. B., AND SANDS, M. 1964. *The Feynman Lectures on Physics: Mainly Electromagnetism and Matter*. Addison-Wesley Publishing Company, Reading, Massachusetts. The Definitive Edition published by Pearson Addison Wesley in 2006.
- MAXWELL, J. C. 1873. *A Treatise on Electricity and Magnetism*. Clarendon Press. Two volumes.
- MORAVEC, H. P. 1981. 3D graphics and the wave theory. *Computer Graphics (Proceedings of ACM SIGGRAPH 81)* 15, 3 (July), 254–261.
- POYNTING, J. H. 1884. On the transfer of energy in the electromagnetic field. *Philosophical Transactions of the Royal Society of London* 175, 343–361.
- RAPAPORT, D. C. 2004. *The Art of Molecular Dynamics Simulation*, second ed. Cambridge University Press.

# Index

---

- absorption coefficient, 71
- conductivity, 67
- conductor, 68
- diamagnetic, 68
- dielectric, 68
- electric susceptibility, 66
- energy density, 63
- Fresnel equations, 74, 75
- homogeneous material, 69
- homogeneous wave, 70
- index of refraction, 70
- inhomogeneous wave, 70
- inverse square law, 65
- isotropic material equations, 67, 68
- law of reflection, 74
- law of refraction, 74
- Lorentz gauge, 65
- magnetic, 68
- magnetisation vector, 66
- Maxwell's equations
  - formal solution, 65
  - Lorentz gauge, 65
  - macroscopic, 67
  - microscopic, 62
  - plane waves, 70
  - time harmonic, 69
- paramagnetic, 68
- photon, 62, 63, 75, 76
- polarisation, 75–76
  - vector, 66
- Poynting's vector, 63
- reflectance, 75
- refractive index, *see* index of refraction
- Snell's law
  - generalised, 74
- wave vector, 68, 70