Chapter *8*

Electromagnetic waves

Introduction

The basic laws of electromagnetism were established in the nineteenth century. Initially, the electric and magnetic interaction were considered to be independent, each of them governed by its own Gauss-like law. Ampere's law provided the first hint of a linkage, since it acknowledged the fact that moving charges produce magnetic fields. The link between electric and magnetic fields was further clarified by Faraday's induction law. Finally, James Clerck Maxwell, completed the set of laws that fully characterize the electromagnetic interaction. Maxwell also went a step further: by combining the electromagnetic equations, he was able to show that electric and magnetic fields in free space satisfy a wave equation in which the wave speed happens to be exactly equal to the measured speed of light. The obvious conclusion was that light is nothing but an electromagnetic wave. In this chapter, we will explore this fascinating discovery.

Maxwell's equations

Maxwell's equations in integral form

When Maxwell started his work on electromagnetism, the known field equations were Gauss' laws for the electric and magnetic fields, Ampère's law, and Faraday's induction law. You probably know these equations in integral form. Gauss' law for the electric field states that

$$\oint_S \mathbf{E} \cdot \mathrm{d}S = \frac{q}{\varepsilon_0} \qquad (1)$$

where **E** is the electric field, q the electric charge enclosed by the surface S and $\varepsilon_0 = 8.854 \times 10^{-12}\ \mathrm{N^{-1}\ m^{-2}\ C^2}$ is the **vacuum permittivity**. Since there are no magnetic charges, Gauss's law for the magnetic field is

$$\oint_S \mathbf{B} \cdot \mathrm{d}S = 0 \qquad (2)$$

Ampère's law relates any current I with the magnetic field **B** it creates:

$$\oint_L \mathbf{B} \cdot \mathrm{d}l = \mu_0 I \quad , \qquad (3)$$

where the magnetic field is integrated along any closed path that encircles the current and μ_0 is the **magnetic permeability of vacuum** $\mu_0 = 4\pi \times 10^{-7}$ m kg C^{-2}. Finally, Faraday's law describes the electric field produced by a changing magnetic flux:

$$\oint_L \mathbf{E} \cdot d\boldsymbol{l} = -\frac{d}{dt}\oint_S B \cdot d\boldsymbol{S} \qquad (4)$$

where the surface integral is over the any surface S enclosed by the line L along which the line integral is performed.

Maxwell noted an inconsistency in the accepted form of Ampère's law. Suppose you apply this law to the circuit illustrated in Fig. 1

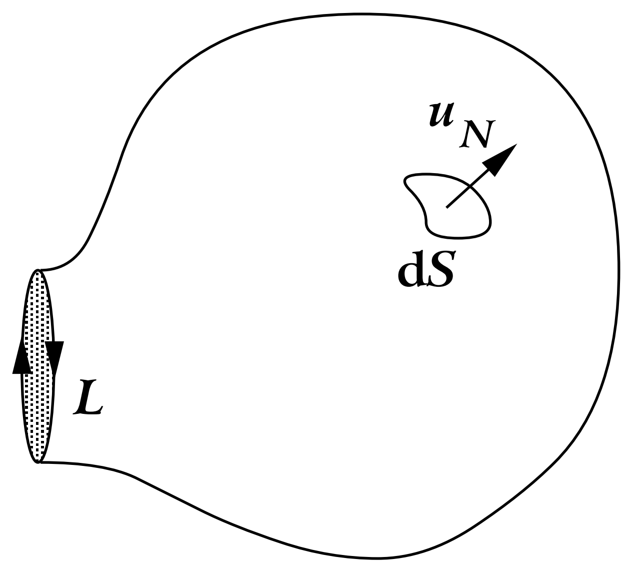

Figure 1 **The line integral of B along L is proportional to the current across any surface S bounded by L. When L shrinks to zero, the surface becomes closed.**

If the line L shrinks to a point, the line integral of $\boldsymbol{B}$ becomes zero. At this point the surface is closed, so that Ampère's law implies that the net current leaving a closed surface is always zero. This is perfectly reasonable if we imagine a wire entering and leaving the surface, for the current that enters the volume enclosed by the surface is exactly equal to the current that leaves the volume. Thus the net current is zero. Imagine, however, that the above surface encloses one of the plates of a capacitor. If the net current flowing across the surface is zero, this means that the net charge enclosed by the surface is constant. In other words, the charge on a capacitor plate should remain constant. This, of course, is not true. Maxwell figured out an *ad hoc* way of fixing this problem. He reasoned that for the case of a closed surface the inconsistency is removed if the right-hand side of Eq. (3) is replaced by μ_0 (I + dq/dt), where q is the charge enclosed by the volume. When the line shrinks to a point, we get I + dq/dt = 0, or I = - dq/dt. This is obviously true, for the net current leaving a volume is *by definition* the rate of change of the charge inside the volume. Using Gauss' law to express the charge in terms of a surface integral of the electric field, Maxwell proposed that Ampère's law be modified to

$$\oint_L \mathbf{B} \cdot \mathrm{d}\boldsymbol{l} = \mu_0 \left(I + \varepsilon_0 \frac{\mathrm{d}}{\mathrm{d}t} \oint_S \mathrm{E} \cdot \mathrm{d}\boldsymbol{S} \right) \qquad (5)$$

For the case of a closed surface, Maxwell's modification of Ampère's law must be correct. But he also assumed that the new expression would be valid even in cases when the surface is not closed. Of course, this would have to be verified by experiment. For cases where the electric field does not depend on time, such as in DC circuits, the new term added by Maxwell is zero and we recover the standard Ampère's law. Eqs. (1), (2), (4), and (140) are known, together, as **Maxwell equations**. They characterize the electromagnetic fields completely. They also lead to the wave equation for the electromagnetic field. However, the wave equation is a differential equation. To show that it follows from Maxwell's equations, we must first rewrite Maxwell's equations in differential form.

Maxwell's equations in differential form

Let us first consider Gauss' law for the electric field. Suppose that we apply Eq. (1) to the infinitesimal volume element in

Fig. (2).

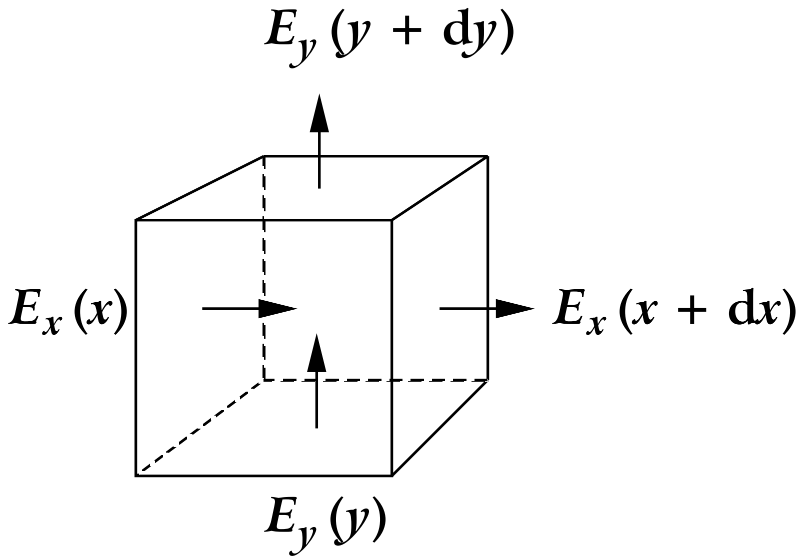

Figure 2 **Infinitesimal volume on which we apply Gauss' law for the electric field**

The surface integral over the six faces of the cube becomes

$$\begin{aligned}\oint_S &= \left[E_x(x+\mathrm{d}x) - E_x(x)\right]\mathrm{d}y\,\mathrm{d}z \\ &+\left[E_y(y+\mathrm{d}y) - E_y(y)\right]\mathrm{d}x\,\mathrm{d}z \\ &+\left[E_z(z+\mathrm{d}z) - E_z(z)\right]\mathrm{d}x\,\mathrm{d}y = \frac{q}{\varepsilon_0}\end{aligned} \qquad ($$

where q is the charge inside the cube. This charge can be written $q = \rho\ \mathrm{d}V = \rho\ \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z.$ On the other hand, the terms inside the square brackets can be written in terms of the derivatives of the field, so that we obtain

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \frac{\rho}{\varepsilon_0} \qquad ($$

This is Gauss' law in differential form. We can now define a special vector that simplifies the notation. Our "vector", denoted as ∇, is defined, in terms of its components, as

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \tag{8}$$

This vector follows the usual rules of vector multiplication, except that multiplication means in this case differentiation. For example, for two vectors **A** and **B** it is known that the dot product can be written as $\mathbf{A}\bullet\mathbf{B} = A_xB_x+A_yB_y+A_zB_z.$ Similarly, the dot product $\boldsymbol{\nabla}\bullet\mathbf{E}$ is given by

$$\nabla \cdot \mathbf{E} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \cdot \left(E_x, E_y, E_z \right) = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \tag{9}$$

Hence we can write Gauss' law for the electric field as

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \tag{10}$$

Similarly, we can write Gauss's law for the magnetic field as

$$\nabla \cdot \mathbf{B} = 0 \tag{11}$$

Faraday's law can also be written in differential form by considering three infinitesimal squares on the planes *XY, YZ,* and *ZX*. Computing the line integral of the electric field along the sides of the squares and equating it to the flux of the electric field, one can show (see homework problem) that Faraday's law in differential form can be written as

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad , \tag{12}$$

where the left-hand side is the vector product between the "vector" $(\partial/\partial x, \partial/\partial y, \partial/\partial z)$ and the real vector **E**. As in the case of Gauss' law, when the vector $\boldsymbol{\nabla}$ "meets" another vector, it differentiates it. Finally, Ampère-Maxwell law can be written as

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \quad , \tag{13}$$

where the vector **j** represents the **current density**. Its magnitude is the current per unit area at a given point, and its direction indicates the direction of the current.

Wave solutions to Maxwell equations

Solution to Maxwell's equations in free space

Let us consider Maxwell's equations in free space, where the charge density ρ as well as the current density **j** are zero. Let us assume that the only non-zero component of the electric field lyes along the y-axis, that is, $\mathbf{E} = (0, E, 0)$ and that the magnetic field has a single component along the z-axis, $\mathbf{B} = (0, 0, B)$. If you don't see any reason for making these assumptions, it's because we are cheating: since the solution to Maxwell's equations will have this form, we will simplify the math a lot by making the assumptions at the outset. However, our solution is no less rigorous: once we find a solution - by whatever dishonest means - we know it must be *the* solution. In advanced courses you will learn how to derive the solutions to Maxwell's equations under the most general conditions.

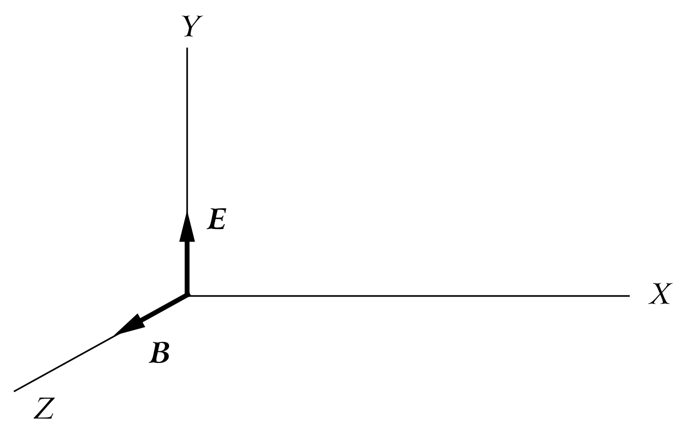

Figure 4 **Orientation of the electric and magnetic fields for the proposed solution to Maxwell's equations in free space.**

Figure 4 shows our electric and magnetic fields, which due to our choice are perpendicular to each other. Notice that although the electric field only has a component along the y-axis, this component will in general be a function of the three coordinates x, y, and z and the time, that is, $E = E(x, y, z, t)$. Similarly, although the orientation of the electric field is along the z-axis, this component will in general be a function of x,y, z, and t: $B = B(x,y,z,t)$. We will now apply the four Maxwell equations to these fields.

A) GAUSS LAW FOR THE ELECTRIC FIELD

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\cdot\left(0,E,0\right)\\ &= \frac{\partial E}{\partial y} = 0\end{aligned} \quad , (14)$$

since the charge density in free space is zero.

B) GAUSS LAW FOR THE MAGNETIC FIELD

$$\begin{aligned}\nabla \cdot \mathbf{B} &= \left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\cdot\left(0,0,B\right)\\ &= \frac{\partial B}{\partial z} = 0\end{aligned} \quad (15)$$

C) FARADAY'S LAW

$$\begin{aligned}\nabla \times \mathbf{E} &= \begin{vmatrix} \mathbf{u}_x & \mathbf{u}_y & \mathbf{u}_z \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ 0 & E & 0 \end{vmatrix} = -\frac{\partial E}{\partial z}\mathbf{u}_x + \frac{\partial E}{\partial y}\mathbf{u}_y \\ &= -\frac{\partial B}{\partial t}\mathbf{u}_z\end{aligned} \quad (16)$$

where $\mathbf{u_x}$, $\mathbf{u_y}$ and $\mathbf{u_z}$ are unit vectors along the directions of the three axes. Notice that Faraday's law is a *vector* equation. It actually contains three scalar equations, one for each component. Equating the different components (and not including the trivial equation $0 = 0$ for the y-direction), we obtain

$$\begin{aligned}\frac{\partial E}{\partial z} &= 0 \\ \frac{\partial E}{\partial x} &= -\frac{\partial B}{\partial t}\end{aligned} \tag{17}$$

D) AMPERE-MAXWELL LAW

$$\nabla \times \mathbf{B} = \begin{vmatrix} \mathbf{u}_x & \mathbf{u}_y & \mathbf{u}_z \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ 0 & 0 & B \end{vmatrix} = \frac{\partial B}{\partial y}\mathbf{u}_x - \frac{\partial B}{\partial x}\mathbf{u}_y \quad , \tag{18}$$

$$= \mu_0 \varepsilon_0 \frac{\partial E}{\partial t}\mathbf{u}_y$$

which implies

$$\begin{aligned}\frac{\partial B}{\partial y} &= 0 \\ \frac{\partial E}{\partial t} &= -\frac{1}{\mu_0 \varepsilon_0}\frac{\partial B}{\partial x}\end{aligned} \tag{19}$$

The results above show that the derivatives of E and B with respect to the y and z coordinates are zero. This means that E and B are functions of x and t only: $E = E(x,t)$ and $B = B(x,t)$. These fields satisfy the coupled equations

$$\begin{aligned}\frac{\partial E}{\partial x} &= -\frac{\partial B}{\partial t} \\ \frac{\partial E}{\partial t} &= -\frac{1}{\mu_0 \varepsilon_0}\frac{\partial B}{\partial x}\end{aligned} \tag{20}$$

These equations can be uncoupled by taking one more derivative and noting that the order in which the derivatives with respect to x and t are performed is irrelevant. This means that the t-derivative of the first equation in (20) equals the x-derivative of the second equation, or

$$\frac{\partial^2 B}{\partial x^2} = \mu_0 \varepsilon_0 \frac{\partial^2 B}{\partial t^2}. \qquad (21)$$

Similarly, by taking the *x*-derivative of the first equation in (20) and comparing with the *t*-derivative of the second equation, we find

$$\frac{\partial^2 E}{\partial x^2} = \mu_0 \varepsilon_0 \frac{\partial^2 E}{\partial t^2} \qquad (22)$$

Thus the electric field and the magnetic field both satisfy a wave equation! The wave speed in this equation is $c = (\mu_0\varepsilon_0)^{-1/2}$. Intrigued, Maxwell used the known values of ε_0 and μ_0 and computed the speed c. He found

$$\begin{aligned} c &= \frac{1}{\sqrt{\mu_0 \varepsilon_0}} \\ &= \frac{1}{\sqrt{1.22\times 10^{-6}\,\mathrm{m\,kg\,C^{-2}} \cdot 8.85\times 10^{-12}\,\mathrm{C^{-2}N^{-1}m^{-2}}}} \\ &= 3.03\times 10^{8}\,\mathrm{m/s} \end{aligned} \qquad (23)$$

This is exactly equal to the speed of light. The conclusion was unavoidable: light is nothing but electromagnetic waves! If confirmed, Maxwell had made one of the most important scientific discoveries of all times. Because he had introduced an *ad hoc* modification to the electromagnetic equations, he felt that his result had to be confirmed experimentally. The confirmation was provided by Heinrich Hertz.

Electromagnetic waves are transverse waves

Since the electric and magnetic fields satisfy the same wave equation as the mechanical waves we have discussed so far, we can immediately write the solutions as traveling or standing waves. For example, one possible solution is $E = E_0 \sin(kx\text{-}\omega t)$. A similar solution would be valid for the magnetic field. The two solutions are related, because the fields are coupled. For example, if $E = E_0 \sin(kx\text{-}\omega t)$, then from Eqs. (20) we obtain $B = B_0 \sin(kx\text{-}\omega t)$, with $kE_0 = \omega B_0$. Using $\omega = ck$, we find

$$E_0 = cB_0 \tag{24}$$

Notice that the wave propagates in the x-direction, whereas the E-field is in the y-direction and the B-field is in the z-direction. Of course, there is nothing special about our choice of axis. We could have oriented the axis in any arbitrary way. This means that our result is completely general: in electromagnetic waves, the electric and magnetic fields are perpendicular to the direction of propagation. Therefore, electromagnetic waves are transverse waves. The electric and magnetic fields are also mutually perpendicular, as shown in Fig. 6

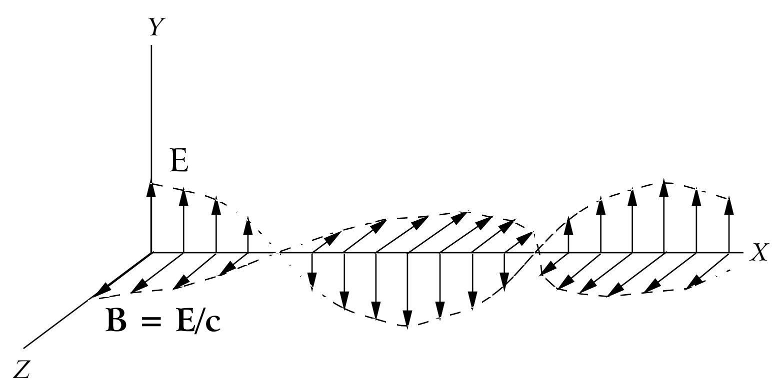

Figure 6 **Schematic diagram of a linearly polarized electromagnetic wave propagating in the *x*-direction**

The direction of the electric field is conventionally regarded as the direction of **polarization**. For the waves in our solution, the field is said to be **linearly** polarized because the electric and magnetic fields are in fixed directions. However, you can show that the fields $\mathbf{E} = [0, E_0 \sin(kx\text{-}\omega t), E_0 \cos(kx\text{-}\omega t)]$ and $\mathbf{B} = [0, B_0 \cos(kx\text{-}\omega t), -B_0 \sin(kx\text{-}\omega t)]$

satisfy the wave equation. In this case, the electric and magnetic field are still mutually perpendicular, but their direction no longer fixed in space: they rotate in the y-z plane. This type of wave is said to be **circularly polarized**. The figure above represents a linearly polarized sinusoidal wave.

The meaning of the wave equation for the fields: From ether to relativity

For the first time, we have encountered a wave equation that is not derived from Newton's second law. The wave equation for electromagnetic waves is a direct consequence of Maxwell's equations. Unlike the wave equation for elastic waves in matter, which is an approximation only valid when the wavelength is much longer than the separation between atoms, the wave equation for electromagnetic waves is exact.

Since elastic waves in matter represent the vibration of a continuous medium, such as air, water, a solid, etc., it was initially thought that the electromagnetic waves represent the vibration of a certain medium, which was called "ether". Aside from the obvious analogy, there was a more disturbing reason why physicists hung to the idea of "ether" for a long time. Physicists expect all physical laws to be identical when tested from different inertial reference frames. For example, if observer A is at rest and observer B moves with constant velocity relative to A, both will measure the same acceleration for a given object. In other words, Newton's laws are the same for the two observers. On the other hand, suppose that observer A is at rest with respect to the medium where an elastic wave propagates, but observer B is in motion with some constant velocity. Now B will not see the same wave equation as A. We have already confronted this issue (although not in the context of the wave equation) when we discussed the Doppler effect. However, there is no conceptual problem here, because A and B are not equivalent "by symmetry". One of them is moving relative to the medium and the other is not.

If we apply the same analysis to electromagnetic waves, we also find that different observers see different wave equations. This is no problem if there is ether, because Maxwell equations would be valid only for a reference frame at rest relative to the ether. But if there is no ether and Maxwell equations are to be valid for *all* inertial observers, then all these observers should "see" the *same* wave equation. The only way to make the wave equation the same in all inertial reference frames is to abandon the Galileo transformations that relate the motion of objects seen by different observers. For example, the well-known (and obvious) addition-of-velocities theorem can no longer be valid. This was absurd and unacceptable to most physicists in the late XIXth century. So they concluded there *must* be something called ether. Hence they devised experiments (the famous Michelson-Morley series) to measure the speed of the Earth relative to the ether. This velocity was found to be zero. In other words, the Earth was found to be *the* object in the Universe which was at rest with respect to the ether. This idea would have been attractive to the Church officials who "helped" people like Galileo correct his "mistakes," but it was unacceptable to the physicists of the early 20th century. So they were forced to abandon the standard transformation laws and embrace the so-called Lorentz transformations, which have the property that they make the wave equation for light look the same in all reference frames. The first to propose this revolution was Albert Einstein in a 1905 paper entitled "On the electrodynamics of moving bodies." You can now understand the reason for the word "electrodynamics" in the first paper on relativity theory: relativity is almost a necessity once we accept Maxwell's equations.

Energy in the electromagnetic field

In your previous physics courses, you found that the energy needed to create an electric field is given by

$$\mathrm{W} = \tfrac{1}{2}\varepsilon_0 \int_{\mathrm{V}} E^2 \mathrm{d}V, \qquad (25)$$

where the integral over the entire volume occupied by the electric field. To avoid confusion between the notation for electric field and energy, we use here W for energy. We can therefore define an electric **energy density** E given by

$$\mathcal{E}_E = {}^1/_2\ \varepsilon_0 E^2 \qquad (26)$$

Similarly, the energy needed to set up a magnetic field **B** is given by

$$W = \frac{1}{2\mu_0}\int_V B^2 \mathrm{d}V\,, \qquad (27)$$

so that the magnetic **energy density** is given by

$$\mathcal{E}_B = \frac{1}{2\mu_0}B^2 \qquad (28)$$

An electromagnetic wave has both electric and magnetic fields, so that its energy density will be the sum of Eq. (26) and Eq. (28). Thus we obtain

$$\mathcal{E} = \mathcal{E}_E + \mathcal{E}_B = {}^1/_2\ \varepsilon_0 E^2 + \frac{1}{2\mu_0}B^2 = \varepsilon_0 E^2 \qquad (29)$$

where we have used $c^2 = 1/(\varepsilon_0\,\mu_0)$ and $B = E/c$. Notice that for an electromagnetic field the magnetic and electric energies are equal. In a previous chapter, we showed that the **intensity** of a wave, that is, the energy per unit time passing across a unit surface perpendicular to the direction of propagation, is $I = c\mathcal{E}$. Hence the intensity of an electromagnetic wave is given by

$$I = c\varepsilon_0 E^2 \qquad (30)$$

Momentum in the electromagnetic field

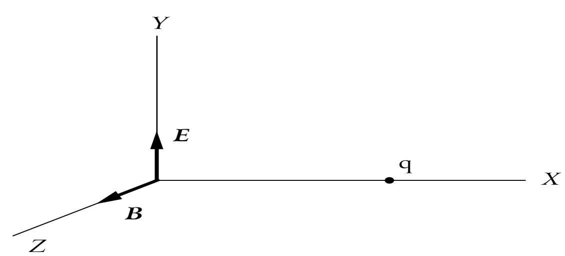

Figure 7 **A charge under the influence of an electromagnetic wave absorbs a net momentum from the fields.**

Let us now consider a charge q, initially at rest at some point on the x-axis. An electromagnetic wave, with the electric field polarized in the y-direction an the magnetic field polarized in the z-direction, approaches the charge. The force acting on the charge is given by the well-know Lorentz formula

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right) \tag{31}$$

If we write $\mathbf{v} = (v_x, v_y, v_z)$, $\mathbf{E} = (0, E_y, 0)$, $\mathbf{B} = (0, 0, B_z)$, we can express this force as

$$\mathbf{F} = \left(qE_y - qv_x B_z\right)\mathbf{u}_y + qv_y B_z \mathbf{u}_x \tag{32}$$

where the $\mathbf{u}$'s are unit vectors along the three axes. Let us now average the force over one cycle of the oscillation. Clearly, $<qE_y> = 0$. On the other hand, we can easily show that v_x changes much more slowly that v_y. The acceleration in the x- direction can be written as

$$\frac{\mathrm{d}v_x}{\mathrm{d}t} = \frac{q}{m} v_y B_z = \frac{q}{m} v_y \frac{E_y}{c} = \left(\frac{v_y}{c}\right)\frac{qE_y}{m} = \left(\frac{v_y}{c}\right)\frac{\mathrm{d}v_y}{\mathrm{d}t} \tag{33}$$

Because $v_y/c << 1$ for non-relativistic objects (our final result will be correct even for relativistic particles, but the derivation must take into account the relativistic expression for momentum. The derivation presented here is only valid for particles that move at speeds much smaller than the speed of light) we conclude that the acceleration in the x-direction is much smaller than the acceleration in the y-direction. This means that v_x remains almost constant over one oscillation cycle of the electromagnetic fields. Hence we can write-$q<v_x B_z> \simeq q v_x <B_z> = 0$. Therefore, the force in the y-direction averages to zero and there is no momentum transfer in the y-direction. We can thus write

$$\left\langle \mathbf{F} \right\rangle = q\left\langle v_y B_z \right\rangle \mathbf{u}_x \tag{34}$$

Let us now compare this expression with the average work done on the charge per unit time. Only the electric force does work, so that we can write

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \mathbf{F}\cdot\mathbf{v} = q\mathbf{v}\cdot\mathbf{E} = qv_y E_y \tag{35}$$

The average of this over a cycle is

$$\left\langle \frac{\mathrm{d}W}{\mathrm{d}t} \right\rangle = q\left\langle v_y E_y \right\rangle \tag{36}$$

If we note that $B_z = E_y/c$ and compare Eq. (34) with Eq. (35), we conclude that

$$\left\langle \mathbf{F} \right\rangle = \left\langle \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} \right\rangle = \frac{1}{c}\left\langle \frac{\mathrm{d}W}{\mathrm{d}t} \right\rangle \mathbf{u}_x \tag{37}$$

Hence the momentum transfer and the energy transfer and directly proportional. In analogy with the energy case, we can define a momentum density $\mathcal{P}$ given by

$$\mathcal{P} = \mathcal{E}/c\ \mathbf{u_x} \tag{38}$$

Here $\mathcal{P}$ is the momentum per unit volume contained in an electromagnetic wave.

Why do fields carry energy and momentum?

We have computed the energy density and the linear momentum density in electromagnetic waves. We can also show that there is a corresponding angular momentum density associated with these waves. At this point you may wonder what is the meaning of all this. When the concepts of energy and momentum were introduced in previous physics courses, they were associated with particles having a certain mass. However, now we seem to be associating the energy and the momentum with the fields themselves. We say that the wave “carries” energy and momentum. In the case of elastic waves, this is no problem because the waves involve the motion of masses, so that the energy we are talking about is the standard energy of vibrating particles. But there are no “particles” vibrating in an electromagnetic wave, it’s the fields themselves which oscillate. How can the fields “have” energy and momentum? The whole idea is even more disturbing when you remember that the fields were introduced as a mathematical artifact. The “real” thing was the force between charges, the fields were viewed only as a practical way to compute those forces. When you derived the "energy" of an electromagnetic field as proportional to the integral of E^2 over the volume, this was little more than a mathematical curiosity. It was always clear that you were talking about the standard potential energy of a system of charges.

When electromagnetic waves are considered in their proper relativistic context, however, they *must* carry energy and momentum. The reason why the total momentum of a system of particles is conserved is Newton's third law. Because any pair of particles exert equal and opposite forces on each other, the momentum gained by one of them is exactly compensated by the momentum lost by the other one. But Newton's third law is not valid for charges in motion, because the force (the electric and magnetic fields) cannot travel faster than the speed of light. When a charge is suddenly shaken, it takes some time for a second particle to "feel" this displacement. If Newton's third law is not valid, then momentum cannot not be conserved *unless* we assume that the missing momentum is "traveling" with the wave. Hence only the sum $\mathbf{p}_{particles} + \mathbf{p}_{fields}$ can remain constant. To the extent that we believe that energy and momentum are conserved, we must accept the fact that the fields carry these quantities. At this point, the fields cease to be a mathematical artifact. They are as "real" as the particles which generate them.

Radiation

We have so far discussed the solution to Maxwell equations in free space, where $\rho = 0$ and $\mathbf{j} = 0$. Of course, a possible solution is also $\mathbf{E} = 0$ and $\mathbf{B} = 0$, *i.e.*, no field at all. In free space, this solution is just as good as the electromagnetic waves we proposed above. The reason why the wave is the right solution is that somewhere in space, maybe far from our "free" space, the charge and/or the current is not zero. Thus $\mathbf{E}$ and $\mathbf{B}$ cannot be zero at those points. On the other hand, if charge and current are zero *everywhere*, then the zero-field solution is the right solution and there are no waves.

This is equivalent to saying that electromagnetic waves originate in charges and currents. If we solve Maxwell's equations in the presence of charge and currents, we can show how the electromagnetic waves are **radiated** by accelerated charges and fluctuating currents. The mathematics of this, however, is complicated enough to make it a topic for advanced courses. However, if we admit the relativistic principle that no "information" can travel faster than the speed of light, we can at least begin to understand how a transverse wave can be generated. Consider the charge in Fig. 9.

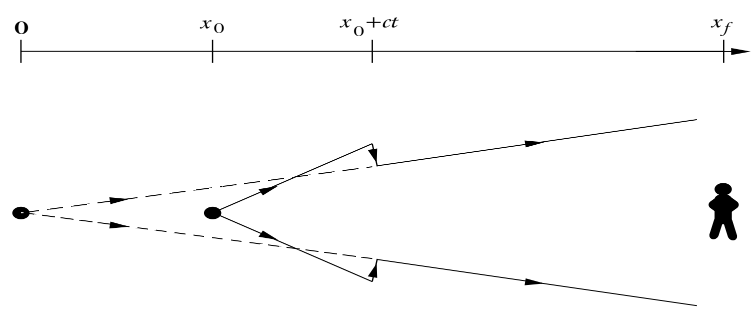

Figure 9 **Transverse electromagnetic waves are produced when the charge at the origin is suddenly displaced to position x_0.**

Before $t = 0$, the charge it at rest at the origin and a distant observer at x_f sees the usual Coulomb field emanating radially from the charge. Note that the electric field is radial, not transverse. At $t = 0$, the charge is suddenly displaced to $x = x_0$. Now the charge will produce a radial Coulomb field whose origin is at x_0. The observer, however, will not see an immediate change in the field he detects, for if he would, he would be able to tell instantaneously that the charge has moved. In other words, the information that the charge has moved would travel in no time from x_0 to x_f. This violates relativity. So the observer must see the "old" Coulomb field for a while. The "new" Coulomb field front cannot travel toward him at a speed higher than c. On the other hand, Gauss' law applies everywhere; in particular, at the boundary between the "old" and the "new" Coulomb field. But because there is no charge at this boundary, the field lines must be continuous. Thus the field lines must bend and the field will be transverse at the boundary between the "old" and the "new" Coulomb field. This transverse section is nothing but the electromagnetic wave produced by the sudden motion of the charge.

If the charge had been moving at constant velocity, it would produce a standard Coulomb electric field plus a magnetic field. The information that the charge is moving would have infinite time to arrive to the observer, and relativity would not be violated. It is only when the velocity is changed, *i.e.,* when the charge is accelerated, that electromagnetic waves are produced. We can thus state

> **Electromagnetic waves are produced by accelerated charges.**

Radiation and Quantum Mechanics

Our conclusion that charges under acceleration radiate electromagnetic waves has a dramatic impact on our understanding of the atomic structure. Experiments suggest that electrons in atoms orbit around the nucleus in much the same way planets orbit around the Sun. But an electron in orbit around the nucleus is under acceleration (centripetal acceleration) and must radiate electromagnetic waves. Because these waves carry energy away from the electron, the electron should rapidly spiral down to the nucleus: the atom should not be stable! This paradox can only be solved with quantum mechanics. According to quantum theory, the electron can be viewed as a wave like the standing waves in a cord. When the electron is in one of the "normal modes" the wave is stationary and no radiation is emitted. When the electron changes from one normal mode to a different one, the electron wave is suddenly modified and radiates electromagnetic waves during the transition.

It is ironic that Maxwell equations, the most impressive achievement of classical physics, brought about the demise of the very fundamentals on which they were built. Today we now that the Galilean transformations have to be replaced by relativistic transformations and that Quantum Mechanics replaces Newtonian Mechanics. Maxwell equations themselves are the only "surviving" element of classical physics.

Maxwell's equations inside a material

Dipole moment and polarization

The **dipole moment** of a system of charges is defined as

$$\mathbf{p} = \sum_i q_i \mathbf{r}_i \quad , \tag{39}$$

where $\mathbf{r}_i$ is the position vector of charge i. When the system consists of two charges q and $-q$, Eq.(39) reduces to the familiar expression $\mathbf{p} = q\boldsymbol{a}$, where $\boldsymbol{a}$ is a vector with its origin at the negative charge and its tip at the positive charge.

The polarization $\mathbf{P}$ of a material is defined as the dipole moment per unit volume. If a certain substance has N atoms or molecules per unit volume, each with a dipole moment $\mathbf{p}$, the polarization is given by $\mathbf{P} = N\mathbf{p}$.

In this discussion, we will consider materials that do not have a permanent polarization, but become polarized in the presence of an external electric field. We will also assume that the *induced* polarization $\mathbf{P}$ is parallel to the external electric field $\mathbf{E}$. This is expected for homogeneous media and for cubic crystals.

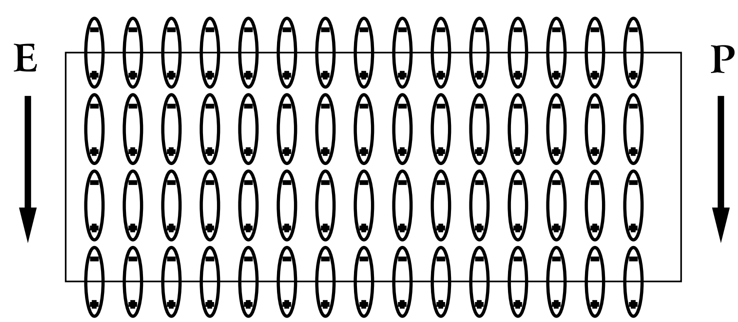

Figure 10 **Microscopic picture of the polarization induced in a material by an external electric field.**

Let us consider any non-conducting material in the presence of an external electric field.The atoms or molecules will be distorted in such a way that a net charge appears at the surface. It is easy to calculate the magnitude of the charge. Suppose that the positive and negative charges in each "atom" of Fig. 10 is q. The electric field separates these charges by a distance a. If the number of "atoms" per unit volume is N and the surface area is S, then the charges at the surface occupy a volume Sa. Thus, the total charge at the surface is $Q = NSaq$. The surface density of charge, defined as $\sigma = Q/S$, is given by $\sigma = Nqa$. But this is precisely the magnitude of the polarization, so that we find $\sigma = P$. If the dipole moment makes an angle θ with the sample's surface, the volume occupied by surface charge becomes $Sa \cos\theta$, so that $\sigma = P\cos\theta$. This can be written in general as

$$\sigma = \mathbf{P} \cdot \mathbf{u}_n \tag{40}$$

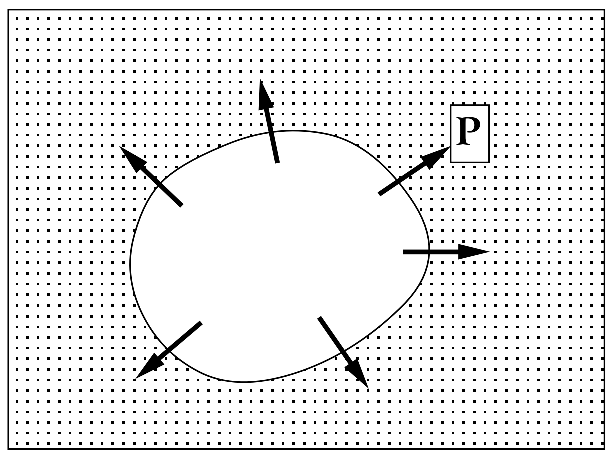

Figure 11 **Any volume inside a dielectric with non-uniform polarization can have an excess or defect of charge**

Let us now consider any volume inside a dielectric material with a non-uniform polarization, as in Fig. 11. At each point at the surface some charge will be displaced just outside the surface. If the polarization is non-uniform, the net charge inside the volume might change because the charges moving in and out of the surface at different points are not equal. Let us compute the net change in charge due to the polarization, which we will call ΔQ_{pol}. The change of charge *inside* the volume is equal to minus the "surface" charge, so that we obtain

$$\Delta Q_{pol} = -\oint_S \mathbf{P} \cdot \mathbf{n} \, \mathrm{d}S \qquad (41)$$

where the integral runs over the entire surface of the volume. On the other hand, we can define a polarization charge density ρ_{pol} as

$$\Delta Q_{pol} = \oint_V \rho_{pol} \, \mathrm{d}V \qquad (42)$$

Combining the two expressions, we obtain

$$\oint_V \rho_{pol} \, \mathrm{d}V = -\oint_S \mathbf{P} \cdot \mathbf{n} \, \mathrm{d}S \qquad (43)$$

This is a Gauss-like form for the polarization. By analogy with the standard forms of Gauss' law, we can immediately convert this equation to a differential form

$$\rho_{pol} = -\nabla \cdot \mathbf{P} \qquad (44)$$

The equation of continuity

Whenever the polarization is a function of time, the charge will also change with time. This means that there will be currents, because the current I leaving a volume is equal to-dQ/dt if Q is the charge inside this volume. This can be converted to a differential form. Let us consider an infinitesimal volume as in Fig. 12

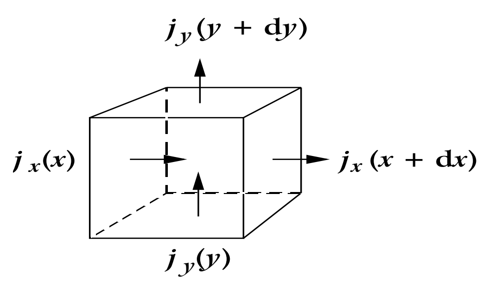

Figure 12 **Infinitesimal volume on which we apply the relation I = -dq/dt.**

Since the current density **j** represents the current per unit area crossing a surface, the net charge loss of the cube per unit time is given by

$$-\frac{dQ}{dt} = \left[j_x(x+dx) - j_x(x) \right] dy\,dz + \left[j_y(y+dy) - j_y(y) \right] dx\,dz + \left[j_z(z+dz) - j_z(z) \right] dx\,dy \qquad (\,$$

where Q is the charge inside the cube. This charge can be written $Q = \rho\,dV = \rho\,dx\,dy\,dz$. On the other hand, the terms inside the square brackets can be written in terms of the derivatives of the current density, so that we obtain

$$\frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z} = -\frac{\partial \rho}{\partial t} \quad , \qquad (\,$$

which can be expressed in a more compact form as

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \qquad (47)$$

This expression is universally valid, and is known as the equation of continuity. It can be viewed as a "charge conservation" statement: because charge is not created or destroyed, the charge loss per unit time in a volume must be equal to the electric currents exiting the volume. Notice that in Chapter 7 we found a very similar equation for the case of water waves. Eq. (4) in Chapter 7 can be written as $\nabla \bullet \boldsymbol{\psi} = 0$. The right-hand side of the continuity equation becomes zero because we assume that water is incompressible and that no bubbles are formed in the liquid.

Maxwell equations with polarization charges

Inside a polarizable medium, it is convenient to split the charge density into two terms:

$$\rho = \rho_{free} + \rho_{pol} \ , \qquad (48)$$

where ρ_{free} corresponds to the charges that are unbound, free to move (as in a metal) and ρ_{pol} is the polarization charge density given by Eq. (44). Similarly, the current density can be split as

$$\mathbf{j} = \mathbf{j}_{free} + \mathbf{j}_{pol} \tag{49}$$

Each term satisfies its own continuity equation. For the polarization charges, for example,

$$\nabla \cdot \mathbf{j}_{pol} = -\frac{\partial \rho_{pol}}{\partial t} = \frac{\partial \nabla \mathbf{P}}{\partial t} \tag{50}$$

which means

$$\mathbf{j}_{pol} = \frac{\partial \mathbf{P}}{\partial t} \tag{51}$$

Suppose that we have a solid with no free charges and currents. A good example would be an insulating material, such as a piece of glass. In this case, we can write Maxwell's equations as

(Gauss' law for **E**) $$\nabla \cdot \mathbf{E} = \frac{\rho_{pol}}{\varepsilon_0} = -\frac{\nabla \mathbf{P}}{\varepsilon_0} \tag{52}$$

(Gauss' law for **B**) $$\nabla \cdot \mathbf{B} = 0 \tag{53}$$

(Induction law) $$\nabla \times E = -\frac{\partial \mathbf{B}}{\partial t} \tag{54}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}_{pol} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \frac{\partial \mathbf{P}}{\partial t} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{55}$$

(Ampère-Maxwell law)

Let us try to find a solution to these equations in the way we found a solution to Maxwell's equations in vacuum. We will assume, as we did in the vacuum case, that the electric field has a single component in the y-direction, $\mathbf{E} = (0.E,0)$ and the magnetic field a single component in the z-direction, $\mathbf{B} = (0,0,B)$. Also, since the polarization is induced in the material by the electric field, we will assume that **P** is proportional to **E**, so that both vectors point in the same direction. Hence the polarization vector is given by **P** =

(0,P,0). If we apply Maxwell's equations in exactly the same way we did in the vacuum case, we end up with the set of equations

$$\frac{\partial E}{\partial y} = -\frac{1}{\varepsilon_0}\frac{\partial P}{\partial y} \tag{56}$$

$$\frac{\partial E}{\partial x} = -\frac{\partial B}{\partial t} \tag{57}$$

$$\frac{1}{\varepsilon_0}\frac{\partial P}{\partial t} + \frac{\partial E}{\partial t} = -\frac{1}{\varepsilon_0 \mu_0}\frac{\partial B}{\partial x} \tag{58}$$

We can satisfy Eq. (56) by assuming that E (and, therefore, its proportional quantity P) is a function of x and t only, so that the partial derivatives with respect to y give zero. Taking the x-derivative of Eq.(57) and the t-derivative of Eq. (58), we eliminate the magnetic field and obtain

$$\frac{1}{\varepsilon_0}\frac{\partial^2 P}{\partial t^2} + \frac{\partial^2 E}{\partial t^2} = c^2 \frac{\partial^2 E}{\partial x^2} \tag{59}$$

Because of the term that contains the polarization, the electric field does not satisfy a standard wave equation. Hence inside a material $\omega \neq ck$ and the electromagnetic waves may become dispersive. Since the polarization is a property of the medium, Maxwell's equations will give different solutions depending on the medium. In order to find solutions to Eq. (59), we must first find the relationship between E and P.

Elementary theory of the polarization

We will consider the simplest possible model. We assume that the polarization arises from the relative displacement of electrons and positive nuclei in atoms. The force between electrons and nuclei is represented by a spring of constant K. (The force is actually a Coulomb force, not a spring force. Our model is not very realistic from a classical point of view, but the final result is very similar to the exact expression obtained from a correct quantum-mechanical treatment of the problem.). Let's assume that an electric field of the form $E = E_0 \sin \omega t$ acts on an electron of mass m_e. The equation of motion is thus

$$m_e \frac{d^2 x}{dt^2} = -Kx - eE_0 \sin \omega t \tag{60}$$

where $-e$ is the electron charge. This equation is identical to the forced oscillations we discussed in Chapter 2. Its solution is

$$x(t) = -\frac{e}{m_e(\omega_0^2 - \omega^2)} E_0 \sin \omega t \tag{61}$$

with $\omega_0^2 = K/m$.The induced electric dipole is $p = -e\, x$; if there are N atoms per unit volume, the polarization is given by

$$P = \frac{e^2 N}{m_e(\omega_0^2 - \omega^2)} E \tag{62}$$

if the electric field is sinusoidal. Let us now go back to Eq.(59). Using Eq. (62) to eliminate the polarization, we obtain

$$\left[\frac{e^2 N}{\varepsilon_0 m_e(\omega_0^2 - \omega^2)} + 1\right] \frac{\partial^2 E}{\partial t^2} = c^2 \frac{\partial^2 E}{\partial x^2} \tag{63}$$

This can be written as

$$\frac{\partial^2 E}{\partial t^2} = \left(\frac{c}{n}\right)^2 \frac{\partial^2 E}{\partial x^2} \tag{64}$$

where we have introduced the **index of refraction** n, given by the expression

$$n^2 = 1 + \frac{e^2 N}{\varepsilon_0 m_e(\omega_0^2 - \omega^2)} \tag{65}$$

The importance of the index of refraction is that the electromagnetic waves in the medium travel with speed c/n. The speed of light in a medium depends on the frequency of the light. Red light and blue light travel with different speeds. This is the reason, as we will see later, for the separation of colors in a prism. When $\omega = \omega_0$, the expression for the refraction index diverges. In this case, we cannot neglect the damping term in the equation of motion of the electrons, as we did in Eq. (60). When damping is present, energy is dissipated. For the case of electromagnetic waves, this corresponds to light being absorbed by the material. Thus ω_0 corresponds to typical frequencies of light absorption in the material. Many insulators, such as glasses, are transparent in the visible. For these materials, ω_0 is in the ultraviolet.

Metallic behavior

The results of the previous section apply to a system with no free charges, such as an insulator. The polarization charges are attached to their sites by the spring constant K. If the spring constant becomes zero, however, this is equivalent to setting the polarization charge free, so that the system becomes a metal. Hence the refraction index of a metal is given by Eq. (65) with $\omega_0 = 0$:

$$n^2 = 1 - \frac{e^2 N}{\varepsilon_0 m_e \omega^2} = 1 - \frac{\omega_p^2}{\omega^2} \qquad (66)$$

where ω_p is the so-called plasma frequency, $\omega_p^2 = \frac{e^2 N}{\varepsilon_0 m_e}$. For $\omega < \omega_p$, $n^2 < 0$, which is impossible. This means that no wave can propagate inside a metal at frequencies below the plasma frequency. If a wave of frequency $\omega < \omega_p$ is incident upon a metal, the wave is completely reflected. In a real metal, $N \sim 10^{28}$ m^{-3}. Hence the plasma frequency corresponds to a wavelength of 3300 Å, which is in the ultraviolet. Visible light is completely reflected from a typical metal.

A good example of a "metal" is the Earth's ionosphere, which consists of air molecules ionized by the radiation emitted by the sun. The ionosphere charge density is maximum at a distance of 200 - 400 km above ground. AM radio waves can reach different points on the surface of the Earth by reflection in the ionosphere. The ionosphere plasma frequency is of the order of 20 MHz. This corresponds to a charge density of 1.2×10^{12} m^{-3}, many orders of magnitude less that the charge density in a standard metal.

For $\omega > \omega_p$, $n < 1$, so that the velocity v of the wave is grater than c! This, however, does not violate relativity because the corresponding group velocity is less than the speed of light in vacuum. Suppose that the electric field is given by $E = E_0 \sin(kx\text{-}\omega t)$. Then Eq. (64) gives

$$n^2 = \frac{c^2 k^2}{\omega^2} \tag{67}$$

Inserting the corresponding expression for n, given in Eq. (66), one obtains, after some rearranging,

$$v_g = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{c^2}{\omega} k \tag{68}$$

which is clearly less than c whenever $\omega/k > c$.

Problems

1.(Alonso 29.1) The electric field (in V/m) of a plane electromagnetic wave in vacuum is represented by $E_x = 0$, $E_y = 0.5 \cos [2\pi \times 10^8 \text{ s}^{-1} (t\text{-}x/c)]$ and $E_z = 0$.
a) Determine (i) the wavelength, (ii) the state of polarization and (iii) the direction of propagation.
b) Calculate the magnetic field of the wave.
c) Calculate the average intensity, or energy flux per unit area and per unit time, of the wave.

2.(Alonso 29.4) A plane sinusoidal linearly polarized electromagnetic wave of wavelength $\lambda = 5.0 \times 10^{-7}$ m travels in vacuum in the direction of the *X*-axis. The average intensity of the wave per unit area is 0.1 W/m^2 and the plane of oscillation of the electric field is parallel to the *Y* axis. Write the equations describing
a) the electric field,
b) the magnetic field.

3. (Alonso 29.5) The electric field of a plane electromagnetic wave has an amplitude of 10^{-2} V/m. Find
a) the magnitude of the magnetic field,
b) the energy per unit volume of the wave.
c) If the wave is completely absorbed when it falls on a body, determine the radiation pressure.
d) What is the radiation pressure if the body is a perfect reflector?

4. Derive Faraday's law and Ampère-Maxwell's laws in differential form.

5. Consider an electric field given by $E_y = A \cos \omega(t\text{-}x/c)$, $E_z = A \sin \omega(t\text{-}x/c)$, $E_x = 0$.
a) Show that this field satisfies the wave equation for waves propagating along the *x*-axis.
b) Calculate the corresponding magnetic field.
c) Is this wave linearly polarized? Are the electric and magnetic fields perpendicular to each other?

6, Repeat Problem 5 for the case where $E_z = -A \cos \omega(t\text{-}x/c)$. All other quantities remain equal.

7. Two sinusoidal electromagnetic waves, both of frequency ν and amplitude E_0, travel in vacuum in the *X* and *Y* direction, respectively. The electric fields if both waves are parallel to the *Z*-axis. Calculate
a) The components of the total electric field
b) The components of the total magnetic field
c) The energy density E. Can it be written as the sum of the energy densities of each individual wave?
d) Determine the planes over which the mean value of E^2 is maximum or minimum. Discuss the connection of this result with your result in part *c)*.

8. Show that for typical values in a gas, the second term in Eq. (65) is small. Show that in that case the index of refraction can be written as

$$n = 1 + \frac{Ne^2}{2m_e\varepsilon_0\left(\omega_0^2 - \omega^2\right)}$$

9. The index of refraction of gaseous hydrogen at normal pressure and temperatures is $n = 1 + 1.400 \times 10^{-4}$ for $\lambda = 5.46 \times 10^{-7}$ m and $n = 1 + 1.547 \times 10^{-4}$ for $\lambda = 2.54 \times 10^{-7}$ m. Assuming that the formula in the previous problem is valid, calculate ω_0 and *N*. How many molecules per unit volume do you expect in this gas?

10. Consider a gas whose molecules behave like dipole oscillators with $K = 300$ N/m. The oscillating particles are electrons. Calculate their characteristic frequency. Express the index of refraction in terms of the frequency, assuming that the gas is at normal pressure and temperature.

11, Show that the group velocity of very high frequency electromagnetic waves is given by

$$v_g = \frac{c}{1 + (Ne^2/2\varepsilon_0 m_e \omega^2)}$$

12. Suppose that the valence electron of silver atom becomes a free electron in solid silver. Find from any *Handbook of Physics and Chemistry* book the valence of silver, its atomic weight, and its mass density. Thus find the number *N* of free electrons per unit volume in solid silver.
a) Calculate the plasma frequency in silver. Show

that for visible light the frequency of the light is below the plasma frequency.
b) For what frequencies should the silver layer become transparent?