

## Chapter 20

# Photons: Maxwell's Equations in a Nutshell

### 20.1 Introduction

Light has fascinated us for ages. And deservedly so. Everything we know about the earth and the universe is because of light. Light from the sun sustains life on earth. Learning to measure and understand the contents of light has enabled us to understand the origins of the universe in the big bang, and talk about its future. And one cannot forget the sheer visual pleasure of a beautiful sunset, a coral reef, or an iridescent flower in full blossom. Indeed, the beauty of light and color is a rare thing that scientists and artists agree to share and appreciate.

Our fascination with light has led to three of the greatest revolutions in 19<sup>th</sup> and 20<sup>th</sup> century physics. Sunlight used to be considered a 'gift of the Gods' and the purest indivisible substance, till Newton observed that passing it through a prism split it into multiple colors. Passing each of the colors through another prism could not split it further. Newton surmised that light was composed of particles, but in the early 19<sup>th</sup> century, Young proved that light was a wave because it exhibited interference and diffraction. Michael Faraday had a strong hunch that light was composed of a mixture of electric and magnetic fields, but could not back it up mathematically. The race for understanding the fabric of light reached a milestone when Maxwell gave Faraday's hunch a rigorous mathematical grounding. Maxwell's theory combined in one stroke electricity, magnetism, and light into an eternal braid<sup>1</sup>. The Maxwell equations predict the existence of light

---

<sup>1</sup>J. R. Pierce famously wrote "To anyone who is motivated by anything beyond the most narrowly practical, it is worthwhile to understand Maxwell's equations simply for the good of his soul."

as a propagating electromagnetic wave. With Maxwell's **electromagnetic theory**, the 'cat' was out of the hat for light.

The second and third revolutions born out of light occurred in early 20<sup>th</sup> century in parallel. Trying to understand blackbody radiation, photoelectric effect, and the spectral lines of hydrogen atoms lead to the uncovering of **quantum mechanics**. And Einstein's fascination with the interplay of light and matter, of space and time led to the theory of **relativity**. Much of modern physics rests on these three pillars of light: that of electromagnetism, quantum mechanics, and relativity. It would be foolhardy to think that we know all there is to know about light. It will continue to amaze us and help probe deeper into the fabric of nature through similar revolutions in the future. In this chapter, we discuss Maxwell's theory of electromagnetism in preparation for the quantum picture, which is covered in the next chapter.

## 20.2 Maxwell's equations

Maxwell's equations connect the electric field  $\mathbf{E}$  and the magnetic field intensity  $\mathbf{H}$  to source charges  $\rho$  and currents  $\mathbf{J}$  via the four relations

$\nabla \cdot \mathbf{D} = \rho,$	Gauss's law	(20.1)
$\nabla \cdot \mathbf{B} = 0,$	Gauss's law	
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},$	Faraday's law	
$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t},$	Ampere's law.	

Here the source term  $\rho$  has units of charge per unit volume ( $\text{C}/\text{m}^3$ ), and current source term  $\mathbf{J}$  is in current per unit area  $\text{A}/\text{m}^2$ .  $\mathbf{H}$  is related to the magnetic flux density  $\mathbf{B}$  via  $\mathbf{B} = \mu_0 \mathbf{H}$ , and the displacement vector is related to the electric field via  $\mathbf{D} = \epsilon_0 \mathbf{E}$ . The constant  $\epsilon_0$  is the permittivity of vacuum, and  $\mu_0$  is the permeability of vacuum. They are related by  $\epsilon_0 \mu_0 = 1/c^2$ , where  $c$  is the speed of light in vacuum.

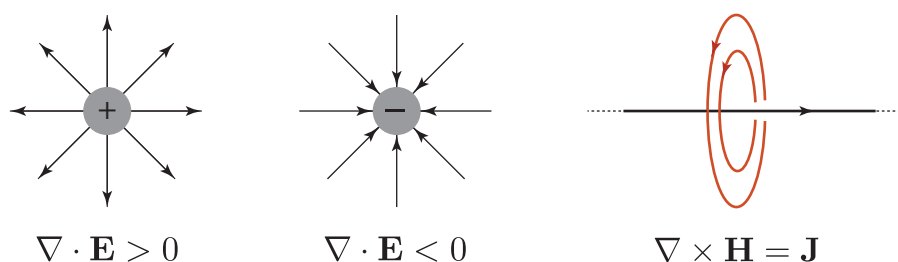


FIGURE 20.1: Electrostatic Fields.

Gauss's law  $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$  says that electric field lines (vectors) due to *static* charges originate at points in space where there are +ve charges, and terminate at negative charges, as indicated in Figure 20.1. Vectors originating from a point in space have a positive divergence. This relation is also called the Poisson equation in semiconductor device physics, and if the charge is zero, it goes by the name of Laplace equation. Gauss's law for magnetic fields tells us that magnetic field lines  $\mathbf{B}$  have no beginnings and no ends: unlike static electric field lines, they close on themselves.

Note that for electrostatics and magnetostatics, we put  $\partial(\dots)/\partial t \rightarrow 0$ , to obtain the static magnetic field relation  $\nabla \times \mathbf{H} = \mathbf{J}$ . The magnetic field lines curl around a wire carrying a dc current, as shown in Figure 20.1. Electrostatic phenomena such as electric fields in the presence of static charge such as p-n junctions, transistors, and optical devices in equilibrium, and magnetostatic phenomena such as magnetic fields near wires carrying dc currents are covered by the condition  $\partial(\dots)/\partial t \rightarrow 0$ , and electric and magnetic fields are *decoupled*. This means a static charge produces just electric fields and no magnetic fields. A static current (composed of charges moving at a *constant* velocity) produces a magnetic field, but no electric field.

Since in electrostatics,  $\nabla \times \mathbf{E} = 0$ , the static electric field vector can be expressed as the gradient of a scalar potential  $\mathbf{E} = -\nabla\Phi$  because  $\nabla \times (\nabla\Phi) = 0$  is an identity.  $\Phi$  is then the scalar electric potential. However, the same cannot be done for the magnetic field vector even in static conditions, because  $\nabla \times \mathbf{H} = \mathbf{J} \neq 0$ . However, the magnetic field can be written as the *curl* of another vector field  $\mathbf{B} = \nabla \times \mathbf{A}$ , where  $\mathbf{A}$  is called the magnetic *vector potential*. Hence from the Maxwell equations,  $\mathbf{E} = -d\mathbf{A}/dt$ .

Faraday's law says that a time-varying magnetic field creates an electric field. The electric field lines thus produced 'curl' around the magnetic field lines. Ampere's law says that a magnetic field intensity  $\mathbf{H}$  may be produced not just by a conductor carrying current  $\mathbf{J}$ , but also by a time-varying electric field in the form of the displacement current  $\partial\mathbf{D}/\partial t$ . The original Ampere's law did not have the displacement current. Maxwell realized that without it, the four constitutive equations would violate current continuity relations. To illustrate, without the displacement current term,  $\nabla \times \mathbf{H} = \mathbf{J}$ , and taking the divergence of both sides, we get  $\nabla \cdot \nabla \times \mathbf{H} = \nabla \cdot \mathbf{J} = 0$  because the divergence of curl of any vector field is zero. But the continuity equation requires

$$\nabla \cdot \mathbf{J} = -\partial\rho/\partial t, \quad \text{Continuity Equation} \quad (20.2)$$

which is necessary for the conservation of charge. With the introduction of the displacement current term, Maxwell resolved this conflict:  $\nabla \cdot \mathbf{J} = -\nabla \cdot \frac{\partial\mathbf{D}}{\partial t} = -\frac{\partial}{\partial t}(\nabla \cdot \mathbf{D}) = -\frac{\partial\rho}{\partial t}$ , which connects to Gauss's law.

### 20.3 Light emerges from Maxwell's equations

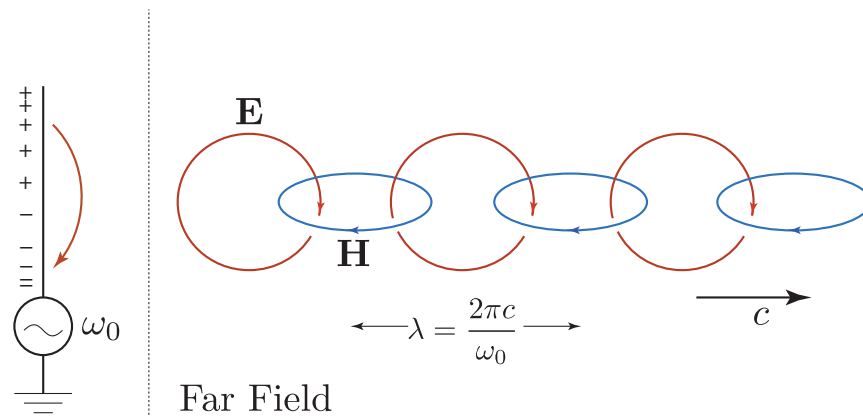


FIGURE 20.2: Antenna producing an electromagnetic wave.

The displacement current term is the crucial link between electricity and magnetism, and leads to the existence of light as an electromagnetic wave. Let's first look at this feature qualitatively. Figure 20.2 shows a metal wire connected to an *ac* voltage source. The battery shoves electrons back and forth from the ground into the wire, causing a charge-density wave as shown schematically. Note that the charge density in the wire is changing *continuously* in time and space. The frequency is  $\omega_0$ . As a result of charge pileups, electric field lines emerge from +ve charges and terminate on -ve charges. This electric field is changing in space and time as well, leading to non-zero  $\nabla \times \mathbf{E}$  and  $\partial \mathbf{E} / \partial t$ . The time-varying electric field *creates* a time-varying magnetic field  $\mathbf{H}$  because of displacement current. The time-varying magnetic field creates a time-varying electric field by Faraday's law. Far from the antenna, the fields detach from the source antenna and become *self-sustaining*: the time-varying  $\mathbf{E}$  creates  $\mathbf{H}$ , and vice versa. An electromagnetic wave is thus born; the oscillations of electric and magnetic fields move at the speed of light  $c$ . For an antenna radiating at a frequency  $\omega_0$ , the wavelength is  $\lambda = 2\pi c / \omega_0$ . That the wave is self-sustaining is the most fascinating feature of light. If at some time the battery was switched off, the far field wave continues to propagate *forever*, unless it encounters charges again. That of course is how light from the most distant galaxies and supernovae reach our antennas and telescopes, propagating through 'light years' in the vacuum of space, sustaining the oscillations<sup>2</sup>.

Now let's make this observation mathematically rigorous. Consider a region in space with no charges ( $\nabla \cdot \mathbf{D} = \rho = 0 = \nabla \cdot \mathbf{E}$ ) and no currents  $\mathbf{J} = 0$ . Take the curl of Faraday's equation to obtain  $\nabla \times \nabla \times \mathbf{E} = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\frac{\partial}{\partial t}(\nabla \times \mathbf{B}) = -\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{E}$ , where we make use of Ampere's law. Since in a source-free region  $\nabla \cdot \mathbf{E} = 0$ , we get the *wave equations*

<sup>2</sup>Boltzmann wrote "... was it a God who wrote these lines ..." in connection to "Let there be light".

$$\boxed{\begin{aligned} (\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}) \mathbf{E} &= 0, & \text{Wave Equations} \\ (\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}) \mathbf{B} &= 0. \end{aligned}} \quad (20.3)$$

Note that the wave equation states that the electric field and magnetic field oscillate both in space and time. The ratio of oscillations in space (captured by  $\nabla^2$ ) and oscillations in time (captured by  $\frac{\partial^2}{\partial t^2}$ ) is the speed at which the wave moves, and it is  $c = 1/\sqrt{\mu_0\epsilon_0}$ . The number is exactly equal to the experimentally measured speed of light, which solidifies the connection that light is an electromagnetic wave. We note that just like the solution to Dirac's equation in quantum mechanics is the electron, the solution of Maxwell's wave equation is light (or photons). Thus one can say that light has 'emerged' from the solution of Maxwell equations.

However, we must be cautious in calling the wave equation above representing light *alone*. Consider a generic wave equation  $(\nabla^2 - \frac{1}{v^2} \frac{\partial^2}{\partial t^2})f(r, t) = 0$ . This wave moves at a speed  $v$ . We can create a sound wave, *and* a water wave that moves at the same speed  $v$ , and  $f(r, t)$  will represent distinct physical phenomena. If a cheetah runs as fast as a car, they are not the same object!

Consider a generic vector field of the type  $\mathbf{V}(\mathbf{r}, t) = V_0 e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \hat{\eta}$ , where  $\hat{\eta}$  is the direction of the vector. This field will satisfy the wave equations 20.3 if  $\omega = c|\mathbf{k}|$ , as may be verified by substitution. This requirement is the first constraint on the nature of electromagnetic waves. The second stringent constraint is that the field must satisfy Gauss's laws  $\nabla \cdot \mathbf{E} = 0$  and  $\nabla \cdot \mathbf{B} = 0$  for free space. In other words, electric and magnetic vector fields are a special class of vector fields. Their special nature is elevated by the physical observation that *no other wave can move at the speed of light*. Einstein's theory of relativity proves that the speed of light is absolute, and unique for electromagnetic waves: every other kind of wave falls short of the speed of light. Thus, Maxwell's wave equation *uniquely* represents light, self-sustaining oscillating electric and magnetic fields.

## 20.4 Maxwell's equations in $(\mathbf{k}, \omega)$ space

Consider an electromagnetic wave of a fixed frequency  $\omega$ . Since  $\mathbf{E}, \mathbf{B} \propto e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}$ , we make two observations. Time derivatives of Faraday and Ampere's laws give  $\frac{\partial}{\partial t} e^{-i\omega t} = -i\omega e^{-i\omega t}$ , which means we can replace  $\frac{\partial}{\partial t} \rightarrow -i\omega$ ,  $\frac{\partial^2}{\partial t^2} \rightarrow (-i\omega)^2$ , and so on. Similarly, the vector operators div and curl act on the  $e^{i\mathbf{k}\cdot\mathbf{r}}$  part only, giving  $\nabla \cdot (e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\eta}) = i\mathbf{k} \cdot (e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\eta})$  and  $\nabla \times (e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\eta}) = i\mathbf{k} \times (e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\eta})$ . These relations may be verified by straightforward substitution. Thus, we can replace  $\nabla \rightarrow i\mathbf{k}$ . With these observations, Maxwell equations in free-space become

$$\begin{aligned}
\mathbf{k} \cdot \mathbf{E} &= 0, \\
\mathbf{k} \cdot \mathbf{B} &= 0, \\
\mathbf{k} \times \mathbf{E} &= \omega \mathbf{B}, \\
\mathbf{k} \times \mathbf{B} &= -\frac{\omega}{c^2} \mathbf{E}.
\end{aligned}
\tag{20.4}$$

Note that we have converted Maxwell's equations in real space and time  $(\mathbf{r}, t)$  to 'Fourier' space  $(\mathbf{k}, \omega)$  in this process. Just as in Fourier analysis where we decompose a function into its spectral components, light of a particular  $\mathbf{k}$  and corresponding frequency  $\omega = c|\mathbf{k}|$  is spectrally pure, and forms the 'sine' and 'cosine' bases. Any mixture of light is a linear combination of these spectrally pure components: for example white light is composed of multiple wavelengths. Since  $\mathbf{B} = \nabla \times \mathbf{A}$ , we can write  $\mathbf{B} = i\mathbf{k} \times \mathbf{A}$ , and hence the magnitudes are related by  $B^2 = k^2 A^2 = (\frac{\omega}{c})^2 A^2$ . The energy content in a region in space of volume  $\Omega$  that houses electric and magnetic fields of frequency  $\omega$  is given by

$$H_{em}(\omega) = \Omega \cdot \left[ \frac{1}{2} \epsilon_0 E^2 + \frac{1}{2} \mu_0 H^2 \right] = \Omega \cdot \left[ \frac{1}{2} \epsilon_0 E^2 + \frac{1}{2} \epsilon_0 \omega^2 A^2 \right].
\tag{20.5}$$

If you have noticed a remarkable similarity between the expression for energy of an electromagnetic field with that of a harmonic oscillator (from Chapter 3)  $H_{osc} = \frac{\hat{p}^2}{2m} + \frac{1}{2} m \omega^2 x^2$ , you are in luck. In Chapter 21, this analogy will enable us to fully quantize the electromagnetic field, resulting in a rich new insights.

Let us now investigate the properties of a spectrally pure, or 'monochromatic' component of the electromagnetic wave. From equations 20.4, we note that  $\mathbf{k} \perp \mathbf{E} \perp \mathbf{B}$ , and the direction of  $\mathbf{k}$  is along  $\mathbf{E} \times \mathbf{B}$ . The simplest possibility is shown in Figure 20.3. If we align the  $x$ -axis along the electric field vector and the  $y$ -axis along the magnetic field vector, then the wave propagates along the +ve  $z$ -axis, i.e.,  $\mathbf{k} = k\hat{z}$ . The electric field vectors lie in the  $x - z$  plane, and may be written as  $\mathbf{E}(\mathbf{r}, t) = E_0 e^{i(kz - \omega t)} \hat{x}$ , which is a plane wave. For a plane wave, nothing changes along the planes perpendicular to the direction of propagation, so the  $\mathbf{E}$  field is the same at all  $x - y$  planes:  $\mathbf{E}(x, y, z) = \mathbf{E}(0, 0, z)$ .

From Faraday's law,  $\mathbf{B} = \mathbf{k} \times \mathbf{E} / \omega$ , and the magnetic field vectors  $\mathbf{B}(\mathbf{r}, t) = \frac{E_0}{c} e^{i(kz - \omega t)} \hat{y}$  lie in the  $y - z$  plane. Note that here we use  $\omega = ck$  and  $k = k_z$ . The amplitudes of the electric and magnetic fields are thus related by  $E_0 = cB_0$ , and the relation to magnetic field *intensity*  $\mathbf{H} = \mathbf{B} / \mu_0$  is  $E_0 = c\mu_0 H_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} H_0 = \eta_0 H_0$ . Since  $E_0$  has units V/m and  $H_0$  has units A/m,  $\eta$  has units of V/A or Ohms.  $\eta_0$  is called the impedance of free space; it has a value  $\eta_0 \approx 377\Omega$ .

The direction of propagation of this wave is *always* perpendicular to the electric and magnetic field vectors and given by the right hand rule. Since the field vectors lie on

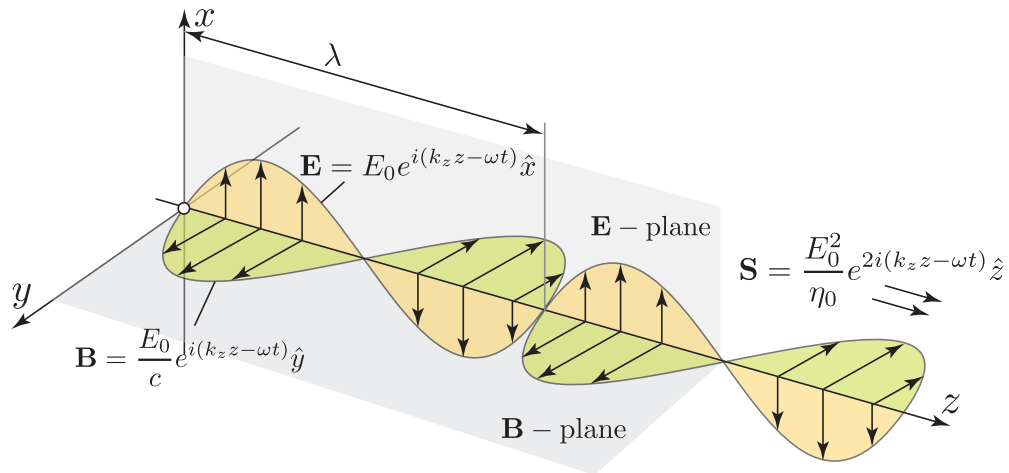


FIGURE 20.3: Electromagnetic wave.

well-defined planes, this type of electromagnetic wave is called *plane-polarized*. In case there was a phase difference between the electric and magnetic fields, the electric and magnetic field vectors will rotate in the  $x - y$  planes as the wave propagates, and the wave would then be called circularly or *elliptically* polarized, depending upon the phase difference.

For the monochromatic wave, Maxwell's wave equation becomes  $(|\mathbf{k}|^2 - (\frac{\omega}{c})^2)\mathbf{E} = 0$ . For non-zero  $\mathbf{E}$ ,  $\omega = c|\mathbf{k}| = ck$ . The electromagnetic field carries energy in the +ve  $z$ -direction. The instantaneous power carried by the wave is given by the Poynting vector  $\mathbf{S}(\mathbf{r}, t) = \mathbf{E} \times \mathbf{H} = \frac{E_0^2}{\eta_0} e^{i(kz - \omega t)} \hat{z}$ . The units are in Watts/m<sup>2</sup>. Typically we are interested in the time-averaged power density, which is given by

$$\mathbf{S} = \langle \mathbf{S}(\mathbf{r}, t) \rangle = \frac{1}{2} \text{Re}[\mathbf{E} \times \mathbf{H}^*] = \frac{E_0^2}{2\eta} \hat{z} = \frac{\eta}{2} H_0^2 \hat{z}, \quad (20.6)$$

where  $\hat{z}$  is the direction of propagation of the wave. In later chapters, the energy carried by a monochromatic wave will for the starting point to understand the interaction of light with matter. In the next chapter, we will discuss how the energy carried by an electromagnetic wave as described by Equation 20.6 actually appears not in continuous quantities, but in quantum packets. Before we do that, we briefly discuss the classical picture of light interacting with material media.

## 20.5 Maxwell's equations in material media

How does light interact with a material medium? Running the video of the process of the *creation* of light in Figure 20.2 backwards, we can say that when an electromagnetic

wave hits a metal wire, the electric field will slosh electrons in the wire back and forth generating an ac current. That is the principle of operation of a receiving antenna. What happens when the material does not have freely conducting electrons like a metal? For example, in a dielectric some electrons are tightly bound to atomic nuclei (core electrons), and others participate in forming chemical bonds with nearest neighbor atoms. The electric field of the electromagnetic wave will deform the electron clouds that are most 'flexible' and 'polarize' them. Before the external field was applied, the centroid of the negative charge from the electron clouds and the positive nuclei exactly coincided in space. When the electron cloud is deformed, the centroids do not coincide any more, and a net dipole is formed, as shown in Figure 20.4. The electric field of light primarily interacts with electrons that are most loosely bound and deformable; protons in the nucleus are far heavier, and held strongly in place in a solid medium. Let us give these qualitative observations a quantitative basis.

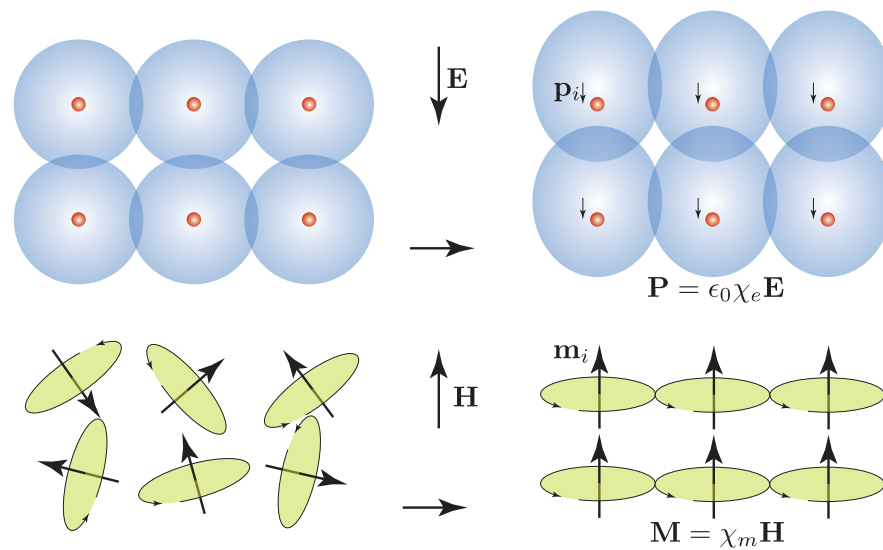


FIGURE 20.4: Dielectric and Magnetic materials. Orientation of electric and magnetic dipoles by external fields, leading to electric and magnetic susceptibilities.

The displacement vector in free space is  $\mathbf{D} = \epsilon_0 \mathbf{E}$ . In the presence of a dielectric, it has an additional contribution  $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$ , where  $\mathbf{P}$  is the polarization of the dielectric. The classical picture of polarization is an electric dipole  $\mathbf{p}_i = qd_i \hat{n}$  in every unit cell of the solid. This dipole has zero magnitude in the absence of the external field<sup>3</sup>. The electric field of light stretches the electron cloud along it, forming dipoles along itself. Thus,  $\mathbf{p}_i$  points along  $\mathbf{E}$ . The net polarization<sup>4</sup> is the volume-averaged

<sup>3</sup>Except in materials that have spontaneous, piezoelectric, or ferroelectric polarization.

<sup>4</sup>This classical picture of polarization is not consistent with quantum mechanics. The quantum theory of polarization requires the concept of Berry phases, which is the subject of Chapter 54.



dipole density  $\mathbf{P} = \frac{1}{V} \sum_V \mathbf{p}_i$ . Based on the material properties of the dielectric, we absorb all microscopic details into one parameter by writing

$$\mathbf{P} = \epsilon_0 \chi_e \mathbf{E}, \quad (20.7)$$

where the parameter  $\chi_e$  is referred to as the electric susceptibility of the solid. With this definition, the displacement vector becomes

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \epsilon_0 \chi_e \mathbf{E} = \epsilon_0 \underbrace{(1 + \chi_e)}_{\epsilon_r} \mathbf{E} = \epsilon \mathbf{E}, \quad (20.8)$$

where the dielectric property of the material is captured by the modified dielectric constant  $\epsilon = \epsilon_0 \epsilon_r = \epsilon_0 (1 + \chi_e)$ . The *relative* dielectric constant is 1 plus the electric susceptibility of the material. Clearly the relative dielectric constant of vacuum is 1 since there are no atoms to polarize and nothing is 'susceptible'.

In exactly the same way, if the material is magnetically polarizable, then  $\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M})$ , where  $\mathbf{M}$  is the magnetization vector. If there are tiny magnetic dipoles  $\mathbf{m}_i = IA\hat{n}$  formed by circular loops carrying current  $I$  in area  $A$  in the material medium (see Figure 20.4), the macroscopic magnetization is given by  $\mathbf{M} = \frac{1}{V} \sum_V \mathbf{m}_i = \chi_m \mathbf{H}$ , which leads to the relation

$$\mathbf{B} = \mu_0 (\mathbf{H} + \chi_m \mathbf{H}) = \mu_0 \underbrace{(1 + \chi_m)}_{\mu_r} \mathbf{H} = \mu \mathbf{H}, \quad (20.9)$$

With these changes, the original Maxwell equations remain the same, but now  $\mathbf{D} = \epsilon \mathbf{E}$  and  $\mathbf{B} = \mu \mathbf{H}$ , so we make the corresponding changes  $\epsilon_0 \rightarrow \epsilon = \epsilon_0 \epsilon_r$  and  $\mu_0 \rightarrow \mu = \mu_0 \mu_r$  everywhere. For example, the speed of light in a material medium then becomes  $v = \frac{1}{\sqrt{\mu\epsilon}} = \frac{c}{\sqrt{\epsilon_r \mu_r}}$ . If the material is non-magnetic, then  $\mu_r = 1$ , and  $v = \frac{c}{\sqrt{\epsilon_r}} = \frac{c}{n}$ , where  $n = \sqrt{\epsilon_r}$  is called the refractive index of the material. Thus light travels slower in a material medium than in free space. Similarly, the wave impedance becomes  $\eta_0 \rightarrow \eta = \sqrt{\frac{\mu}{\epsilon}} = \frac{\eta_0}{n}$  where the right equality holds for a non-magnetic medium.

If the material medium is conductive, or can absorb the light through electronic transitions, then the phenomena of *absorption* and corresponding attenuation of the light is captured by introducing an imaginary component to the dielectric constant,  $\epsilon \rightarrow \epsilon_R + i\epsilon_I$ . This leads to an imaginary component of the propagation vector  $\mathbf{k}$ , which leads to attenuation. We will see in Chapters 26 and 27 how we can calculate the absorption coefficients from quantum mechanics.

Electric and magnetic field lines may cross interfaces of different material media. Then, the Maxwell equations provide rules for tracking the magnitudes of the tangential and perpendicular components. These **boundary conditions** are given by

$$\begin{array}{l} \mathbf{E}_{1t} - \mathbf{E}_{2t} = 0, \\ \mathbf{H}_{1t} - \mathbf{H}_{2t} = \mathbf{J}_s \times \hat{\mathbf{n}}, \\ D_{1n} - D_{2n} = \rho_s, \\ B_{1n} - B_{2n} = 0. \end{array} \quad (20.10)$$

In words, the boundary condition relations say that the tangential component of the electric field  $\mathbf{E}_t$  is *always* continuous across an interface, but the normal component is discontinuous if there are charges at the interface. If there are no *free* charges at the interface ( $\rho_s = 0$ ),  $\epsilon_1 E_{1n} = \epsilon_2 E_{2n}$ , implying the normal component of the electric field is larger in the material with a smaller dielectric constant. This feature is used in Si MOSFETs, where much of the electric field drops across an oxide layer rather than in the semiconductor which has a higher dielectric constant. Similarly, the normal component of the magnetic field is always continuous across an interface, whereas the tangential component can change if there is a surface current flowing at the interface of the two media.

The force in Newtons on a particle of charge  $q$  in the presence of an electric and magnetic field is given by the Lorentz equation

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (20.11)$$

Since the energy of the charged particle changes as  $W = \int \mathbf{F} \cdot d\mathbf{r}$ , the rate of change of energy is  $\mathbf{F} \cdot \mathbf{v} = q\mathbf{E} \cdot \mathbf{v}$ , which is the power delivered to the charged particle by the fields. Note that a static magnetic field cannot deliver power since  $\mathbf{v} \times \mathbf{B} \cdot \mathbf{v} = 0$ . Thus a time-independent magnetic field cannot change the energy of a charged particle. But a time-dependent magnetic field creates an electric field, which can.

When a point charge is *accelerated* with acceleration  $a$ , it radiates electromagnetic waves. Radiation travels at the speed of light. So the electric and magnetic fields at a point far from the charge are determined by a *retarded* response. Using retarded potentials, or more intuitive approaches<sup>5</sup>, one obtains that the radiated electric field goes as

<sup>5</sup>An intuitive picture for radiation by an accelerating charge was first given by J. J. Thomson, the discoverer of the electron.

$$\mathbf{E}_r = \left( \frac{qa}{4\pi\epsilon_0 c^2} \right) \frac{\sin\theta}{r} \hat{\theta}, \quad (20.12)$$

expressed in spherical coordinates with the charge at the origin, and accelerating along the  $x$ -axis. The radiated magnetic field  $\mathbf{H}_r$  curls in the  $\hat{\phi}$  direction and has a magnitude  $|\mathbf{E}_r|/\eta_0$ . The radiated power is obtained by the Poynting vector  $\mathbf{S} = \mathbf{E} \times \mathbf{H}$  as

$$\mathbf{S} = \left( \frac{\mu_0 q^2 a^2}{16\pi^2 c^2} \right) \left( \frac{\sin\theta}{r} \right)^2 \hat{r}, \quad (20.13)$$

Note that unlike static charges or currents that fall as  $1/r^2$  away from the source, the *radiated*  $\mathbf{E}$  and  $\mathbf{H}$  fields fall as  $1/r$ . If they didn't, the net power radiated very far from the source will go to zero since  $\oint \mathbf{S} \cdot d\mathbf{A} \sim S(r)4\pi r^2 \rightarrow 0$ . Integrating the power over the angular coordinates results in the famous **Larmor Formula** for the net electromagnetic power in Watts radiated by an accelerating charge:

$$\boxed{P = \frac{\mu_0 q^2 a^2}{6\pi c}} \quad (20.14)$$

## 20.6 Need for a quantum theory of light

Classical electromagnetism contained in Maxwell's equations can explain a remarkably large number of experimentally observed phenomena, but not all. We discussed in the beginning of this chapter that radiation of electromagnetic waves can be created in an antenna, which in its most simple form is a conducting wire in which electrons are sloshed back and forth. The *collective* acceleration, coupled with the Larmor formula can explain radiation from a vast number of sources of electromagnetic radiation.

By the turn of the 20th century, improvements in spectroscopic equipment had helped resolve what was originally thought as broadband (many frequencies  $\omega$ ) radiation into the purest spectral components. It was observed that different gases had different spectral signatures. The most famous among them were the spectral features of the hydrogen atom, then known as the hydrogen gas. There is nothing *collective* about hydrogen gas, since it is not a conductor and there are not much electrons to slosh around as a metal. The classical theory for radiation proved difficult to apply to explain the spectral features. Classical electromagnetism could not explain the photoelectric effect, and the spectrum of blackbody radiation either. The search for an explanation led to the quantum theory of light, which is the subject of the next chapter.

## Chapter 26

# Optical Transitions in Bulk Semiconductors

### 26.1 Introduction

In this chapter, we explore fundamental optical transitions in bulk 3-dimensional semiconductors. We approach the topic by first investigating the optical absorption spectrum. The spectrum will direct us to a rich range of electron state transitions affected by the electron-photon interaction. Then, we explore the most important of these transitions: interband (valence  $\rightarrow$  conduction) transitions in more detail. We derive expressions for the equilibrium interband absorption coefficient  $\alpha_0(\hbar\omega)$  for bulk semiconductors. With the understanding of the physics of optical absorption, in the next chapter we extend the concept to non-equilibrium situations to explain optical emission, optical gain, inversion, and lasing conditions. The key to understanding these concepts is a clear quantum-mechanical picture of optical transitions, and the role of non-equilibrium conditions. We begin with the fundamental quantum-mechanical optical transitions by recalling the electron-photon Hamiltonian.

### 26.2 Electron-photon matrix elements for semiconductors

In Chapter 21, we justified that the Hamiltonian

$$H = \frac{1}{2m_0}(\hat{\mathbf{p}} + e\mathbf{A})^2 + V(\mathbf{r}) \quad (26.1)$$

captures the interaction of electrons with light. We used this Hamiltonian in Chapter 25 to investigate the interaction of light with atoms, and explained the optical spectra of atoms. We found that the spectra of atoms are typically very sharp because of the discrete energy eigenvalues of electrons. Here, we apply the same idea to bulk semiconductors, in which the energy eigenvalues form bands separated by energy gaps.

We recall that the electromagnetic wave enters the Hamiltonian via the magnetic vector potential  $\mathbf{A}$ , which is related to the electric field via

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial}{\partial t} \mathbf{B}(\mathbf{r}, t) \quad \xrightarrow{\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t)} \quad \mathbf{E}(\mathbf{r}, t) = -\frac{\partial}{\partial t} \mathbf{A}(\mathbf{r}, t), \quad (26.2)$$

and we work in the Coulomb gauge

$$\nabla \cdot \mathbf{A} = 0. \quad (26.3)$$

This enables the vector potential  $\mathbf{A}$  to commute with the momentum operator  $\hat{\mathbf{p}}$

$$[\hat{\mathbf{p}}, \mathbf{A}] = 0 \rightarrow \hat{\mathbf{p}} \cdot \mathbf{A} = \mathbf{A} \cdot \hat{\mathbf{p}}, \quad (26.4)$$

which leads to the electron-photon Hamiltonian

$$H = \underbrace{\left[ \frac{\hat{p}^2}{2m_0} + V(\mathbf{r}) \right]}_{\hat{H}_0} + \underbrace{\frac{e}{m_0} \mathbf{A} \cdot \hat{\mathbf{p}}}_W + \underbrace{\frac{e^2 A^2}{2m_0}}_{\text{neglect}}. \quad (26.5)$$

We have written out the Hamiltonian in terms of the electron Hamiltonian  $\hat{H}_0$ , and the ‘perturbation’ term seen by the electron due to the electromagnetic wave. For an electron in a semiconductor crystal, the potential energy term in the unperturbed Hamiltonian is the periodic crystal potential  $V(\mathbf{r} + \mathbf{a}_0) = V(\mathbf{r})$ , where  $\mathbf{a}_0$  is a lattice constant. We neglect the perturbation term that goes as the square of the magnetic vector potential for ‘weak’ intensities of light. This is justified when the condition  $|e\mathbf{A}| \ll |\mathbf{p}| \sim \hbar\pi/a_0$  is met; in other words, we neglect the term  $\frac{e^2 A^2}{2m_0}$  w.r.t.  $\frac{\hat{p}^2}{2m_0}$ . The net Hamiltonian we retain then has the electron experiencing a perturbation

$$\boxed{\hat{W} = \frac{e}{m_0} \mathbf{A} \cdot \hat{\mathbf{p}}} \quad (26.6)$$

due to its interaction with light. The magnetic vector potential for an EMag wave is of the form<sup>1</sup>

$$\mathbf{A}(\mathbf{r}, t) = \hat{e}A_0 \cos(\mathbf{k}_{\text{op}} \cdot \mathbf{r} - \omega t) \quad (26.7)$$

$$= \hat{e} \frac{A_0}{2} e^{+i\mathbf{k}_{\text{op}} \cdot \mathbf{r} - i\omega t} + \hat{e} \frac{A_0}{2} e^{-i\mathbf{k}_{\text{op}} \cdot \mathbf{r} + i\omega t}, \quad (26.8)$$

where  $\omega$  is the angular frequency of the EMag wave,  $\hat{e}$  is the unit vector along the electric (and vector potential) field, and  $\mathbf{k}_{\text{op}}$  is the propagation wave vector of magnitude  $2\pi/\lambda$ . The electron-photon interaction Hamiltonian is then given by

$$\hat{W}(\mathbf{r}, t) = \frac{e}{m_0} \mathbf{A} \cdot \hat{\mathbf{p}} \quad (26.9)$$

$$= \hat{W}(\mathbf{r})e^{-i\omega t} + \hat{W}^+(\mathbf{r})e^{+i\omega t} \quad (26.10)$$

$$\hat{W}(\mathbf{r}) = \frac{eA_0 e^{i\mathbf{k}_{\text{op}} \cdot \mathbf{r}}}{2m_0} \hat{e} \cdot \hat{\mathbf{p}} \quad (26.11)$$

$$\hat{W}^+(\mathbf{r}) = \frac{eA_0 e^{-i\mathbf{k}_{\text{op}} \cdot \mathbf{r}}}{2m_0} \hat{e} \cdot \hat{\mathbf{p}} \quad (26.12)$$

The electron-photon matrix elements for bulk semiconductors are thus of the form  $\langle \mathbf{k}_c | \hat{W} | \mathbf{k}_v \rangle$  and  $\langle \mathbf{k}_c | \hat{W}^+ | \mathbf{k}_v \rangle$ , where the *unperturbed* electron states  $|\mathbf{k}_c\rangle$  and  $|\mathbf{k}_v\rangle$  are solutions of the unperturbed Hamiltonian  $\hat{H}_0 = \frac{\hat{p}^2}{2m_0} + V(\mathbf{r})$ . But this is precisely what we discussed in chapters 16 and 17 for semiconductors. The electron states in the valence and conduction bands in the effective mass approximation are  $\psi_c(\mathbf{r}) = \langle \mathbf{r} | \mathbf{k}_c \rangle = \left( \frac{e^{i\mathbf{k}_c \cdot \mathbf{r}}}{\sqrt{V}} \right) u_c(\mathbf{r})$  for bulk semiconductors. The term in the round bracket is a slowly varying envelope function, and  $u_c(\mathbf{r})$  is the periodic part of the Bloch function. The effective mass approximation transforms the unperturbed electronic Hamiltonian into the much simpler form  $\frac{\hat{p}^2}{2m_0} + V(\mathbf{r}) \rightarrow \frac{\hat{p}^2}{2m^*}$ , and the corresponding effective-mass Schrodinger equation is  $\frac{\hat{p}^2}{2m^*} \psi_c(\mathbf{r}) = (E - E_c) \psi_c(\mathbf{r})$ . We will work in this effective-mass theory. The advantage of working in the effective-mass theory is that the light-matter interaction matrix elements for electrons confined in low-dimensional structures such as quantum wells, wires, or dots follows in a simple way from the bulk results. We will need the matrix elements shortly to explain the absorption spectra of bulk semiconductors, which we discuss next.

<sup>1</sup>This approach of treating the electron-photon interaction is *semi-classical*, justified for *classical* electromagnetic fields when the number of photons is much larger than unity. It is semi-classical because electrons receive the full quantum treatment, but we neglect the quantization of the electromagnetic field, treating it as a classical infinite source or sink of energy, albeit in quantum packets of  $\hbar\omega$ . The electromagnetic field will be quantized in Chapters 46 and beyond in this book.

### 26.3 The absorption spectrum of bulk semiconductors

We learn early of the Beer-Lambert ‘law’, which states that if light of intensity  $I_0$  is incident on a material that absorbs, the intensity will decay inside the material as  $I(z) = I_0 e^{-\alpha z}$ . Here  $\alpha$  is the absorption coefficient, in units of inverse length. Typically the unit used for  $\alpha$  is  $\text{cm}^{-1}$ . Let us consider the following experiment: take a piece of bulk semiconductor, say GaAs or GaN, and using a tunable light source, measure  $\alpha$  as a function of the photon energy  $\hbar\omega$ . Then we obtain the absorption spectrum  $\alpha(\hbar\omega)$ . The absorption spectrum of most semiconductors looks like what is shown in the schematic Figure 26.1.

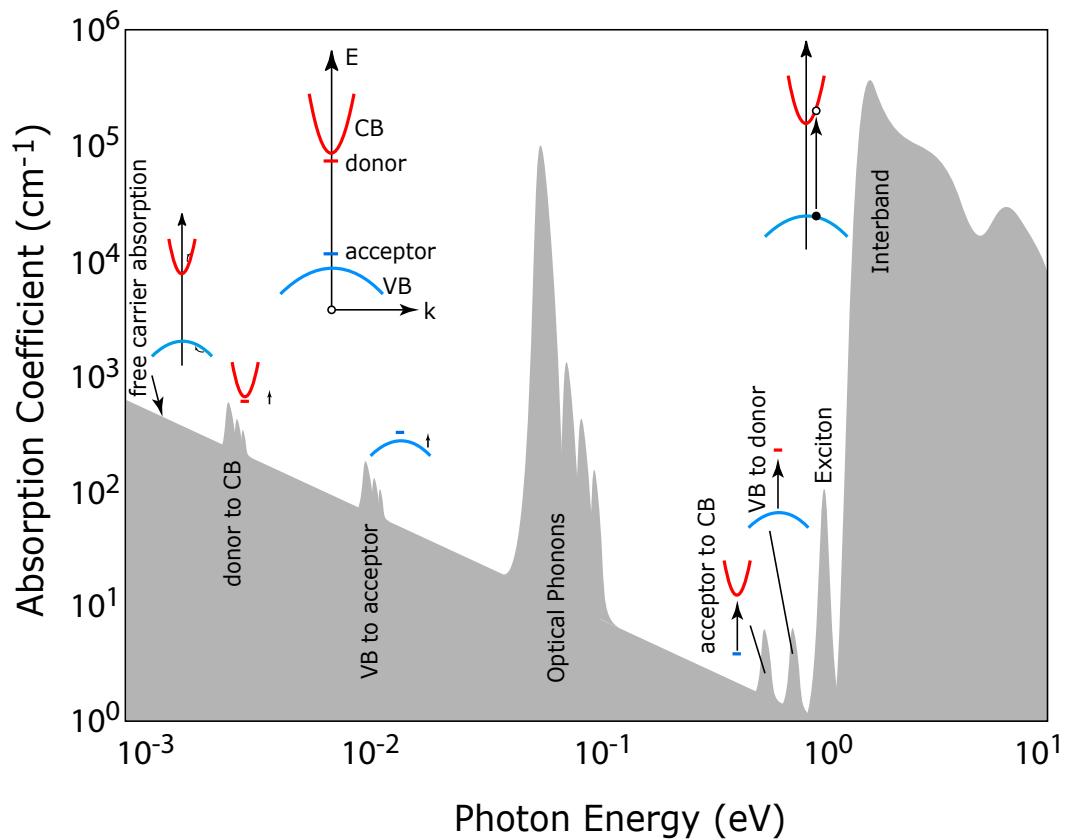


FIGURE 26.1: Schematic absorption spectrum  $\alpha(\hbar\omega)$  of bulk semiconductors. The insets depict various state transitions upon absorption of photons.

The inset of Figure 26.1 indicates the electron bandstructure of the bulk semiconductor, including states corresponding to donor and acceptor dopants. The transitions between electron states caused by photon absorption are indicated. The floor of the absorption spectrum is due to intraband transitions caused by the absorption of low energy ( $\sim$  few meV) photons by free carriers. Transitions between dopant and band states are shown, in addition to the below-bandgap excitonic transition. Such optical measurements provide a sensitive experimental determination of dopant and excitonic energies with respect

to the fundamental band edge energies. Photons can excite *mechanical* vibrations of the bulk semiconductor crystal by creating optical phonons: the absorption peak for this process is rather strong, and forms the basis of Raman spectroscopy. By far, the strongest absorption occurs for interband transitions, which is the focus of this chapter.

The absorption spectrum is quantitatively defined as

$$\alpha(\hbar\omega) = \frac{\text{Number of photons absorbed per unit volume per second}}{\text{Number of photons incident per unit area per second}} = \frac{R(\hbar\omega)}{N_{ph}(\hbar\omega)}. \quad (26.13)$$

In the next section we derive an expression for the denominator  $N_{ph}(\hbar\omega)$ , and in the following section we deal with the numerator  $R(\hbar\omega)$ .

## 26.4 The number of photons in light

Consider a monochromatic EMag wave of frequency  $\omega$  and corresponding wavevector  $\mathbf{k}_{op} = \frac{2\pi}{\lambda}\hat{n}$ . For a plane wave, the magnetic vector potential is

$$\mathbf{A}(\mathbf{r}, t) = \hat{e}A_0 \cos(\mathbf{k}_{op} \cdot \mathbf{r} - \omega t), \quad (26.14)$$

from where the electric field is obtained by using

$$\mathbf{E}(\mathbf{r}, t) = -\frac{\partial}{\partial t}\mathbf{A}(\mathbf{r}, t) \quad (26.15)$$

$$= -\hat{e}\omega A_0 \sin(\mathbf{k}_{op} \cdot \mathbf{r} - \omega t), \quad (26.16)$$

and the magnetic field intensity is

$$\mathbf{H}(\mathbf{r}, t) = \frac{1}{\mu}\nabla \times \mathbf{A}(\mathbf{r}, t) \quad (26.17)$$

$$= -\frac{1}{\mu}\mathbf{k}_{op} \times \hat{e}A_0 \sin(\mathbf{k}_{op} \cdot \mathbf{r} - \omega t). \quad (26.18)$$

Here we have used  $\nabla \times (\dots) \equiv -\mathbf{k}_{op} \times (\dots)$  for plane waves, as described in Chapter 20. Then, the energy carried by the plane wave per unit area per unit time is given by the Poynting vector



$$\mathbf{S}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, t) \times \mathbf{H}(\mathbf{r}, t) \quad (26.19)$$

$$= \mathbf{k}_{\text{op}} \frac{\omega A_0^2}{\mu} \sin^2(\mathbf{k}_{\text{op}} \cdot \mathbf{r} - \omega t) \quad (26.20)$$

Where we use the identity  $\hat{e} \times \mathbf{k}_{\text{op}} \times \hat{e} = \mathbf{k}_{\text{op}}$ . Since the frequency of typical UV-visible-IR light is very high, we time-average the Poynting vector over a period to obtain

$$\langle \mathbf{S}(\mathbf{r}, t) \rangle = \frac{\omega A_0^2}{2\mu} \mathbf{k}_{\text{op}}, \quad (26.21)$$

and its magnitude is

$$S = |\langle \mathbf{S}(\mathbf{r}, t) \rangle| = \frac{\omega A_0^2}{2\mu} k_{\text{op}} = \frac{n_r c \epsilon_0 \omega^2 A_0^2}{2} = \frac{E_0^2}{2\eta} \quad (26.22)$$

where  $\mu = \mu_0$  and  $n_r = \sqrt{\mu_r \epsilon_r}$  is the refractive index of the media, in which the speed of light is  $c/n_r$ . Also note that  $E_0 = \omega A_0$ , and  $\eta = \sqrt{\mu/\epsilon}$  is the field impedance. This relation gives us a way to find the magnitude of the vector potential  $A_0$  if we know the power carried per unit area by the electromagnetic wave. Since energy in electromagnetic waves is carried in quantum packets (photons) of individual energy  $\hbar\omega$ , the *number of photons* that cross unit area per unit time is then given by

$$N_{ph}(\hbar\omega) = \frac{S}{\hbar\omega} = \frac{n_r c \epsilon_0 \omega^2 A_0^2}{2\hbar\omega} = \frac{E_0^2}{2\eta\hbar\omega}. \quad (26.23)$$

The *intensity* of light is proportional to the *square* of the electric (or magnetic) field amplitude, and thus the number of photons is a measure of the intensity of radiation. Equation 26.23 provides the denominator of the expression for absorption coefficient Equation 26.13. The numerator term is discussed in the next section.

## 26.5 Photon absorption rate in bulk semiconductors

To find the rate of photon absorption in the bulk semiconductor, we apply Fermi's golden rule derived in Chapter 24. We first note that the numerator of Equation 26.13 has units of number of photons absorbed per unit volume per second. Consider Figure 26.2. An electron in the valence band state  $|a\rangle$  absorbs a photon of energy  $\hbar\omega$  and transitions into state  $|a\rangle$  in the conduction band. Each such transition results in the *annihilation* of a

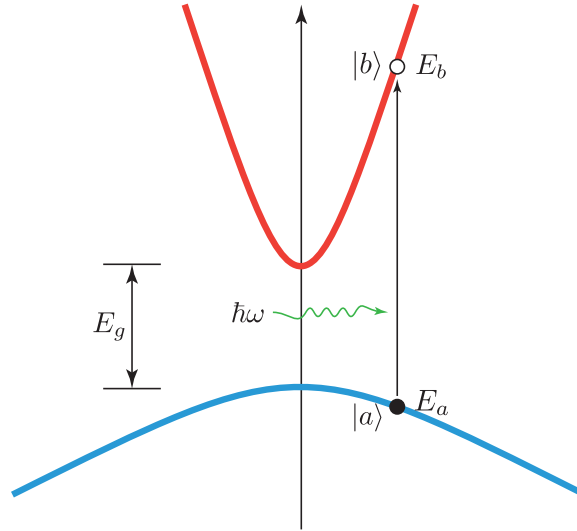


FIGURE 26.2: The absorption process of a single photon by interband transition.

photon from the EMag field. The rate at which this happens is given by Fermi's golden rule as

$$\frac{1}{\tau_{a \rightarrow b}} = \frac{2\pi}{\hbar} |\langle b | W(\mathbf{r}) | a \rangle|^2 \delta[E_b - (E_a + \hbar\omega)], \quad (26.24)$$

where  $\langle b | W(\mathbf{r}) | a \rangle$  is the perturbation matrix element, and the Dirac-delta function is a statement of energy conservation in the process. The reverse process of photon emission is also allowed, which results in the *creation* of a photon in the EMag field at the rate

$$\frac{1}{\tau_{b \rightarrow a}} = \frac{2\pi}{\hbar} |\langle a | W(\mathbf{r}) | b \rangle|^2 \delta[E_a - (E_b - \hbar\omega)], \quad (26.25)$$

which must be *subtracted* because an emission process makes a negative contribution to the number of photons absorbed. The above results are for the single states  $|a\rangle$  and  $|b\rangle$ . A semiconductor crystal has a large number of states in the respective bands, so let's sum the rates for all possible transitions, and divide it by the net volume  $V$  to obtain the absorption rate per unit volume (in  $\text{s}^{-1} \cdot \text{cm}^{-3}$ ). Add in the electron spin degeneracy  $g_s = 2$  for each  $\mathbf{k}$  state<sup>2</sup>. For the absorption process to occur, the lower state  $|a\rangle$  has to be occupied (probability =  $f_a$ ) and the higher state  $|b\rangle$  has to be empty (probability

<sup>2</sup>Photons carry an angular momentum of  $\pm\hbar$  depending upon their polarization. Therefore, the conservation of angular momentum couples specific spin states. Here we are considering light with photons of mixed polarization. Angular momentum conservation dictates which bands can be involved in the absorption or emission process, thus providing a way to selectively excite say the light hole, heavy hole, or split-off bands because they differ in their net angular momentum.

$= (1 - f_b)$ ), where  $f$ 's are the occupation functions. The net absorption rate per unit volume is then given by

$$R_{abs} = \frac{2}{V} \sum_{\mathbf{k}_a} \sum_{\mathbf{k}_b} \frac{2\pi}{\hbar} |W_{ba}|^2 \delta[E_b - (E_a + \hbar\omega)] f_a (1 - f_b), \quad (26.26)$$

and the net emission rate per unit volume is

$$R_{em} = \frac{2}{V} \sum_{\mathbf{k}_a} \sum_{\mathbf{k}_b} \frac{2\pi}{\hbar} |W_{ab}|^2 \delta[E_a - (E_b - \hbar\omega)] f_b (1 - f_a). \quad (26.27)$$

The summation runs over *all* valence band electron states  $\mathbf{k}_a$  and conduction band electron states  $\mathbf{k}_b$ , including those that do not meet the criteria  $E_b - E_a = \hbar\omega$ . The energy conservation requirement is automatically taken care of by the Dirac-delta functions. We note now that the Dirac-delta functions are the same for emission and absorption process because  $\delta[+x] = \delta[-x]$ ,  $|W_{ab}| = |W_{ba}|$ , and  $f_a(1 - f_b) - f_b(1 - f_a) = f_a - f_b$ . Therefore, the net photon *absorption* rate per unit volume is the difference

$$R(\hbar\omega) = R_{abs} - R_{em} = \frac{2}{V} \sum_{\mathbf{k}_a} \sum_{\mathbf{k}_b} \frac{2\pi}{\hbar} |W_{ab}|^2 \delta[E_b - (E_a + \hbar\omega)] \times (f_a - f_b) \quad (26.28)$$

To evaluate the sum over states, we must first obtain an expression for the matrix element, which is given by the electron-photon perturbation term

$$W_{ab} = \langle b | \frac{e}{m_0} \mathbf{A} \cdot \hat{\mathbf{p}} | a \rangle. \quad (26.29)$$

At this stage, we need to know the wavefunctions corresponding to the band states  $|a\rangle$  and  $|b\rangle$ . In the effective mass approximation, the electron wavefunction = (envelope function)  $\times$  (Bloch function). The valence band state wavefunction is then

$$\psi_a(\mathbf{r}) = C(\mathbf{r}) u_v(\mathbf{r}) = \underbrace{\frac{e^{i\mathbf{k}_v \cdot \mathbf{r}}}{\sqrt{V}}}_{\text{Envelope } C(\mathbf{r})} \underbrace{u_v(\mathbf{r})}_{\text{Bloch}}, \quad (26.30)$$

and the conduction band state wavefunction is

$$\psi_b(\mathbf{r}) = C'(\mathbf{r}) u_c(\mathbf{r}) = \frac{e^{i\mathbf{k}_c \cdot \mathbf{r}}}{\sqrt{V}} u_c(\mathbf{r}). \quad (26.31)$$

Since the spatial part of the vector potential for the EMag wave is  $\mathbf{A} = \hat{e} \frac{A_0}{2} e^{i\mathbf{k}_{\text{op}} \cdot \mathbf{r}}$ , we obtain the matrix element  $W_{ab} = \langle b | \frac{e}{m_0} \mathbf{A} \cdot \hat{\mathbf{p}} | a \rangle$  to be

$$W_{ab} = \frac{eA_0}{2m_0} \hat{e} \cdot \left( \int \psi_b^* e^{i\mathbf{k}_{\text{op}} \cdot \mathbf{r}} \hat{\mathbf{p}} \psi_a d^3\mathbf{r} \right) \quad (26.32)$$

$$= \frac{eA_0}{2m_0} \hat{e} \cdot \int \left[ \frac{e^{i\mathbf{k}_{\mathbf{c}} \cdot \mathbf{r}}}{\sqrt{V}} u_{\mathbf{c}}(\mathbf{r}) \right]^* \underbrace{\left( e^{i\mathbf{k}_{\text{op}} \cdot \mathbf{r}} \hat{\mathbf{p}} \right)}_{\text{operator}} \left[ \frac{e^{i\mathbf{k}_{\mathbf{v}} \cdot \mathbf{r}}}{\sqrt{V}} u_{\mathbf{v}}(\mathbf{r}) \right] d^3\mathbf{r} \quad (26.33)$$

$$= \frac{eA_0}{2m_0} \hat{e} \cdot \int \left[ e^{-i\mathbf{k}_{\mathbf{c}} \cdot \mathbf{r}} u_{\mathbf{c}}^*(\mathbf{r}) \right] \underbrace{\left( e^{i\mathbf{k}_{\text{op}} \cdot \mathbf{r}} \hat{\mathbf{p}} \right)}_{\text{operator}} \left[ e^{+i\mathbf{k}_{\mathbf{v}} \cdot \mathbf{r}} u_{\mathbf{v}}(\mathbf{r}) \right] \frac{d^3\mathbf{r}}{V} \quad (26.34)$$

$$= \frac{eA_0}{2m_0} \hat{e} \cdot \int \left[ e^{-i\mathbf{k}_{\mathbf{c}} \cdot \mathbf{r}} u_{\mathbf{c}}^*(\mathbf{r}) \right] \left( e^{i\mathbf{k}_{\text{op}} \cdot \mathbf{r}} \right) \left[ e^{+i\mathbf{k}_{\mathbf{v}} \cdot \mathbf{r}} (\hbar \mathbf{k}_{\mathbf{v}} u_{\mathbf{v}}(\mathbf{r}) - i\hbar \nabla u_{\mathbf{v}}(\mathbf{r})) \right] \frac{d^3\mathbf{r}}{V} \quad (26.35)$$

$$= \frac{eA_0}{2m_0} \hat{e} \cdot \underbrace{\int e^{i(-\mathbf{k}_{\mathbf{c}} + \mathbf{k}_{\text{op}} + \mathbf{k}_{\mathbf{v}}) \cdot \mathbf{r}} \left[ u_{\mathbf{c}}^*(\mathbf{r}) u_{\mathbf{v}}(\mathbf{r}) \right] (\hbar \mathbf{k}_{\mathbf{v}})}_{\text{forbidden}} \frac{d^3\mathbf{r}}{V} + \quad (26.36)$$

$$\underbrace{\frac{eA_0}{2m_0} \hat{e} \cdot \int e^{i(-\mathbf{k}_{\mathbf{c}} + \mathbf{k}_{\text{op}} + \mathbf{k}_{\mathbf{v}}) \cdot \mathbf{r}} \left[ u_{\mathbf{c}}^*(\mathbf{r}) \hat{\mathbf{p}} u_{\mathbf{v}}(\mathbf{r}) \right]}_{\text{allowed}} \frac{d^3\mathbf{r}}{V} \quad (26.37)$$

The first term is labeled *forbidden* because the integral is  $\approx \hbar \mathbf{k}_{\mathbf{v}} \langle \mathbf{k}_{\mathbf{c}} | \mathbf{k}_{\mathbf{v}} \rangle = 0$  if we neglect the photon momentum. This is because the states belong to different bands, and are orthogonal. The ‘allowed’ transition matrix element is:

$$W_{ab} = \frac{eA_0}{2m_0} \hat{e} \cdot \int_V e^{i(-\mathbf{k}_{\mathbf{c}} + \mathbf{k}_{\text{op}} + \mathbf{k}_{\mathbf{v}}) \cdot \mathbf{r}} \left[ u_{\mathbf{c}}^*(\mathbf{r}) \hat{\mathbf{p}} u_{\mathbf{v}}(\mathbf{r}) \right] \frac{d^3\mathbf{r}}{V} \quad (26.38)$$

$$= \frac{eA_0}{2m_0} \hat{e} \cdot \int_V \underbrace{e^{i(-\mathbf{k}_{\mathbf{c}} + \mathbf{k}_{\text{op}} + \mathbf{k}_{\mathbf{v}}) \cdot \mathbf{r}}}_{\text{slow}} \underbrace{\left[ u_{\mathbf{c}}^*(\mathbf{r}) \hat{\mathbf{p}} u_{\mathbf{v}}(\mathbf{r}) \right]}_{\text{periodic}} \underbrace{\frac{d^3\mathbf{r}}{V}}_{N\Omega} \quad (26.39)$$

$$= \frac{eA_0}{2m_0} \hat{e} \cdot \int_V \underbrace{e^{i(-\mathbf{k}_{\mathbf{c}} + \mathbf{k}_{\text{op}} + \mathbf{k}_{\mathbf{v}}) \cdot \mathbf{r}}}_{\text{slow}} \underbrace{\left[ u_{\mathbf{c}}^*(\mathbf{r}) \hat{\mathbf{p}} u_{\mathbf{v}}(\mathbf{r}) \right]}_{\text{periodic}} \frac{d^3\mathbf{r}}{\Omega} \quad (26.40)$$

To visualize the slow and periodic parts inside the integral, refer to Figure 26.3. The periodic term functions  $u_{\mathbf{c}}(\mathbf{r}) \hat{\mathbf{p}} u_{\mathbf{v}}(\mathbf{r})$  repeats in every unit cell in real space of volume  $\Omega$  of the crystal. But the slowly varying function of the form  $e^{i\mathbf{k} \cdot \mathbf{r}}$  hardly changes inside a unit cell, it changes appreciably only over many many cells. Then, we treat the slowly varying function as constant inside a the unit cell located at  $\mathbf{R}_i$ , but the value to change from cell to cell. Then, the integral decomposes to

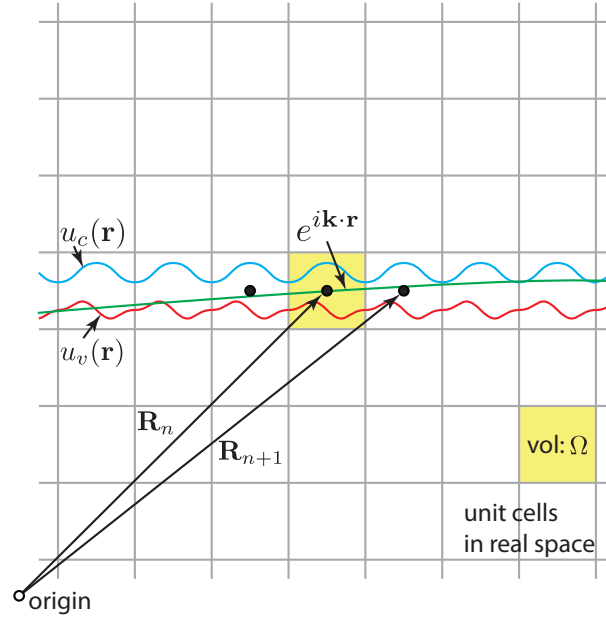


FIGURE 26.3: Explanation of the decomposition of the optical matrix element. Because the matrix element consists of a product of a plane wave part that varies slowly over unit cells, and a part that is periodic in unit cells, the product decomposes into a sum and a cell-periodic integral.

$$W_{ab} = \frac{eA_0}{2m_0} \hat{\mathbf{e}} \cdot \underbrace{\left[ \frac{\sum_{n=1}^N e^{i(-\mathbf{k}_c + \mathbf{k}_{op} + \mathbf{k}_v) \cdot \mathbf{R}_n}}{N} \right]}_{\delta_{\mathbf{k}_c, \mathbf{k}_v + \mathbf{k}_{op}}} \underbrace{\int_{\Omega} [u_c^*(\mathbf{r}) \hat{\mathbf{p}} u_v(\mathbf{r})] \frac{d^3\mathbf{r}}{\Omega}}_{\mathbf{P}_{cv}}. \quad (26.41)$$

The sum runs over all unit cells in real space. Since  $\sum_{n=1}^N e^{i(-\mathbf{k}_c + \mathbf{k}_{op} + \mathbf{k}_v) \cdot \mathbf{R}_n}$  is the sum of the complex exponential at every unit cell site  $\mathbf{R}_n$ , and there are a *lot* of them, let us visualize this sum. Refer to Figure 26.4 to see why the sum  $\frac{\sum_{n=1}^N e^{i(-\mathbf{k}_c + \mathbf{k}_{op} + \mathbf{k}_v) \cdot \mathbf{R}_n}}{N}$  is zero for all cases except when  $-\mathbf{k}_c + \mathbf{k}_v + \mathbf{k}_{op} = 0$ , in which case it is evidently unity. The complex numbers  $e^{i\theta_n}$  are all on the unit circle on the complex plane, and if there are a *lot* of them, they distribute uniformly around the origin. Thus, their sum tends to have zero real and imaginary parts; further, they are divided by a large number  $N$ . But when  $\theta_n = 0$ , all the points fall at  $e^{i0} = 1 + 0i$ , and thus the sum is unity.

The optical matrix element is thus given by the very important result

$$W_{ab} = \frac{eA_0}{2m_0} [\delta_{\mathbf{k}_c, \mathbf{k}_v + \mathbf{k}_{op}}] (\hat{\mathbf{e}} \cdot \mathbf{P}_{cv}) \quad (26.42)$$

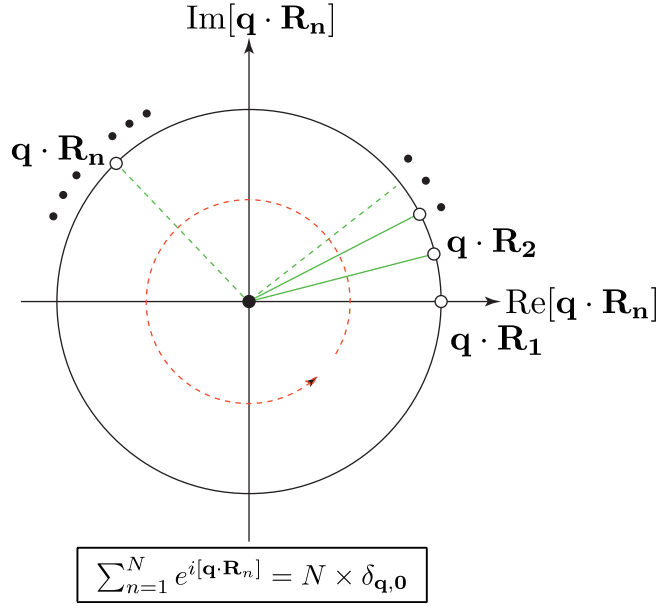


FIGURE 26.4: The sum of complex exponentials of the form  $e^{i\theta_n}$ . If the sum is over a large number of phases, the sum  $\sum_n e^{i\theta_n}$  is zero, unless  $\theta_n = 0$ , in which case  $\sum_n e^{i\theta_n} = N$ . This statement is captured in  $\sum_n e^{i\theta_n} = N\delta_{\theta_n,0}$ .

Note that the Kronecker-delta function ensures momentum conservation because  $\hbar\mathbf{k}_v + \hbar\mathbf{k}_{op} = \hbar\mathbf{k}_c$ . With this form of the optical matrix element, the net absorption rate from equation 26.28 becomes

$$R(\hbar\omega) = \left(\frac{eA_0}{2m_0}\right)^2 \frac{2}{V} \sum_{\mathbf{k}_c} \sum_{\mathbf{k}_v} \frac{2\pi}{\hbar} |\hat{\mathbf{e}} \cdot \mathbf{p}_{cv}|^2 \delta_{\mathbf{k}_c, \mathbf{k}_v + \mathbf{k}_{op}}^2 \delta[E_c(\mathbf{k}_c) - (E_v(\mathbf{k}_v) + \hbar\omega)] \times [f_v(\mathbf{k}_v) - f_c(\mathbf{k}_c)] \quad (26.43)$$

We note that the square of the Kronecker-delta function is the same as the Kronecker-delta  $\delta_{\mathbf{k}_c, \mathbf{k}_v + \mathbf{k}_{op}}^2 = \delta_{\mathbf{k}_c, \mathbf{k}_v + \mathbf{k}_{op}}$ . We also note at this point that  $|\mathbf{k}_c|, |\mathbf{k}_v| \gg \mathbf{k}_{op}$ . This is because the band-edge states occur around reciprocal lattice vectors  $2\pi/a_0$ , and the lattice constants  $a_0 \ll \lambda$ , the wavelength of light. This is the rationale behind the commonly stated fact: *direct optical transitions are vertical in  $E(\mathbf{k})$  diagrams*. Using the Kronecker delta function to reduce the summation over  $\mathbf{k}$  states assuming  $\mathbf{k}_c = \mathbf{k}_v = \mathbf{k}$ , and  $\mathbf{k}_{op} \approx 0$ , and taking the term  $\hat{\mathbf{e}} \cdot \mathbf{p}_{cv}$  out of the sum because it does not depend on  $\mathbf{k}$ , we obtain the net absorption rate per unit volume to be given by the following form, which actually holds also for lower-dimensional structures such as quantum wells, wires, or dots:

$$R(\hbar\omega) = \frac{2\pi}{\hbar} \left(\frac{eA_0}{2m_0}\right)^2 |\hat{\mathbf{e}} \cdot \mathbf{p}_{cv}|^2 \frac{2}{V} \sum_{\mathbf{k}} \delta[E_c(\mathbf{k}) - (E_v(\mathbf{k}) + \hbar\omega)] \times [f_v(\mathbf{k}) - f_c(\mathbf{k})] \quad (26.44)$$

## 26.6 The Equilibrium Absorption Coefficient $\alpha_0(\hbar\omega)$

We are now ready to evaluate the absorption coefficient. Using the expression for  $R(\hbar\omega)$  with the photon flux  $N_{ph}(\hbar\omega)$  from Equation 26.23, the expression for the absorption coefficient from Equation 26.13 becomes

$$\alpha(\hbar\omega) = \underbrace{\left(\frac{\pi e^2}{n_r c \epsilon_0 m_0^2 \omega}\right)}_{C_0} |\hat{e} \cdot \mathbf{p}_{cv}|^2 \frac{2}{V} \sum_{\mathbf{k}} \delta[E_c(\mathbf{k}) - (E_v(\mathbf{k}) + \hbar\omega)] \times [f_v(\mathbf{k}) - f_c(\mathbf{k})]. \quad (26.45)$$

Notice that the absorption coefficient thus formulated becomes *independent* of the intensity of the incident photon radiation  $I \propto A_0^2$  because both  $N_{ph}(\hbar\omega) \propto A_0^2$  and  $R(\hbar\omega) \propto A_0^2$ , and the  $A_0^2$  factor thus cancels in the ratio. This is a signature of a linear process - i.e., the linear absorption coefficient of the semiconductor is a property of the semiconductor alone, and does not depend on the excitation intensity. With the coefficient  $C_0 = \frac{\pi e^2}{n_r c \epsilon_0 m_0^2 \omega}$  we re-write the absorption coefficient again as the following compact expression which will be used also for lower-dimensional structures such as quantum wells, wires, or dots in chapter 27:

$$\alpha(\hbar\omega) = C_0 |\hat{e} \cdot \mathbf{p}_{cv}|^2 \frac{2}{V} \sum_{\mathbf{k}} \delta[E_c(\mathbf{k}) - (E_v(\mathbf{k}) + \hbar\omega)] \times [f_v(\mathbf{k}) - f_c(\mathbf{k})] \quad (26.46)$$

To evaluate the  $\mathbf{k}$ -sum, we need to identify the occupation functions  $f_v(\mathbf{k})$  and  $f_c(\mathbf{k})$ . If the semiconductor is in equilibrium, there is one Fermi level  $E_F$ , and the occupation is given by the Fermi-Dirac function  $f(E) = (1 + \exp[(E - E_F)/k_B T])^{-1}$  at temperature  $T$ . When the semiconductor is pushed to non-equilibrium by either optical excitation or electrical injection of excess carriers, the occupation functions are conveniently modeled by retaining the Fermi-Dirac form. But the single Fermi-energy  $E_F$  splits to two quasi-Fermi levels: one for electrons in the conduction band  $F_c$ , and the other for electrons in the valence band  $F_v$ . The occupation functions are then given by

$$f_v(\mathbf{k}) = \frac{1}{1 + \exp\left(\frac{E_v(\mathbf{k}) - F_v}{kT}\right)} \quad (26.47)$$

$$f_c(\mathbf{k}) = \frac{1}{1 + \exp\left(\frac{E_c(\mathbf{k}) - F_c}{kT}\right)} \quad (26.48)$$

We will consider non-equilibrium conditions in the next chapter. Under thermal equilibrium,  $F_c = F_v = E_F$ , and there is only one Fermi level<sup>3</sup>. For an undoped semiconductor,  $E_F$  locates close to the middle of the bandgap. Then, as  $T \rightarrow 0$  K,  $f_v(\mathbf{k}) \rightarrow 1$  and  $f_c(\mathbf{k}) \rightarrow 0$ . Actually, these conditions hold fine even at room temperature for wide-bandgap semiconductors with little error. Converting the sum to an integral using the usual prescription, we get the *equilibrium* absorption coefficient to be

$$\alpha_0(\hbar\omega) = C_0 |\hat{\mathbf{e}} \cdot \mathbf{p}_{cv}|^2 \frac{2}{V} \times \int_{\mathbf{k}} \frac{d^3\mathbf{k}}{(2\pi)^3} \delta[E_c(\mathbf{k}) - (E_v(\mathbf{k}) + \hbar\omega)] \quad (26.49)$$

Note that the volume term  $V$  cancels.

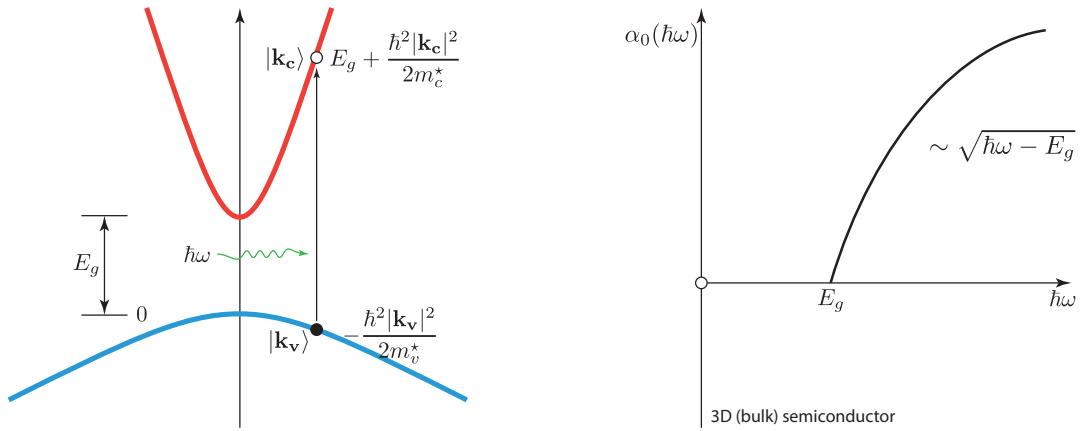


FIGURE 26.5: Definition of various energies, and the equilibrium absorption spectrum of bulk (3D) semiconductors  $\alpha_0(\hbar\omega)$ .

From Figure 26.5, the optical transition can occur only for  $\mathbf{k}$  states that satisfy

$$E_c(\mathbf{k}) = E_g + \frac{\hbar^2 k^2}{2m_e^*} \quad (26.50)$$

$$E_v(\mathbf{k}) = -\frac{\hbar^2 k^2}{2m_h^*} \quad (26.51)$$

$$E_c(\mathbf{k}) - E_v(\mathbf{k}) = E_g + \frac{\hbar^2 k^2}{2m_r^*} \quad (26.52)$$

$$\frac{1}{m_r^*} = \frac{1}{m_e^*} + \frac{1}{m_h^*} \quad (26.53)$$

<sup>3</sup>When photons are incident on the semiconductor, it is by definition not in equilibrium and  $F_c \neq F_v$ . But we assume that the intensity of the EM wave is low enough to ensure that  $F_c \approx F_v \approx E_F$ .



Using spherical coordinates in the 3D  $\mathbf{k}$ -space,  $d^3\mathbf{k} = k^2 \sin\theta dk d\theta d\phi$ , we convert the variables from wavevector to energy. Assuming  $E = \frac{\hbar^2 k^2}{2m_r^*}$ , we break up  $k^2 \sin\theta dk d\theta d\phi$  into three parts:  $k^2 \cdot dk = \left(\frac{2m_r^*}{\hbar^2}\right) E \cdot \frac{1}{2} \left(\frac{2m_r^*}{\hbar^2}\right)^{\frac{1}{2}} \frac{dE}{\sqrt{E}} = \frac{1}{2} \left(\frac{2m_r^*}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{E} dE$ , the second part being  $\sin\theta d\theta$  and the third part  $d\phi$ . When we integrate over all  $k$ -space, the angular parts evaluate to  $\int_0^\pi \sin\theta d\theta = 2$  and  $\int_0^{2\pi} d\phi = 2\pi$ .

The absorption coefficient then becomes

$$\alpha_0(\hbar\omega) = C_0 |\hat{\mathbf{e}} \cdot \mathbf{p}_{\mathbf{cv}}|^2 \frac{2}{(2\pi)^3} \cdot (2\pi) \cdot (2) \cdot \frac{1}{2} \left(\frac{2m_r^*}{\hbar^2}\right)^{\frac{3}{2}} \underbrace{\int_0^\infty dE \sqrt{E} \times \delta[E - (\hbar\omega - E_g)]}_{\sqrt{\hbar\omega - E_g}} \quad (26.54)$$

which reduces to

$$\boxed{\alpha_0(\hbar\omega) = C_0 |\hat{\mathbf{e}} \cdot \mathbf{p}_{\mathbf{cv}}|^2 \underbrace{\frac{2}{(2\pi)^2} \left(\frac{2m_r^*}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{\hbar\omega - E_g}}_{\rho_r(\hbar\omega - E_g)}} \quad (26.55)$$

where we have defined the joint optical density of states function for bulk 3D semiconductors as

$$\rho_r(u) = \frac{g_s}{(2\pi)^2} \cdot \left(\frac{2m_r^*}{\hbar^2}\right)^{\frac{3}{2}} \cdot \sqrt{u} \quad (26.56)$$

Figure 26.5 shows the equilibrium absorption spectrum  $\alpha_0(\hbar\omega)$  of a bulk 3D semiconductor. Using typical values of effective masses and material constants, it may be verified that the absorption coefficient for GaN for example are of the order of  $\sim 10^5 \text{ cm}^{-1}$ , as indicated in Fig 26.1 at the beginning of this chapter. The absorption coefficient is zero for photon energies below the bandgap of the semiconductor, as is intuitively expected.

Instead of leaving the expression for the absorption coefficient in terms of the unphysical parameter  $C_0$ , we use the fundamental Rydberg energy  $R_\infty = \frac{e^2}{4\pi\epsilon_0(2a_B)}$ , the Bohr radius  $a_B = \frac{\hbar}{m_0\alpha}$ , and the fine structure constant  $\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c}$  to write the absorption coefficient as

$$\boxed{\alpha_0(\hbar\omega) = \left(\frac{4\pi^2\alpha}{n_r}\right) \cdot (R_\infty a_B^2) \cdot \left(\frac{2|\hat{\mathbf{e}} \cdot \mathbf{p}_{\mathbf{cv}}|^2}{\hbar\omega}\right) \cdot \rho_r(\hbar\omega - E_g)} \quad (26.57)$$

where we have split off the dimensionless term  $2|\hat{\epsilon} \cdot \mathbf{p}_{\mathbf{cv}}|^2/m_0\hbar\omega$ . Note that as discussed in chapter 14, the rough order of  $2|\hat{\epsilon} \cdot \mathbf{p}_{\mathbf{cv}}|^2/m_0 \approx 20$  eV for most bulk semiconductors. The coefficients decompose to reveal a proportionality to the fine-structure constant. The term  $R_\infty a_B^2$  has units eV.cm<sup>2</sup>, and the reduced density of states is in 1/eV.cm<sup>3</sup>, which leads to the net units cm<sup>-1</sup>. This general form of the equilibrium absorption coefficient holds even for low-dimensional structures with the suitable DOS  $\rho_r(\hbar\omega - E'_g)$ , where  $E'_g$  accounts for ground state quantization shifts in the bandgap. Many interesting effects happen when the semiconductor is pushed out of equilibrium: it is the subject of the next chapter.

Debdeep Jena: [www.nd.edu/~djena](http://www.nd.edu/~djena)