# Chapter 1

# Electrons, Photons and Phonons

Version 2.12 (July 31, 2012), ©Prentice Hall

**Outline**

Quantum mechanics, statistical mechanics and solid state physics constitute the fundamental underpinnings of modern semiconductor device physics. In-depth understanding of the detailed processes lurking inside microelectronic devices demands a firm foundation in these disciplines. Fortunately for microelectronics engineers, it is possible to go quite far in the analysis and design of devices armed solely with an intuitive understanding of a few basic results. The goal of this chapter is to present selected basic concepts in an intuitive and mathematically non-rigorous fashion.

In this chapter we will concentrate our interest on the nature of the electron and its behavior in a crystalline semiconductor. Additionally, we will introduce two more "particles": the *photon* as a quantum of light and the *phonon* as a quantum of vibrational energy of the semiconductor lattice. These particles play a decisive role in semiconductor device physics.

The brief summary presented in this chapter obviously cannot do justice to the depth and breadth of these disciplines. In most educational institutions, senior or graduate-level courses are offered to provide a working-level knowledge of these topics. Students who have already studied these subjects should be able to skip the corresponding sections of this chapter.

## 1.1 Selected concepts of quantum mechanics

The formulation of quantum mechanics at the beginning of the 20th century revolutionized the understanding of the world that existed at that time. The "new physics" would have profound implications on science and technology to this date. Classical long-standing problems such as the blackbody radiation, to select one of interest to us in this chapter, were finally solved. Many of the semiconductor pioneers had superior command of quantum mechanics as well as solid-state physics and statistical mechanics. This is splendidly illustrated in Shockley's book *Electrons and Holes in Semiconductors*, first published in 1950 (for more details on this and other suggested reading, see Section 1.5 at the end of the Chapter).

### 1.1.1 The dual nature of the photon

There are multiple aspects to quantum mechanics. Many are very relevant to the way semiconductor devices operate; most, unfortunately, are not very intuitive. In fact, many of them contradict outright common human perceptions of the physical world.

Light is a good example. In the late 19th century in spite of the great progress achieved in the understanding of electromagnetic radiation, a few puzzling observations about light could not be explained. There was a great deal of interest, for example, in the spectrum of light emitted by objects. It was observed that heated objects had a prominent light emission spectrum that depended on their temperature. After correcting for the reflectivity of the surface, it was found that there was a "universal" spectrum of light emission for all objects that received the name of *blackbody radiation.*

A blackbody is an ideal object which does not reflect any light at all. It emits light, however, with a spectrum that depends only on its temperature. In fact an ideal blackbody appears all

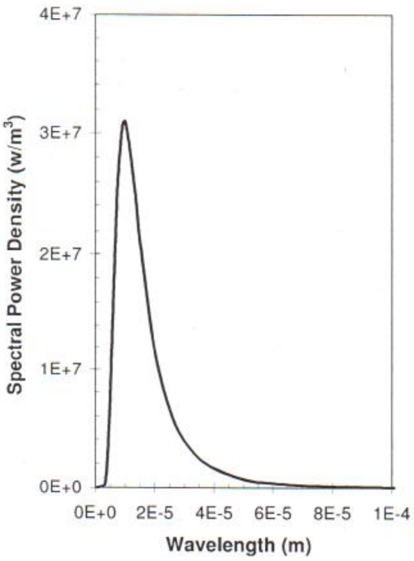


Figure 1.1: Blackbody light emission spectrum at 300 K.

black to the human eye unless it is so hot that it emits light in the visible portion of the spectrum. Familiar examples of blackbody radiation are hot coals in a campfire or red-hot iron in a foundry.

Fig. 1.1 sketches the spectrum of an ideal blackbody at 300 K. The figure depicts the spectral power density as a function of wavelength (the units of the vertical axis are $W/m^3$, that is, 1 $W$ of power falls over a 1 $m^2$ of surface normal to the radiation per $m$ of wavelength). In the visible range, there is very little energy being emitted by a blackbody at room temperature. In fact, the peak of the blackbody radiation at room temperature is in the infrared. [1] At the turn of the century, the peculiar spectral distribution of the radiated energy of a blackbody seen in Fig. 1.1, with a sharp threshold at short wavelengths and a long tail at high wavelengths, defied understanding. Both the shape of the curve and its evolution with temperature could not be explained by the classical electromagnetic theory of light.

A breakthrough in understanding blackbody radiation came from Planck in 1901. Planck postulated that light was emitted or absorbed by electromagnetic "oscillators" of an energy that is restricted to "quantized" values that are multiples of a fundamental energy amount:

$$E_{ph} = h\nu \tag{1.1}$$

where $\nu$ is the frequency of the light, and $h$ is Planck constant ($6.63 \times 10^{-34}\ J \cdot s$, or $4.14 \times 10^{-15}\ eV \cdot s$ in units more common to microelectronics engineers; see Table I at end of book). In Planck's hypothesis, if an oscillator lowers its energy from $nh\nu$ to $(n-1)h\nu$ (where $n$ is an integer), it emits a quantum of electromagnetic radiation, or **photon**, of energy $E_{ph}$ given by

---

[1] This has practical consequences. Infrared cameras create a temperature image that enables night vision. They are also used to measure temperature in integrated circuits.

Eq. 1.1. Similarly, an oscillator vibrating with energy $nh\nu$ can absorb a photon of the same frequency in order to acquire a final energy $(n+1)h\nu$. Planck postulated light to be made out of tiny "lumps" that pack an amount of energy that is proportional to the frequency of the light. This revolutionary hypothesis represented a radically different way of thinking about light and became the key for the explanation of blackbody radiation. Planck's hypothesis explained for the first time, the wavelength dependence of the blackbody radiation depicted in Fig. 1.1. The details are left to an introductory course on quantum mechanics.

**Exercise 1.1**: *Find the relationship between the energy of a photon in $eV$ and its wavelength in $\mu m$.*

Using Eq. 1.1, the relationship between wavelength and frequency of light, $\nu = \frac{c}{\lambda}$, and the values of the fundamental constants listed in Table I, we can easily find:

$$\begin{aligned} E_{ph} &= h\frac{c}{\lambda} = 4.14 \times 10^{-15}\ eV \cdot s \times \frac{3.00 \times 10^{10}\ cm/s}{\lambda\ (cm)} = \frac{1.24 \times 10^{-4}}{\lambda (cm)}\ eV = \frac{1.24 \times 10^{-4}}{\lambda(\mu m) \times 10^{-4}} eV \\ &= \frac{1.24}{\lambda\ (\mu m)} eV \end{aligned} \tag{1.2}$$

This is a simple relationship that is handy to remember. For an energy of 1.1 eV (we will soon appreciate the significance of this particular number), the wavelength is 1.13 $\mu m$, which is in the infrared.

**Exercise 1.2**: *Calculate the photon fluence (integrated flux) emitted by a GaAs laser that is delivering 1 $mW$ of light power at a wavelength of 0.85 $\mu m$.*

This is another exercise where care is needed in the handling of units. The power delivered by a beam of light is equal to the photon fluence, $F_{ph}$, times the photon energy. Using the simple relationship derived in the previous exercise, a 0.85 $\mu m$ photon has an energy of 1.46 eV. We can then calculate the fluence of the light beam in the following way:

$$F_{ph} = \frac{P}{E_{ph}} = \frac{1\ mW}{1.46\ eV} = \frac{10^{-3}\ J/s}{1.46\ eV} = \frac{10^{-3}\ J/s}{1.46\ eV} \frac{1}{1.60 \times 10^{-19}\ J/eV} = 4.28 \times 10^{15}\ s^{-1}$$

This result means that every second there are over 4,000 trillion photons with a wavelength of 0.85 $\mu m$ being emitted by the laser.

Another important aspect of quantum mechanics is the *particle-wave duality.* In quantum mechanics, there is no distinction between a wave and a particle. Waves can behave like particles and vice versa. Light is again a good example. [2] Under the right conditions, electrons can be ejected from a metal by light impinging on its surface, a process called the *photoelectric effect.* When this experiment was first carried out, it was observed that the kinetic energy of the photogenerated electrons, or simply photoelectrons, was independent of the light intensity but exhibited a peculiar behavior as a function of the frequency of the incoming light. This is sketched in Fig. 1.2. There was a threshold frequency, $\nu_c$, characteristic of each metal, below which no photoelectron was emitted. Beyond this threshold, the electron kinetic energy was found to increase with the frequency of the light.

---

[2] Interestingly, at the time of Newton, a particle view of light prevailed. It was Huygens that demonstrated the

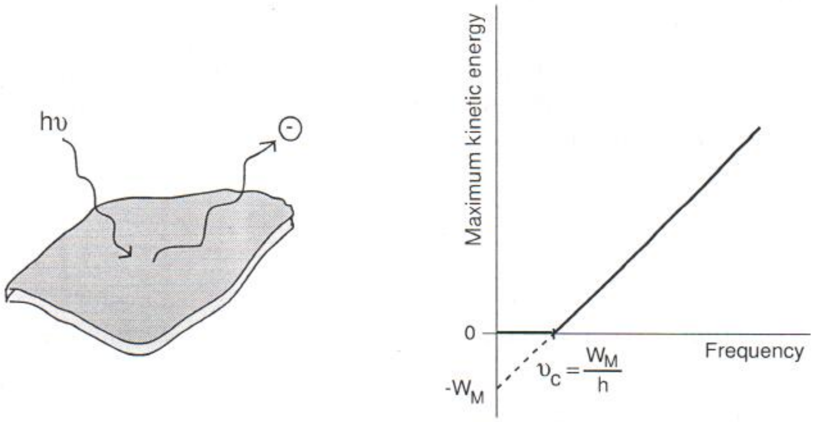


Figure 1.2: Left: a beam of light impinging on a metal can provoke the emission of electrons from it (the photoelectric effect). Right: sketch of kinetic energy of electrons ejected from a metal surface as a function of the frequency of impinging light.

The classical explanation was that for an electron to escape from the metal, it had to acquire enough energy to overcome the binding forces. This energy is called the *work function* of the metal, $W_M$. Since an electromagnetic wave has an energy density associated with it, it is reasonable to think that exposing the metal to the radiation for sufficient time would allow the metal to accumulate enough energy so as to ultimately eject an electron. In the classical picture, this should happen regardless of the frequency of the light.

An explanation based on the quantum nature of light was advanced by Einstein in 1905. Einstein postulated that the energy of an individual incoming photon is given to a single electron in a two-body collision. If the energy of the photon is lower than the metal work function, the electron does not acquire enough energy to escape from the metal. If, on the other hand, the energy of the photon is higher than $W_M$, the electron is allowed to escape and the excess energy over $W_M$ is provided to the electron as kinetic energy. Following this explanation, the threshold frequency for the light to eject an electron can be found using Eq. 1.1 to be $\nu_c = W_M/h$. The extrapolation of the electron kinetic energy to zero frequency gives $W_M$. In extensive experiments that followed Einstein's hypothesis, Millikan established the correctness of Einstein's predictions. As a by-product of these experiments, the work functions of many metals were measured.

At the core of Einstein's hypothesis is the exchange of energy between one electron and one photon. Since this treats the photon as a well defined particle, it was an unprecedented concept at the time. The most direct evidence of the particle nature of electromagnetic radiation was provided by Compton, who studied the scattering of X-rays through a thin metal foil (X-rays are simply high-energy photons). The *Compton effect*, described in detail in many quantum mechanical textbooks, could only be explained by invoking billiard-ball type interactions between photons and electrons, highlighting one more time the particle nature of light.

---

wave nature of light. In some way, quantum mechanics brought together the particle and wave views of light.

### 1.1.2 The dual nature of the electron

In 1925, inspired by the dual wave-particle nature of the photon, de Broglie postulated (in his PhD thesis) that particles also exhibit wave properties. He formulated an expression for the wavelength associated with a particle that stated:

$$\lambda_B = \frac{h}{p} \tag{1.3}$$

where $p$ is the momentum of the particle. $\lambda_B$ is known as the *de Broglie wavelength*, a cornerstone of modern quantum theory.

De Broglie hypothesized that the wave nature of a particle is only relevant when its de Broglie wavelength is comparable to the physical dimensions of its environment. If the wavelength is much smaller than the dimensions of the physical environment, classical particle behavior provides an adequate description of the relevant physics. When the wavelength is comparable or much larger than the environment, the wave nature of the particles becomes relevant and wave-like phenomena such as diffraction and interference should be observed.

In search of a confirmation of de Broglie's predictions, in 1927 Davisson and Germer studied diffraction of low-energy electrons from a crystal of Ni. They obtained diffraction patterns with a shape that could not be explained by classical physics in which electrons are considered as particles. In fact, a peak in the diffraction pattern was observed precisely when the momentum of the incoming electrons had an associated de Broglie wavelength equal to the interatomic spacing of Ni. This constituted an experimental proof of the wave nature of the electron.

**Exercise 1.3**: *Calculate the de Broglie wavelength of a free electron in vacuum with a kinetic energy of 26 meV.*

A relationship between the kinetic energy of an electron and its momentum is obtained through (relativistic effect neglected):

$$E_K = \frac{1}{2} m_o v^2 = \frac{p^2}{2m_o} \tag{1.4}$$

If we solve for $p$ and plug into Eq. 1.3 above, we get a relationship between the de Broglie wavelength and the kinetic energy:

$$\lambda_B = \frac{h}{\sqrt{2m_o E_K}} \tag{1.5}$$

Inserting numbers into this:

$$\lambda_B = \frac{4.14 \times 10^{-15}\ eV.s}{\sqrt{2 \times 5.69 \times 10^{-16}\ eV \cdot s^2/cm^2 \times 0.026\ eV}} = 7.61 \times 10^{-7}\ cm \simeq 7.6\ nm$$

The experiment of Davisson and Germer is particularly relevant for semiconductor engineers because, as we will see below, typical semiconductors are solids with interatomic distances of the order of a fraction of a $nm$. As Exercise 1.3 above showed, these dimensions are substantially smaller than the de Broglie wavelength of electrons with energies normally encountered in the operation of semiconductor devices. The quantum nature of the electron therefore plays a dominant role in many phenomena associated with semiconductors.

The wave nature of particles presents difficulties when trying to precisely pinpoint the position of any individual particle. The particle is somehow spread out in space. Quantum mechanics therefore only talks in probabilistic terms. Although in the quantum regime there is a sense of uncertainty about the precise behavior of individual particles, the properties of interest in semiconductor engineering involve the average behavior of a large population of electrons. The probabilistic language of quantum mechanics is perfectly suited to this task.

Quantum mechanics handles the wave nature of particles by defining a *wave function* whose modulus square is the probability of finding the particle at any one point in space at any one time. Knowing the wave function of an electron in a given system is the key to understanding the electronic properties of the system in question. In 1926, Schrödinger postulated an equation to describe the spatial distribution and the dynamics of the wave function. Schrödinger's wave equation was not deduced from first principles - he hypothesized it. Seventy years of testing against experiments have amply established the validity and usefulness of this equation (for non-relativistic physics). One of the greatest success of Schrödinger's wave equation has been in its contribution to the understanding of the physics of solids. Its impact in this field has been far reaching. Progress in semiconductor physics and technology have brought about the electronics and communications revolutions that continue on after more than three decades.

Since we will not use Schrödinger's wave equation in this book, it is not useful to reproduce it here. The reader can find it in all books on quantum mechanics. It is however necessary to appreciate a few implications of Schrödinger's equation which are discussed in the following section. Before closing this section, it is important to note that there is still an advanced quantum mechanical concept, *tunneling*, which is of great relevance to modern devices. Tunneling describes the finite probability that particles penetrate into regions that are classically forbidden to them. Tunneling is a manifestation of quantum mechanics that is pervasive in semiconductor devices. In fact, as we discuss in this book, tunneling is at the heart of most ohmic contacts and it is also responsible for certain breakdown phenomena and parasitic currents in devices. Tunneling is also the basis of some important devices, such as EEPROMs (Electrically-Erasable Programable Read-Only Memory) and Zener diodes. In spite of its importance, the eminently quantum-mechanical nature of tunneling prevents us from treating it rigorously in this book.

### 1.1.3 Electrons in confined environments

A certain class of electronic systems is of great interest to us. These are *bound systems* in which a field of forces would classically restrict the electron movement to a region of space. A good example of such a system is an atom (Fig. 1.3). In an atom, the Coulombic attraction between the electrons and the protons in the nucleus effectively confines the electrons to the vicinity of the nucleus, much as the Moon is restricted to rotate around the Earth due to their mutual gravitational attraction. Figure 1.3 (a) sketches the Coulombic potential of an atom. This is an energy representation of the electrostatic field of forces which acts between electrons and the

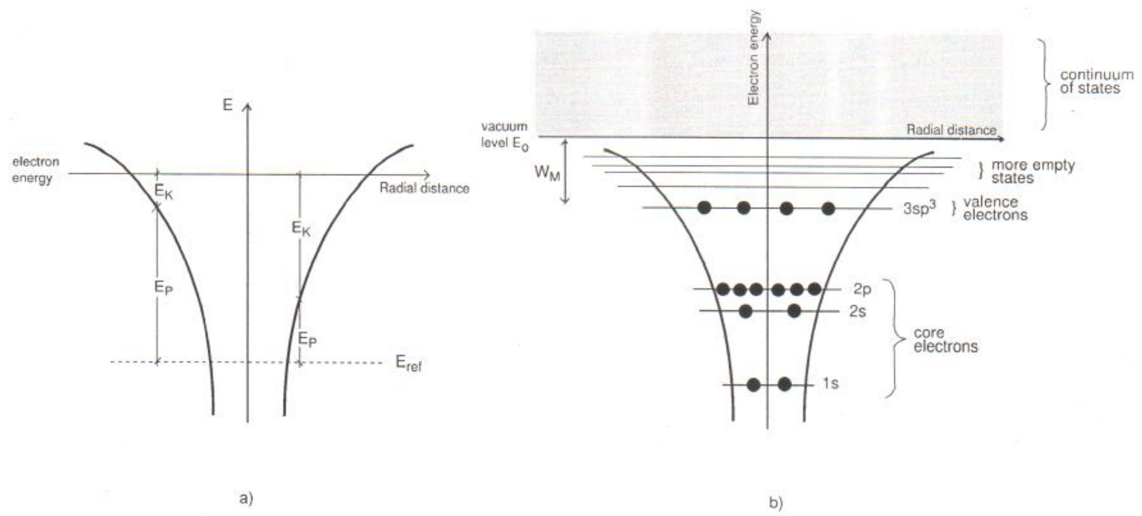


Figure 1.3: a) Sketch of the Coulombic potential of an atom. As an electron wanders around the nucleus, it trades potential energy for kinetic energy. b) Electron energy levels of a Si atom.

nucleus. As the electron moves within the sphere of influence of the nucleus, it trades back and forth potential energy $E_P$ for kinetic energy $E_K$.

The wave nature of the electron imposes further restrictions to its movement. A string in a guitar, for example, can only oscillate in modes or frequencies that depend on the length, composition, and tension of the string. In a similar way, the confinement of the electron around a nucleus results in only a discrete set of allowed "orbits" (using a particle view), each one characterized by a certain energy. An energy representation of the situation is shown in Fig. 1.3 (b) for a Si atom. This figure sketches the various levels allowed for electrons when bound to it. They are obtained from a solution of Schrödinger's wave equation with the appropriate potential for Si.

The allowed orbits of electrons around an atom are classified using a set of characteristic *quantum numbers*. These numbers arise from solving Schrödinger's equation. They are handy indices to catalog the various "shells", "subshells" "orbitals" and "spin" of the various *quantum states* that the electrons can occupy. The number of electrons of an atom and their distribution among the various quantum states determines to a great extent the electrical and chemical properties of the atom. It also greatly impacts the electronic behavior of solids made out of these atoms.

The arrangement of the electrons in the quantum states of the atom is restricted by the *Pauli exclusion principle*, postulated by Pauli in 1925. According to this principle, no two electrons in a given quantum system can occupy the same quantum state. Without this postulate, most experimental observations, such as the shape of the blackbody spectrum, cannot be explained. In fact, without Pauli's exclusion principle, most microelectronics phenomena of interest to us cannot be understood.

The Pauli exclusion principle implies that the electron distribution in a system is found by filling quantum states with one electron each starting from the lowest energy. This is shown for the 14 electrons of the Si atom in Fig. 1.3 (b). The lowest energy shell, labeled *1s*, contains two quantum states and therefore can hold two electrons. The next shell, contains two subshells at different energies. The *2s* subshell can hold two electrons, while the *2p* subshell can accommodate six with different quantum numbers. These two subshells therefore become completely full. At this point there are four electrons left. Ideally, two would go to the next *3s* subshell (which would become full) and two more to the *3p* subshell (which would remain only partially full). However, there is a minor complication. The *3s* and *3p* subshells become "hybridized", or mixed, into four $sp^3$ subshells with the same energy. Since each one can hold two electrons, the four $sp^3$ subshells are only partially occupied - four states are full while four additional ones remain empty. This fact will have an important consequence when Si atoms bond together to form a solid, as we will see. Above these subshells, there are other shells and subshells that are empty.

It is important to have a sense of how tightly bound the electrons are in a Si atom. Imagine for a moment that we try to remove one electron from a Si atom by hitting it with energetic photons, such as X-rays [3]. It takes at least 1847 eV to free up a *1s* electron. Any excess energy above this value will provide extra kinetic energy to the released electron. At least 158 eV and 108 eV are needed to free a *2s* or a *2p* electron, respectively. In contrast, it takes only 8.2 eV to free an electron from the four $sp^3$ uppermost orbitals. Clearly the four electrons in the outer shell are far more weakly bound to the atom than the remaining ten. These four *valence electrons*, as they are known, play a key role in the chemical reactions of Si and in the formation of bonds in a Si crystal. On the other hand, the ten *core electrons* are much harder to dislodge and remain, in all situations that we encounter in microelectronics, tightly bound to the Si nucleus.

An electron that has just enough energy to overcome the attraction of the nucleus is represented at an energy $E_o$, the *vacuum level*, as shown in Fig. 1.3 (b). An electron with this energy is at rest infinitely far away from the atom. Most likely, the process that frees the electron also gives it some extra kinetic energy. This electron can move around and it is represented in a "continuum" of quantum states at energies above $E_o$, as represented in Fig. 1.3 (b). The meaning of the work function, first introduced in Section 1.1.1, now becomes clear: it is the minimum energy required to free the valence electrons, that is, the energy difference between $E_o$ and the highest electron energy.

Before closing this section, let us note that the crucial role valence electrons play in the properties of atoms is recognized in the very organization of the periodic table. The vertical columns of the table contain atoms with the same number of valence electrons. Fig. 1.4, for example, shows a small section of the periodic table of particular interest to us as it contains the key elements constitutive of common semiconductors. The column number (in roman numbers) indicates the number of valence electrons in each atom in that column. Si belongs to column IV, with Ge and C. As one proceeds down the rows of the periodic table, the atomic number (at the upper right hand of each symbol in Fig. 1.4) increases and more shells become completed. For example, C has shell 1 completely full and shell 2 only partially full. In the case of Si, as we just saw, shells 1 and 2 are full while shell 3 is partially full. Down the table, Ge has shells 1, 2 and 3

---

[3] The technique of using X-rays to measure the electron binding energies of atoms is known as X-ray Photoelectron Spectroscopy, or XPS. Since the electron binding energies are peculiar to each element, XPS is widely used in the electronics industry to study the composition of materials.

| IIB | IIIA | IVA | VA | VIA |
|---|---|---|---|---|
| | 5<br>B | 6<br>C | 7<br>N | 8<br>O |
| | 13<br>Al | 14<br>Si | 15<br>P | 16<br>S |
| 30<br>Zn | 31<br>Ga | 32<br>Ge | 33<br>As | 34<br>Se |
| 48<br>Cd | 49<br>In | 50<br>Sn | 51<br>Sb | 52<br>Te |

Figure 1.4: Subsection of the periodic table showing the most common elements constitutive of semiconductors. The atomic number is indicated at the upper right of each chemical symbol.

full while shell 4 is partially full, and so on. We will refer back to this figure several times later in this book.

## 1.2 Selected concepts of statistical mechanics

In the operation of semiconductor devices, a very large number of electrons are typically involved. In devices we are interested in the collective behavior of the electron population rather than the details of any individual electron. Statistical mechanics is a discipline concerned with large systems of indistinguishable particles. Statistical mechanics describes the macroscopic properties of a whole system in a probabilistic fashion without inquiring about the detailed behavior of any one particle at any one time.

### 1.2.1 Thermal motion and thermal energy

It is a common observation that at finite temperatures, nothing in the world is really standing still. Molecules in a gas or a liquid, atoms in a solid, and electrons in a transistor are all jiggling around. This is called *Brownian motion* in honor of botanist Robert Brown who discovered it in 1827. At a finite temperature, particles in an ensemble have a certain average kinetic energy. Some have more, some have less. There is a certain kinetic energy distribution around this average. While it is not possible to measure the kinetic energy of a given particle, it is not difficult to measure the average kinetic energy of an ensemble of particles. We can, for example, confine a known number of gas molecules at a given temperature in a given volume and measure the pressure that they exert on the walls of the vessel. For a gas under ideal conditions, if we do this we find that the mean value of the kinetic energy of a particle is related only to temperature and not to the nature of the particle such as its mass. This fact is used to *define* temperature itself. There is some arbitrariness in this definition and that is why there are different temperature scales. In science and engineering, the following convention for the average kinetic energy of a molecule of an ideal gas is widely used:

$$< E_K >= \frac{3}{2}kT \tag{1.6}$$

where $k$ is the Boltzmann constant and $T$ is the temperature in degrees Kelvin (the temperature in degrees centigrade is equal to the value in Kelvin minus 273.15). The value of the Boltzmann constant is $k = 1.38 \times 10^{-23}\ J/K$ or, with energies measured in electron volts, $k = 8.62 \times 10^{-5}\ eV/K$. Eq. 1.6 applies to an ideal ensemble of indistinguishable classical particles with three degrees of freedom. It makes a statement about the *average* kinetic energy of a molecule. There are certainly particles with higher and lower kinetic energy than this average. [4] For a classical particle at "room temperature" (defined in this book as 300 K or $27°C$), we find, using Eq. 1.6, that its average kinetic energy is 39 meV.

The value $\frac{3}{2}kT$ is an important energy reference. Processes that occur with characteristic energies smaller than $\frac{3}{2}kT$ are unlikely to be distinguishable. Processes that involve a well defined energy exchange still will appear "blurred" in energy by the thermal energy. For example, this is the case for the energy spectrum of light emission in light-emitting diodes (LEDs). Since the product $kT$ alone (without the $\frac{3}{2}$ factor) shows up in many equations in a great variety of physical processes, it is often used as an order-of-magnitude estimate of the average kinetic energy of a particle. $kT$ is called the *thermal energy* and its value at room temperature is 26 meV, about 1/40th of an eV.

### 1.2.2 Thermal equilibrium

Pervasive in our discussion above, but not explicitly mentioned, is the concept of *thermal equilibrium.* Let us define this concept. A system of particles is in thermal equilibrium if it fulfills two conditions: first, it does not exchange energy with the outside world, and second, it is in steady state. Let us examine these conditions in more detail.

When we picture an ideal gas confined in a vessel at a certain temperature, such as sketched in Fig. 1.5, we assume that no energy can be lost from or added to the gas ensemble through the walls of the container. This means that, the collisions that the gas particles experience against the walls of the vessel are perfectly "elastic", so no energy is lost or added as a result of them. Also, external energy, such as light, cannot penetrate to the interior of the vessel. This is called a *closed system* and it is a necessary but not sufficient condition for thermal equilibrium.

An additional requirement for a situation to be characterized as thermal equilibrium, as Fig. 1.5 illustrates, is that all time derivatives of all ensemble averages, global and local, are zero. That is, the system is in *steady state.* It is possible for a system to be closed and yet not to be in thermal equilibrium. Suppose that at $t = 0^-$ we shine light onto a vessel containing gas, that is otherwise completely isolated from its surrounding, with an energy that is absorbed by the gas molecules. Assume furthermore that only a fraction of the vessel is exposed to the radiation. At

---

[4] The factor of $\frac{3}{2}$ in Eq. 1.6 was introduced in the theory of ideal gases as a matter of convenience.

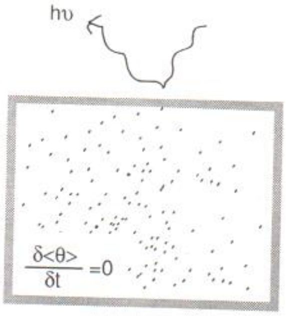


Figure 1.5: Illustration of a system of particles in thermal equilibrium. The system is "closed" and all time derivatives of all ensemble averages ($\theta$ in the figure) are zero.

$t = 0^+$, all energy exchanges with the outside world are stopped. In spite of this, for $t > 0^+$ the gas ensemble is not in thermal equilibrium. On average, in a region of the vessel the gas molecules have higher kinetic energy than in the other. As time goes on, collisions between the molecules randomize the average kinetic energy. After some time, the average kinetic energy of the molecules is the same in all regions of the vessel. Only then we can properly talk about thermal equilibrium. The time that it takes for this condition to be established depends on the average collision time and the dimensions of the vessel.

The combination of a closed system and steady state is a very rigorous definition of thermal equilibrium[5]. When studying semiconductor devices, we will often discuss the thermal equilibrium condition in some level of detail. In strict thermal equilibrium, a device does not perform any useful function. However, this state is important because when disturbed by some outside influence, semiconductor devices react by trying to restore thermal equilibrium. If we wish to identify the bottlenecks to the reestablishment of thermal equilibrium and answer questions such as the time that it takes for equilibrium to be reached, we need to understand this state well. There are many examples in nature that illustrate this. For example, if we wish to know the frequency at which a string in a piano vibrates when struck by one of the hammers, clearly an event outside equilibrium, we must know the tension of the string. This is a parameter that applies to the string in equilibrium. A higher tension will result in a higher pitch tone.

In semiconductor devices, we will frequently encounter situations that we will denote as *quasi-equilibrium*. This is a concept that is slightly harder to understand but enormously important. Basically, in many circumstances, perturbing the device (by applying a voltage or shining light on it, for example), will not significantly modify some of the properties that we know well in equilibrium. This allows us to "stretch" results obtained in equilibrium and apply them to situations outside equilibrium.

An analogy might help to clarify the concept of quasi-equilibrium. Consider carrying out an experimental study of the physics of canoeing. An ideal place to conduct experiments is a perfectly still lake on a day without any wind. Under these conditions, one could evaluate, among other things, the movement of the water around the canoe and the energy conversion efficiency of the canoe/human system. Those ideal experimental conditions might be hard to get for a period

---

[5] At times, we will refer to it simply as *equilibrium*, for short.

of time long enough to complete the study. It is clear that one could also carry out many valid experiments on a gently flowing river or under light wind. If the water current or the wind speed are weak enough, one can surely make many useful observations about the canoe/water system. Conversely, whatever one might have learned about the canoe/water behavior in ideal conditions probably applies too to the less ideal situations of light wind or weak current. A perfectly still lake in the absence of wind can be considered a thermal equilibrium situation. A gently flowing river or a lake under a light breeze are situations outside thermal equilibrum. For many purposes, we can think of these though as thermal equilibrium situations. In contrast, if one were to study the movement of a canoe on a fast moving river with rapids, or under strong wind that produce sizable waves, the behavior of the canoe is likely to be very different. These are clearly situations out of equilibrium that require special consideration.

The concept of quasi-equilibrium will be extensively used in this book in different contexts. Judiciously invoking quasi-equilibrium greatly simplifies the development of a wide range of analytical models for microelectronic device operation.

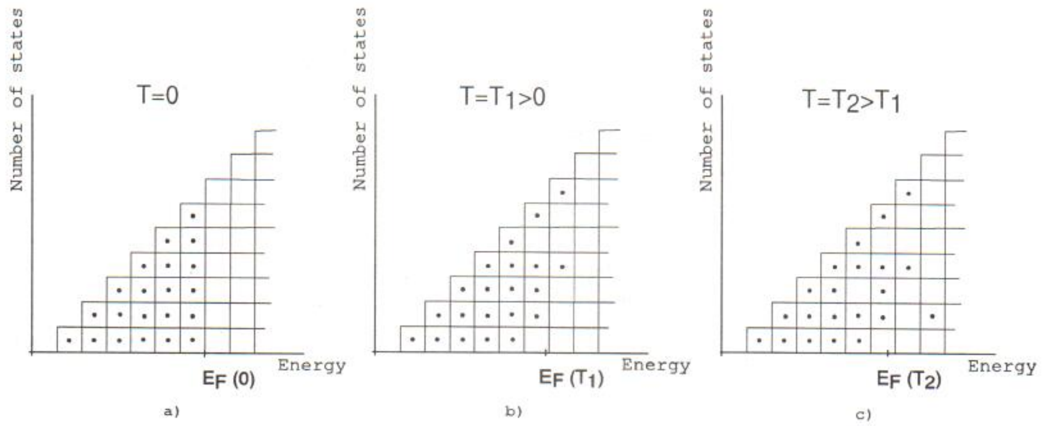
### 1.2.3 Electron statistics

In the discussion above on the thermal motion of gas molecules, no constraints were imposed on the motion of the molecules. We have already seen, however, that electrons are rarely entirely free. In any electronic system, there are only certain quantum states that electrons can occupy. Furthermore, electrons must follow the Pauli exclusion principle in occupying these states. These two restrictions must be kept in mind when dealing with the thermal properties of electrons.

Consider a confined gas of electrons in thermal equilibrium and assume that collisions among electrons are very rare, that is, the electron gas is very dilute. Now let us inquire about the energy distribution of these electrons at absolute zero temperature. The confined system in question will have a certain number of allowed quantum states. Classically, at absolute zero temperature, one would place the entire supply of electrons at the lowest energy available. Quantum mechanically, however, electrons obey Pauli exclusion principle and only one electron can occupy each quantum state. Therefore, the electron energy distribution at 0 K is obtained by filling the available quantum states with electrons starting with the lowest energy and proceeding up in energy until all electrons are placed. The energy of the top-most filled state in the distribution is called the *Fermi energy*, or *Fermi level* $E_F$.

We can visualize the meaning of the Fermi level by looking at the example depicted in Fig. 1.6. Here we have an electronic system that consists of 21 electrons and an electron state distribution with energy in a staircase shape, as sketched in the figure. At 0 K, the lowest 21 states in energy get filled (see Fig. 1.6(a)). The energy at which the state occupancy goes from 1 to 0 (or 100% to 0%) is the location of the Fermi level $E_F$.

At finite temperatures the electrons possess a finite amount of kinetic energy that increases with temperature. This implies that some states above the Fermi energy, which were empty at 0 K, are now occupied, while others below $E_F$, which were full at 0 K, now become empty. Statistical mechanics provides us with an expression that allows us to compute the probability that a certain state is occupied in thermal equilibrium. This is called the *Fermi-Dirac distribution*



a) b) c)

Figure 1.6: Example of a quantum state distribution and its electron occupation as a function of temperature. (a) depicts the situation at T=0 K, (b) and (c) at finite temperature $T_1$ and $T_2$, respectively, with $T_2 > T_1$.

*function*:

$$f(E) = \frac{1}{1 + \exp \frac{E - E_F}{kT}} \tag{1.7}$$

which is depicted in Fig. 1.7.

The Fermi-Dirac distribution function is a fundamental result of statistical mechanics. It has several interesting properties. First, for states with energies far below $E_F$, the probability of finding them occupied by electrons approaches unity asymptotically as the energy is reduced: deep down, all states are occupied. For states with energies well above $E_F$, on the other hand, the occupation probability becomes negligible: way up in energy, the states are empty. At $E = E_F$, the occupation probability is exactly 1/2. This is actually a good definition of the Fermi energy at finite temperatures [6]. A second property of the Fermi-Dirac distribution function is its symmetry. You can easily verify that $1 - f(E)$ has a mirror shape to $f(E)$ around $E = E_F$. This means that the probability of finding a state at a certain energy *above* the Fermi level occupied is identical to the probability of having another state at the same energy *below* $E_F$ empty.

A third property to remark about the Fermi-Dirac distribution function is that the transition of $f(E)$ from 1 to 0 around $E_F$ is sharper at lower temperatures. This is better seen in Fig. 1.8 where the Fermi-Dirac distribution function is graphed for various temperatures. At T=0 K, $f(E)$ abruptly jumps from 1 to 0 at $E_F$. As $T$ increases above 0 $K$, the transition becomes softer. The width of the transition region is about *3kT* for a 20% criterion (that is, the energy difference

---

[6] In statistical mechanics a distinction is made between the Fermi energy, which is only defined at T=0 K in the way we did above, and the *chemical potential*, defined at any temperature as the energy at which the occupation probability is $\frac{1}{2}$. In the field of semiconductors this distinction is not usually made and we use the term "Fermi energy" at any temperature. Furthermore, in statistical mechanics, an *electrochemical potential* is also defined in situations in which an electric field is present and a potential energy due to the field is added to every particle. This does not modify the average kinetic energy of an electron and is equivalent to a local shift of the energy distribution. Because of this, this distinction is also typically disregarded in the field of semiconductors.

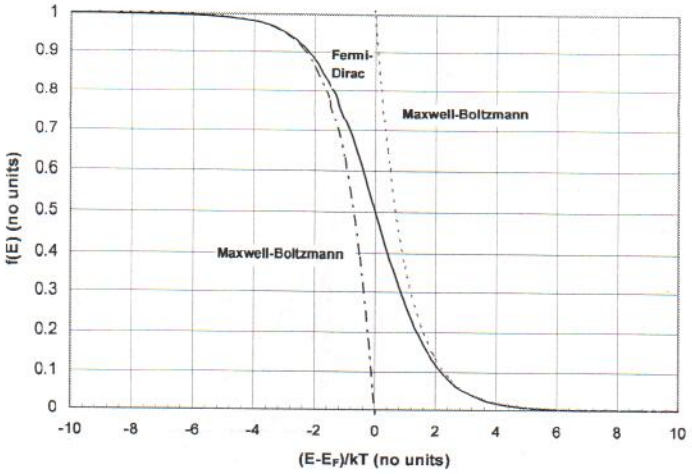


Figure 1.7: Fermi-Dirac distribution function. Approximate Maxwell-Boltzmann distribution functions are also shown.

between $f(E) = 0.8$ and 0.2). This is a handy order of magnitude to keep in mind.

**Exercise 1.4**: *Calculate the probability that a state at 0.1 eV above the Fermi level is occupied: a) at 300K, and b) at 1200K.*

Use Fermi-Dirac distribution function for both parts. For part *a)*:

$$f(E) = \frac{1}{1 + \exp\frac{E-E_F}{kT}} = \frac{1}{1 + \exp\frac{0.1}{0.0259}} = 0.0206$$

This is a very small probability because the state is several $kT$'s above $E_F$ at room temperature.

To do part *b)*, we must first compute $kT$ at $1200K$:

$$kT|_{1200K} = 4kT|_{300K} = 4 \times 0.0259 = 0.104\ eV$$

We now compute the probability of occupation as above:

$$f(E) = \frac{1}{1 + \exp\frac{0.1}{0.104}} = 0.28$$

The probability of occupation is now substantially higher because in the scale of $kT$, the state in question is relatively close to $E_F$.

The properties of the Fermi-Dirac distribution function imply that the Fermi energy in most electronic systems will be a function of temperature. We can visualize the impact of finite temperatures on the electron energy distribution by looking again at the example shown in Fig. 1.6. For a system with such few states and electrons, the statistical fluctuations among different possible configurations are significant. The sketches of Fig. 1.6(b) and (c) represent just one of the possible configurations that can be attained. Note that, at finite temperatures, in contrast with the situation at 0 K depicted in Fig. 1.6(a), a few states below the Fermi energy are empty while others above are occupied. The Fermi energy is located where the probability of occupation is $\frac{1}{2}$. Since in this particular example the number of states increases with $E$, the Fermi level moves

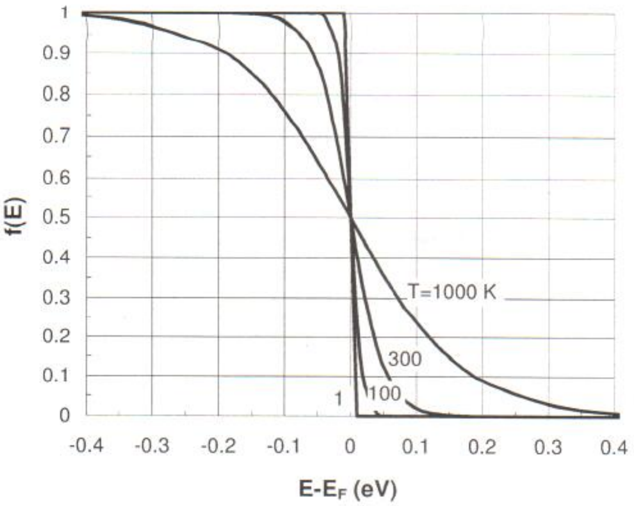


Figure 1.8: Fermi-Dirac distribution function for various temperatures.

to lower energies as the temperature increases, as can be seen by comparing Fig. 1.6 (a), (b) and (c).

It is interesting to examine the form of the Fermi-Dirac distribution function at energies well above $E_F$. In this case, the exponential in the denominator is much larger than 1 and Eq. 1.7 can be approximated by:

$$f(E) \simeq \exp(-\frac{E-E_F}{kT}) \qquad E-E_F \gg kT \tag{1.8}$$

This is called the *Maxwell-Boltzmann distribution function* (also graphed in Fig. 1.7). This approximation tells us that at energies much higher than the Fermi energy, the probability of finding a state occupied by an electron decreases exponentially with a characteristic energy $kT$. The Maxwell-Boltzmann distribution function is an excellent approximation to the Fermi-Dirac distribution function for energies a few $kT$s above $E_F$. For example, for $E-E_F = 2kT$, the error is 13%. If $E-E_F = 3kT$, the error drops to 5%.

Similarly, we can also examine the probability that a given state is occupied at energies much lower than $E_F$. This can be approximated by:

$$f(E) \simeq 1 - \exp\frac{E-E_F}{kT} \qquad E-E_F \ll -kT \tag{1.9}$$

which we can actually rewrite as:

$$1-f(E) \simeq \exp\frac{E-E_F}{kT} \qquad E-E_F \ll -kT \tag{1.10}$$

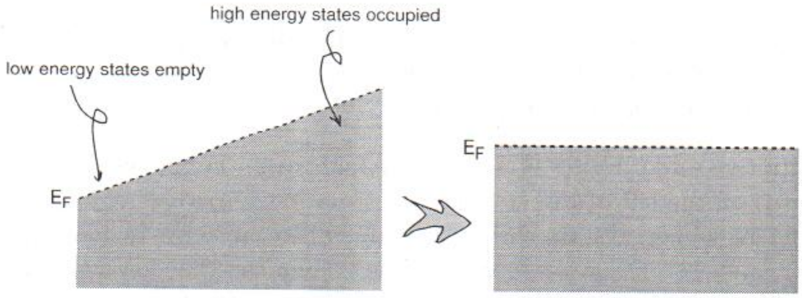


Figure 1.9: A system in thermal equilibrium is characterized by a Fermi energy that is constant everywhere (right). If that was not the case, as in the situation depicted on the left diagram, electrons would flow from a region where high-energy states are occupied to other regions where lower-energy states are empty. In this way, the system attains an overall lower energy situation.

There is a parallel between Eq. 1.10 and Eq. 1.8. Eq. 1.8 describes the probability that a state at an energy substantially above the Fermi energy is full. Eq. 1.10, on the other hand, gives the probability that a state located much below the Fermi energy is *empty*. In both cases the probability decreases exponentially at energies away from the Fermi energy. Furthermore, the characteristic energy of the exponential function is $kT$ in both cases. The symmetry of Eqs. 1.8 and 1.10 arises from the symmetry of the Fermi function around $E_F$, that was mentioned above. Mathematically, it is much easier to use Eqs. 1.8 and 1.10 than the exact Eq. 1.7. Because of this, these two approximations are extensively used in the study of semiconductor devices.

**Exercise 1.5**: *Solve Exercise 1.4 again under the assumption of Maxwell-Boltzmann statistics. Compute the relative error when compared with the result using Fermi-Dirac statistics.*

For part *a)*, Maxwell-Boltzmann statistics yields:

$$f(E) \simeq \exp(-\frac{E-E_F}{kT}) = \exp(-\frac{0.1}{0.0259}) = 0.0210$$

The relative error is:

$$\epsilon = 100 \times \frac{0.21-0.206}{0.206} = 1.9\%$$

The error is small because the distance between the state in question and the Fermi level (0.1 $eV$) is about $4kT$.

For part *b)*, again using Maxwell-Boltzmann statistics results in:

$$f(E) \simeq \exp(-\frac{0.1}{0.104}) = 0.38$$

The relative error is now:

$$\epsilon = 100 \times \frac{0.38-0.28}{0.28} = 36\%$$

This error is significant because the distance between the state and the Fermi level (0.1 $eV$) is just about equal to $kT$ (0.105 $eV$).

In a final note before closing this section, the very definition of Fermi energy has an important and useful consequence. In an electronic system in thermal equilibrium, there is a unique Fermi energy that is *constant throughout the entire system*. In other words, if an electronic system is characterized by a Fermi energy that varies with location, then it is not in equilibrium. This applies regardless of the complexity of the system.

This makes intuitive sense. Consider a closed system with a Fermi energy that changes with position, as depicted on the left of Fig. 1.9. In this situation, there are regions where the Fermi energy is high and high-energy states are occupied. There are also regions where the Fermi energy is low, with empty states at low energies. The system can thus lower its overall energy by allowing electrons that occupy high-energy states to move to low-energy states that are empty. This net flow of electrons inside the system only stops when the Fermi energy becomes constant everywhere, as in the right of Fig. 1.9. This is the true equilibrium condition. The constancy of the Fermi energy in a system is a good working definition of thermal equilibrium itself.

## 1.3 Selected concepts of solid-state physics

Most semiconductor materials used in modern microelectronics belong to the category of crystalline solids (important exceptions are amorphous semiconductors, such as amorphous silicon, widely used in photovoltaics and flat-panel display applications). A perfect *crystalline solid* is a solid with an elemental atomic arrangement, or *unit cell*, that repeats itself *ad infinitum* in the three dimensions. Fig. 1.10 shows a sketch of the atomic arrangement of the unit cell of the two most important semiconductors in microelectronics: Si and GaAs. For both of them, the unit cell has a cubic shape. Si is an *elemental semiconductor* since it is only made of a single kind of atom. GaAs is made out of two, Ga and As, and is thus referred to as a *compound semiconductor*. There are also compound semiconductors with more than two different kinds of atoms, such as InGaAs or InGaAsP. What is common to all these materials is a certain atomic arrangement that repeats itself in all directions in space in a perfect *lattice*. The atomic arrangement in the unit cells depicted in Fig. 1.10 appears rather complicated. In fact, a simple way to visualize it is to realize that it is made of two face-centered cubes ("fcc" or a cube with atoms at its vertices plus in the middle of all faces) with one cube displaced along the diagonal of the other a quarter of its diagonal dimension. For Si, both fcc cubes are made out of Si atoms. For GaAs, one is made out of Ga atoms and the second one is made out of As atoms.

The lateral dimension of the unit cell is called the *lattice constant* (labeled $a$ in Fig. 1.10). For Si and GaAs at room temperature, the lattice constant is 0.543 and 0.565 $nm$, respectively. Since in general, a unit cell contains several atoms, the interatomic distance is typically smaller than the lattice constant. For Si and GaAs, the distance between neighboring atoms is 0.235 and 0.245 $nm$, respectively. This is an important length scale to keep in mind.

The number of atoms per unit volume, or *atomic density*, of a crystalline solid is an important reference scale for volume concentrations. For Si and GaAs, the atomic densities at room temperature are $5.0 \times 10^{22}$ and $4.42 \times 10^{22}$ $cm^{-3}$ respectively (these and other important physical parameters of Si and GaAs at room temperature are summarized in Appendix B at the end of the book).

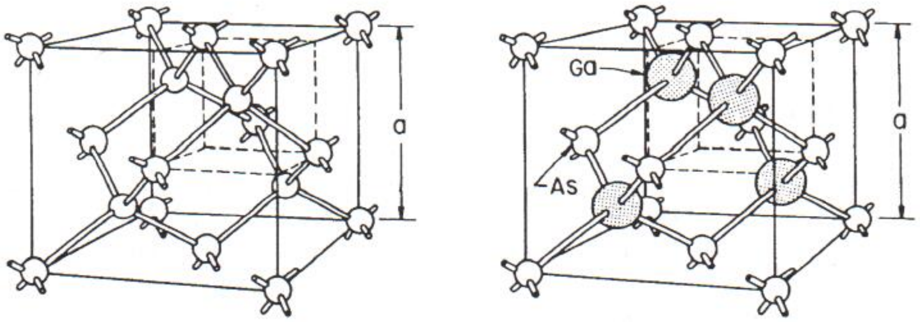


Figure 1.10: Sketch of unit cell of Si (left) and GaAs (right).

The concept of a perfect crystalline solid may appear to be a pure academic fiction. In a real semiconductor crystal, there are going to be unavoidable disruptions of a perfect periodicity such as missing atoms (called *vacancies*) or atoms in wrong locations (for example Si atoms in between lattice sites, called *interstitials*, or Ga atoms in place of As, called *Ga antisite defects*). Furthermore, foreign atoms are always present in a random way in real samples. In addition to this, at finite temperatures, the atoms in a semiconductor are not standing still but are randomly jiggling around, so that strictly speaking, we never have a perfect periodic structure. In addition, the very finite size of a device means that there are surfaces that break perfect atomic periodicity in a drastic way.

These disruptions to the ideal crystalline solid are in most cases just that, disruptions, relatively small perturbations to an otherwise very ideal structure[7]. For example, the spectacular engineering success of modern semiconductor technology is rooted in its ability of producing extremely pure and nearly perfect semiconductor crystals. Si wafers can now be fabricated with foreign atom concentrations below 1 part-per-million, or 1 every $10^6$ Si atoms. Similar levels of crystalline defects can be attained. Actually, one does not need to claim such engineering prowess as a justification for the use of the "perfect" crystal concept. After all, many of the newer semiconductors have not reached the level of purity and crystalline perfection that Si enjoys. And even in Si, as the next chapter presents, selected foreign atoms at concentrations approaching 1 ppm (part per million) to 1% are often deliberately introduced.

The ideal crystalline solid model remains a very useful view of a semiconductor because of the "effective size" of an electron in the lattice. We carried out a first order estimate of this key length in vacuum in a numerical example in Section 1.1.2, about 7.6 $nm$ at room temperature. This figure is not too different inside a semiconductor. Over how many lattice-sites is such an electron spread out? One can get an estimate for this by calculating the number of atoms in a sphere of 7.6 $nm$ diameter. For typical atomic densities this number is over 10,000 atoms. This means then that an electron in a semiconductor is really spread out over many atomic sites and it therefore "averages out" the relatively rare crystalline disruptions.

Hence, the ideal crystalline solid model is a good idealization of a modern semiconductor

---

[7] A surface is not a small perturbation. Special care is needed in its treatment. This is dealt with in Ch. 8.

crystal. The unavoidable disruptions to the perfect periodic lattice are best captured in the form of "collisions" with the electrons, as we will see later in this chapter and in subsequent chapters.

### 1.3.1 Bonds and bands

A typical representation of a semiconductor lattice, such as that of Fig. 1.10, shows atoms connected by "sticks". The sticks represent bonds between atoms. In Si, GaAs and most semiconductors, every atom is bonded to four neighbors. This number is particularly important. The top-most partially occupied subshell in the Si atom contains four empty states that can accommodate four electrons. In an elemental semiconductor such as Si, by binding with four neighbors, a Si atom shares one of its four valence electrons with each of four neighbors. On average, then, a Si atom ends up with a total of eight electrons in its valence shell. In atomic physics, this is known to be a particularly low energy configuration. For instance, most noble gases, which are notoriously unreactive elements, have eight electrons in their outer layer. This type of bonding in which electrons are shared among neighboring atoms is called *covalent bonding*.

For a compound semiconductor such as GaAs, the situation is slightly more complex. Looking at the periodic table in Fig. 1.4, we see that Ga has three electrons in its outer layer while As has five. In a solid in which each Ga atom is surrounded by four As atoms and each As atom is surrounded by four Ga atoms (see Fig. 1.10) each constituent atom shares on average a total of eight electrons in its outer layer. This, again, is a low energy configuration that permits the formation of a stable crystal. In this case however, in addition to covalent, the bonding has a small *ionic character* to it. This is because a Ga site becomes slightly negatively charged while an As site gets slightly positively charged as electrons are shared but protons are not. The resulting electrostatic attractions among these ions further helps to hold the crystal together.

A perfect crystalline solid is a very special electronic system. One can think of it as an entirely new system, but one can also view it as an ensemble of many smaller subsystems, the individual atoms. This mixed approach helps to understand the peculiar energy distribution of electron states in a solid that is sketched in Fig. 1.11. In a crystalline solid, the potential experienced by the electron is periodic in space. This potential largely results from the overlap of the individual potentials associated with each atom. It is also slightly affected by the electron distribution itself.

The core electrons of each atom are very tightly bound to each nucleus. Furthermore, their spatial extent is much smaller than the interatomic distance in the solid. As a result, the core electrons are for the most part unaware of the existence of the solid. The core states largely maintain their atomic character. The atomic bonding only results in a slight broadening in the energies that are allowed. For higher energies, as sketched in Fig. 1.11, the atomic level broadening becomes more significant as the atomic potentials overlap more.

For the valence electrons, the situation is very different. The spatial extent of their wavefunction is larger than the interatomic distance. When dealing with the valence electrons, it is not advantageous to think of a solid in terms of individual atoms: one should view the entire crystal instead as a whole new electronic system. In a crystalline solid one cannot associate a valence electron with a particular pair of atoms. Electrons belong to the crystal as a whole; in a way, they are shared by all atoms. This is contrary to the core electrons that remain associated with

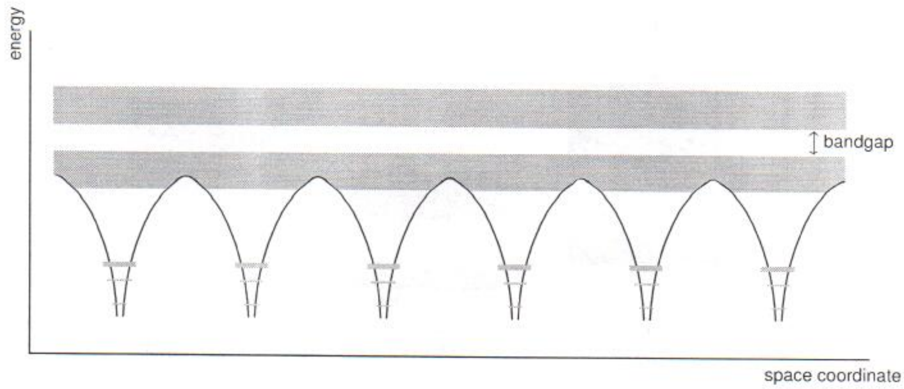


Figure 1.11: Sketch of bands and bandgaps in the electron state distribution of an idealized solid along a certain crystallographic direction. This energy picture can be viewed as an assembly of the Coulombic potentials of the individual atoms.

each atom and are not shared. The quantum states for the valence electrons extend over the whole crystal and Pauli exclussion principle must be satisfied across all these states.

The defining characteristic of a crystalline solid is the perfect periodicity of the electrostatic potential that the electrons experience. This arises from the fact mentioned above that the unit cell repeats itself in space in the three dimensions. Under these "periodic boundary conditions", as they are called, a fundamental result of solid-state physics states that the allowed electron states "cluster" in sharply defined *bands* leaving *bandgaps* of forbidden energies in between. This is represented in Fig. 1.11 by the continuous gray areas that extend over the entire crystal.

### 1.3.2 Metals, insulators, and semiconductors

So far we have been talking about crystalline solids in general. However, we know that there are materials with widely different electronic properties. Where do the differences among metals, semiconductors, and insulators, for example, arise from?

In order to provide a simple qualitative answer to this question, one must consider how the states in a solid get filled with the available electrons. As with an individual Si atom in Section 1.1.3, one starts by filling up the lower energy states and stops when there are no more electrons. Three different situations can arise, as sketched in Fig. 1.12.

First, let us consider a case in which a band ends up partially filled after all electrons are in place (Fig. 1.12a). In this situation there are empty states right above occupied states. As will become clearer in subsequent chapters, in these circumstances electrons can easily move around the crystal in response to an electric field. The conductivity of materials with a partially filled band is high. These are *metals*. Metals exhibit a resistivity in the $10^{-6} - 10^{-4}\ \Omega \cdot cm$ range. In a great analogy from Shockley, this situation resembles a partially filled parking lot where all cars are identical. The availability of empty slots allows any customer to move his car around the lot

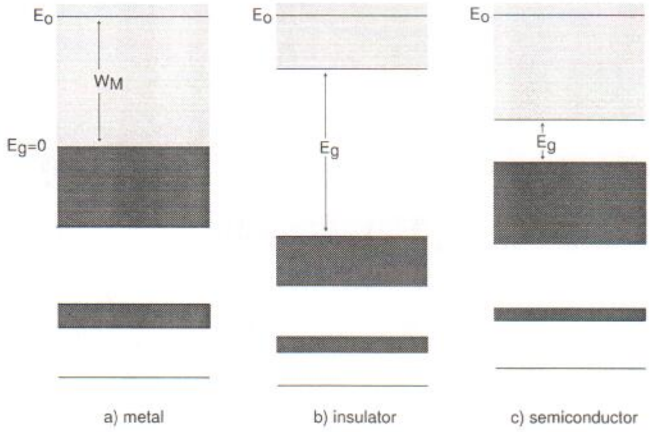


Figure 1.12: Sketch of band occupation for several solids: (a) metal, (b) insulator, and (c) semiconductor. Dark shading represents states occupied by electrons. Light shading represents empty states.

and choose the slot of his pleasing. In response to a stimulus, say heavy snow on the top exposed level, one expects many customers to move their cars to lower lying protected levels.

In a second case, all available electrons precisely fill a band. If a "wide" energy bandgap separates this full band from the next empty band, conduction is impossible because there are no available states into which electrons can easily move. This is an *insulator* and is depicted in Fig. 1.12b. The resistivity of insulators is in the $10^9 - 10^{17}\ \Omega \cdot cm$ range. In Shockley's parking lot analogy, a completely full lot permits cars to swap places, but since all cars are identical one cannot observe any changes in the picture. Quantum mechanics does not allow us to tell electrons apart. If electrons merely trade locations, nothing has really changed.

We can also conceive a third situation in which one ends up with a completely full band that is separated from the next empty band by a "narrow" energy gap (Fig. 1.12c). At zero absolute temperature, this situation is no different from the one just described and the material is perfectly insulating. At finite temperatures, however, the distinction between "wide" and "narrow" energy gap matters. If the bandgap is narrow enough, at high temperatures a few states of the upper band may actually be occupied by electrons, as implied by the Fermi-Dirac distribution discussed in Section 1.2.3. The electrons in the partially filled upper band can move around in response to a stimulus. In addition, there are empty states or *holes* (more about them in the next chapter) in the lower band which now allow electrons to move about. The net result is a moderate conductivity typical of a *semiconductor*. The resistivity of semiconductors is in the $10^{-4} - 10^{4}\ \Omega \cdot cm$ range.

The parking lot analogy for semiconductors requires some elaboration. We should now consider two parking lots next to each other, one more expensive than the other because of some amenities, such as elevators, wider slots, or better lighting. If the price differential is too high, all customers flock to the cheaper one which again gets full. With smarter pricing, some customers will be willing to pay the higher price for the up-scale lot and will park there. This leaves some slots open in the popular lot. In response to a snow storm, customers in both lots will be able to move their cars around since both have slots available.

It is clear from the above discussion that the difference between an insulator and a semiconductor is only qualitative. It really comes down to the width of the energy gap between the last filled band and the next empty band relative to the thermal energy. If the ratio between these two energies is high, very few electrons will reside in the upper band and the material will be highly insulating. If this ratio is low, substantial conductivity is expected, particularly as the temperature is raised. In fact, a pure semiconductor can behave as an insulator at low enough temperatures while an insulator may have non-negligible conductivity at high temperatures. It is perhaps for this reason that semiconductors are named as they are. They never quite match the conductivity of metals, but their conductivity changes (and, as we will see very soon, can be engineered) over many orders of magnitude. Also in contrast with metals, the conductivity of semiconductors *increases* with temperature.

The bandgap that exists at 0 $K$ between the last filled and the next empty band plays such a crucial role in semiconductor physics that it is often called the "fundamental energy gap" (in this book, we will refer to it simply as "the bandgap"). To calibrate our energy scale, most common semiconductors have bandgaps between 0.5 and 2.5 eV while most common insulators have bandgaps of the order of 5 to 10 eV. The bandgaps of Si and GaAs at room temperature are 1.1 and 1.4 eV respectively.

Fig. 1.12 also indicates the *vacuum level* $E_o$. As in the case of the Si atom in Section 1.1.3, this is the energy at which electrons escape from the crystal. For metals, the energy view of Fig. 1.12 provides an intuitive picture for the photoelectric effect and the concept of work function introduced in Sec. 1.1.1. The work function is the minimum energy that needs to be provided to an electron in order for it to escape from the solid. In the energy view of a metal sketched in Fig. 1.12, this is clearly the energy difference between the vacuum level and the highest lying occupied state. In the photelectric effect, only photons with an energy equal or in excess of this difference would be capable of extracting electrons from the metal. For semiconductors and insulators, the concept of work function requires that we know more about the electron distribution in these materials. This is discussed in Ch. 7.

### 1.3.3 Density of states

The energy bands of a solid are composed of individual states that are located in energy very close to each other. At practical temperatures, the energy separation between single states is much smaller than $kT$ and it is of no relevance to try to distinguish individual states. The total number of states present in a given window of energy is a much more important parameter. In this regard, it is useful to define a density of states, $g(E)$, such that $g(E)dE$ is the number of states in a unit volume of material that lie in a window of energy between $E$ and $E + dE$, where $dE$ is a small differential of energy. The units of $g(E)$ are $eV^{-1}cm^{-3}$. [8] In general $g(E)$ depends on energy. A sketch of $g(E)$ vs. $E$ is shown in Fig. 1.13. This picture corresponds to the sketch in real space shown in Fig. 1.11.

In this figure, the core states are characterized by narrow bands with a delta-function density

---

[8] Many important properties of solids scale with the volume of the solid. Because of this, for certain variables, it is convenient to normalize the volume away and define them per unit volume. This will be done extensively in this book.

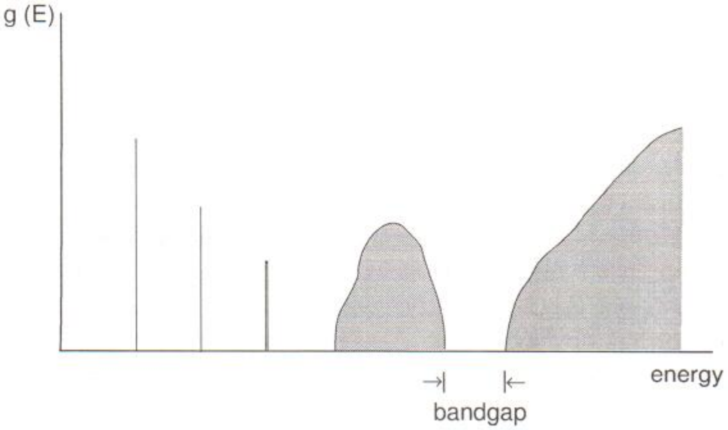


Figure 1.13: Sketch of density of states corresponding to electron state distribution of Fig. 1.11.

of states. States at higher energies give rise to bands with a finite density of state distribution in energy. The bandgap regions are defined by $g(E) = 0$. These regions are "forbidden" to the electrons because there are no states for them to reside in.

The concept of density of states is important. Many semiconductor physical parameters are profoundly affected by the density of states in the bands that surround the fundamental bandgap. We will see several examples of this in the next Chapter.

**Exercise 1.6**: *Calculate the total number of states that lie at the bottom 26 meV of a system with a density of states given by* $g(E) = 1 \times 10^{22}\sqrt{E}\ eV^{-1} \cdot cm^{-3}$.

The number of states that lie between energies $E_1$ and $E_2$ in a certain state distribution is simply given by the integral:

$$N(E_1, E_2) = \int_{E_1}^{E_2} g(E)dE$$

The units of $N(E_1, E_2)$ are $cm^{-3}$.

In the present example, if we denote as $C$ the factor in front of $\sqrt{E}$, we must perform the integral:

$$N(0, E) = \int_0^E C\sqrt{E}\ dE = \frac{2}{3}CE^{3/2}\bigg|_0^E = \frac{2}{3}CE^{3/2}$$

Putting numbers in this expression, we obtain:

$$N(0, 0.026\ eV) = \frac{2}{3} \times 1 \times 10^{22}\ eV^{-3/2} \cdot cm^{-3} \times (0.026\ eV)^{3/2} = 2.8 \times 10^{19}\ cm^{-3}$$

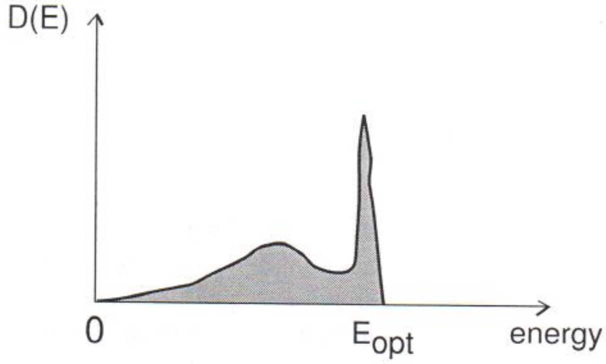


Figure 1.14: Sketch of the phonon spectrum of a typical semiconductor. $D(E)$ is the density of modes per unit energy per unit volume.

### 1.3.4 Lattice vibrations: phonons

Atoms in a solid also possess thermal energy. At ordinary temperatures, the atoms which constitute a solid vibrate around their equilibrium positions in the lattice. In semiconductors at room temperature, the average amplitude of the vibration is of the order of a few hundredths of a *nm*. Since this is a sizable fraction of the interatomic distance, lattice vibrations affect the material properties of the semiconductor.

The lattice of a solid can be viewed as a network of atoms connected together by springs. At absolute zero temperature, the atoms will be fixed in space at the equilibrium points of these springs. At finite temperatures, they will be oscillating around these equilibrium positions. An individual atom cannot oscillate in complete ignorance of what its neighbours are doing. This is because the equilibrium position of an atom in the lattice is set by the balance between the attractive and repulsive Coulombic forces that it experiences with its neighbours. A solid is thus a large *coupled* system.

In a solid, the lattice can take many different vibrational modes. This is similar to the membrane of a drum or a string in a piano. Each mode is characterized by its wavelength and its mechanical energy. At any one time, a multiplicity of modes will be active with various energies. The resulting vibration pattern can be very complex.

The lattice can exchange energy with free electrons in the solid. Energy exchange can go two ways. An electron can give energy to the lattice or it can receive energy from it. If the electron gives some energy to the lattice, it must *excite* an available vibrational mode and provide the precise amount of energy required by that mode. When the lattice gives energy to the electron, a certain vibrational mode is extinguished and its mechanical energy is transferred to the electron. The quantized nature of the vibrational energy of the lattice makes it particularly convenient to visualize the vibrational modes as particles. They are called **phonons**. Energy exchanges between electrons and the lattice can be viewed as collisions between electrons and phonons. When an electron donates energy to the lattice, a phonon is "emitted." In the reverse process, a phonon is "absorbed."

The phonon energy spectrum, that is, the density of modes per unit energy per unit volume, plays an important role in many properties of semiconductors. Examples are their electrical and thermal resistivity. It is determined by the nature of the atoms that make the solid and the bonding arrangement among them. A typical phonon spectrum for a semiconductor is sketched in Fig. 1.14. There are some aspects to it that we need to appreciate.

In a typical semiconductor a continuum of phonons is available all the way down to zero energy. These low-energy phonons are called *acoustical phonons.* Comparatively, there are not very many of such phonons. There is also a maximum energy beyond which no phonons exist. Interestingly, there are many phonons at this maximum energy which are associated with very peculiar lattice vibrations. They are called *optical phonons* [9] and their energy is denoted as $E_{opt}$. $E_{opt}$ is an important landmark in the energy scale of relevance to semiconductors devices. In Si, $E_{opt}$ is 63 meV while in GaAs, its value is 35 meV. We will explore the significance of these numbers later in this book. At this time, let us note that $E_{opt}$ is slightly larger than $kT$ at room temperature but is much smaller than the bandgap of typical semiconductors.

It is interesting to compare phonons against photons, and their interactions with electrons in a semiconductor. Phonons are particles with little energy and a large momentum. In contrast, photons are particles with very small momentum but relatively high energy. When an electron collides with a phonon, the electron momentum can change significantly but not its energy. In contrast, when an electron collides with a photon, the momentum of the electron is not affected but its energy can change substantially.

In thermal equilibrium, there is no net energy exchange between the lattice and the electrons. On average, an equal number of phonons are emitted and are absorbed at every energy. In other words, the electron "gas" is in equilibrium with the lattice.

## 1.4 Summary

- The electron has a dual particle-wave nature.
- The Pauli exclusion principle imposes restrictions to the dynamics of electrons. In an isolated atom, for example, certain electrons are tightly confined to the vicinity of the nucleus (core electrons) while others reside in much looser orbits (valence electrons).
- Concept of Fermi energy: at zero absolute temperature, a gas of electrons occupies all lowest energy available states up to the Fermi energy. At finite temperatures, the transition between occupied and empty states is gradual, with a softness described by a key characteristic energy, the thermal energy $kT$. The Fermi level is defined as the energy for which electron occupation probability is 0.5.
- A system of particles is in thermal equilibrium if it does not exchange energy with the outside world (it is closed) and if the time derivatives of all local and global ensemble

---

[9] *Optical phonons* are given such a name because their vibrational modes can be excited by the electric field of a light wave. In contrast, the vibrational modes of *accoustical phonons* cannot be excited by light.

averages are zero (it is in steady state). A system in thermal equilibrium has a Fermi level that is constant throughout.

- Electron occupation probability in thermal equilibrium is given by Fermi-Dirac distribution function:

$$f(E) = \frac{1}{1 + \exp\frac{E-E_F}{kT}}$$

For $E \gg E_F$, or $E \ll E_F$, the Fermi-Dirac distribution function can be approximated by the Maxwell-Boltzmann distribution function:

$$f(E) \simeq \exp(-\frac{E-E_F}{kT}) \qquad E - E_F \gg kT$$

$$1 - f(E) \simeq \exp\frac{E-E_F}{kT} \qquad E - E_F \ll -kT$$

- A crystalline solid imposes periodic boundary conditions to the electrons. A consequence of this is that electron energy states cluster in bands leaving bandgaps forbidden to electrons.

- A semiconductor is a solid in which at 0 $K$, electrons completely fill a band leaving a bandgap between the highest energy filled state and the next available empty state. Additionally, the width of this fundamental energy gap is such that at room temperature there are a few electrons in the band just above the bandgap.

- A semiconductor lattice vibrates in certain modes. Each mode can be characterized by a phonon or quantum of mechanical energy. Electrons in a semiconductor can exchange energy with the lattice by emitting or absorbing phonons. Typically, the maximum energy of a phonon is a few tens of meV, which is much smaller than the bandgap energy. Electrons can also exchange energy with photons or quanta of light.

## 1.5 Further Reading

**The Feynman Lectures on Physics** by R. P. Feynman, R. B. Leighton, and M. Sands, Addison-Wesley, 1964 (ISBN 0-201-02116-1-P, QC23.F435) have excellent reading material on elemental quantum mechanics and statistical mechanics. The style of this classic series emphasizes physical intuition. Chapters 37 and 38 of Vol. I are a great introduction to quantum mechanics, the wave-particle duality, electron behavior, and atomic structure. Chapters 39, 40, and 41 also of Volume I deal with the kinetic theory of gasses, the principles of statistical mechanics and Brownian motion. Chapter 41, in particular, brings together results of quantum mechanics and statistical mechanics to derive the physics of blackbody radiation. Highly recommended reading.

**Modern Physics and Quantum Mechanics** by E. E. Anderson, Saunders, 1971 (ISBN 0-7216-1220-2, QC174.1.A527) is an excellent textbook on introductory quantum mechanics. It represents a good balance between physical intuition and mathematical rigor. Several chapters are organized with a sense of history in them which makes them exciting and informative. The book has many examples, lots of charts and pictures, displays data from original experiments and is in general well written and clear. Chapters 2-4 are particularly relevant to the discussion of interest to us here.

**Introduction to Solid State Physics** by C. Kittel, Wiley, 1976 (ISBN 0-471-49024-5, QC176.K5) is also a classic textbook. Chapters worthy of reading at this point are Ch. 1 on crystal structure, Ch. 3 on crystal binding, and Ch. 6 on electron statistics. You may also want to browse through Ch. 7 on energy bands and Ch. 4 on lattice vibrations and phonons. The book is rather rigorous mathematically, yet physically intuitive. If you have not taken a course on solid-state theory, you will have to make an effort to uncover the physics out of the mathematical clutter.

**Crystal Fire - The Birth of the Information Age** by M. Riordan and L. Hoddeson, Norton, 1997 (ISBN 0-393-04124-7, TK7809.R56) is an enjoyable and easy to read book telling the story of the invention of the transistor and the integrated circuit. Its first few chapters presents an excellent description of the puzzle that physicists faced at the beginning of the 20th century and the resulting development of quantum mechanics. The book covers in simple and intuitive terms many of the topics discussed in the present Chapter.

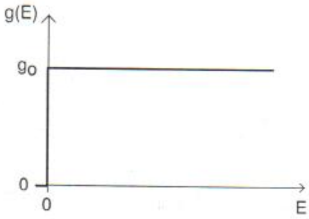
**History of Semiconductor Engineering** by B. Lojek, Springer, 2007 (ISBN 3-540-34257-5, TK7871.85.L65) is also a fun to read account of the early days of the semiconductor industry. There is a lot of detail about the beginnings of integrated circuits and the personalities involved. It includes many historical pictures and diagrams.

## Problems

**1.1** Compute the kinetic energy in $eV$ and the de Broglie wavelentgh in $cm$ of a) a 57 $gr$ tennis ball served at 160 km/hr, b) a 3000 pound car traveling at 65 miles/hr, and c) an electron in vacuum moving at $10^9$ cm/s. Comment on the suitability of the particle-type description of these objects if they are, respectively, in a) a tennis court that is 78 feet by 36 feet, b) a road 15 m wide, and c) a particle accelerator with a 1-inch diameter core.

**1.2** Calculate the wavelength and the frequency of a photon with the lowest possible energy required to free up a $1s$ core electrons from a Si atom.

**1.3** Consider a simple quantum mechanical system with an electron concentration $n$ and a uniform density of states $g_o$ as sketched below.



Calculate:

**a)** The position of the Fermi level at 0 $K$

**b)** The average energy per electron at 0 $K$.

**c)** The position of the Fermi level at a finite temperature $T$. Derive first an exact expression, then simplify it for the case in which $E_F \gg kT$.

**d)** The *increase* in average energy per electron as the temperature is raised from 0 to $T$, defined in the following way:

$$< \Delta E >= \frac{1}{n} \int_0^\infty E \; g(E)[f(E,T) - f(E,0)] \, dE$$

where $f(E,T)$ and and $f(E,0)$ are the Fermi-Dirac distribution function at temperature $T$ and at zero temperature, respectively (assume $E_F \gg kT$).

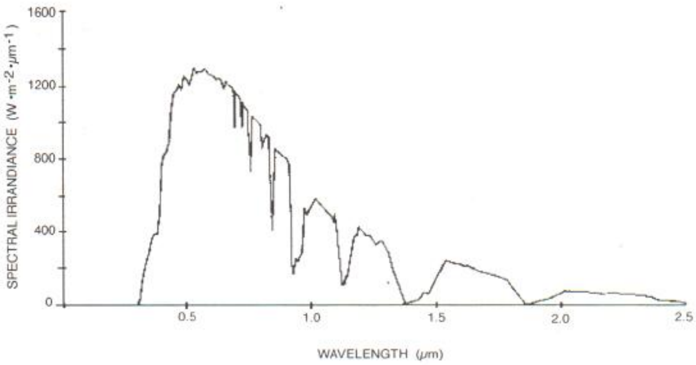
Mathematical help:

$$\int_0^\infty \frac{x dx}{1+e^x} = \frac{\pi^2}{12}$$

Also, when doing the above integral, do not break it through the minus sign inside the square brackets, but rather break it in energy at $E_F$.

**1.4** An **atom laser** is a coherent beam of atoms. Its recent demonstration (see, for example, http://cua.mit.edu/ketterle_group/home.htm) represents a spectacular manifestation of matter waves. An important step towards the realization of the atom laser was the achievement of Bose-Einstein condensation. This is a cold and coherent condensate of atoms. The atom laser is realized when atoms are extracted in a coherent way from the Bose-Einstein condensate. The proof of coherence was obtained by observing an interference pattern when two Bose-Einstein condensates overlapped. Let's put some numbers to this. In its first demonstration, Na atoms were used. The diffraction pattern that was observed had a wavelength of 15 $\mu m$. *Estimate* the temperature at which the atoms were cooled down to for this to be possible? For comparison, *estimate* the de Broglie wavelength of a Na atom at room temperature. At all temperatures, consider the system of Na atoms as an ideal gas.

**1.5** The spectral irradiance of the sun under standard conditions at the Earth's surface is shown in the Figure below. The peak of the radiation occurs at a wavelength of 0.6 $\mu m$ where the spectral irradiance is about 1300 $W \cdot m^{-2} \cdot \mu m^{-1}$.



a) Compute the spectral density of the photon flux at the surface of the Earth at a wavelength of 0.6 $\mu m$.

b) Compute the photon flux at the Earth's surface for photons of wavelength between 0.59 and 0.61 $\mu m$ wavelength.

c) What is the maximum bandgap that a semiconductor can have and still be able to absorb 0.6 $\mu m$ wavelength photons?

**1.6** Starting from the atomic density of Si given in Appendix B at the end of the book, *estimate* the interatomic distance of Si atoms in a Si crystal in two ways. First, assume that the solid is made out of closely packed hard spherical balls that represent the Si atoms. Second, assume that the each atom can be considered as a hard cube. Compare with the actual value of the interatomic distance given in Appendix B.

**1.7** This problem is about estimating the size of a Hydrogen atom and the binding energy of its electron. This is an easy calculation and the result is surprisingly accurate. This is how this can be done. Consider the electron as a classical particle that is bound to the proton by the attractive electrostatic force. The electron performs a circular orbit around the proton. The quantum mechanics are introduced by assuming that the length of the orbit is equal to the de Broglie wavelength of the electron. The solution to the problem lies in computing the total energy of the electron, the sum of its kinetic energy plus its potential energy, and finding the radius that minimizes it.

Proceed as follows.

a) Using Eqs. 1.3 and 1.4, express the kinetic energy of the electron in terms of its de Broglie wavelength. Assume that the de Broglie wavelength is equal to the circumference of the orbit and derive an expression for the kinetic energy in terms of the orbital radius.

b) From elemental electrostatics, write an expression for the potential energy of the electron in terms of the radius of its orbit. Get the total energy of the electron. Find the radius that minimizes it. Derive a simple expression for the total energy of the electron in terms of fundamental parameters.

c) Put numbers to these expressions. Use the SI system. Give the final result in $nm$ and $eV$.

**1.8** Derive an expression for the relative error of Maxwell-Boltzmann statistics with respect to Fermi-Dirac statistics as a function of energy. Graph in a quantitative fashion.

**1.9** Atoms have been guided through hollow optical fibers (see M. J. Renn *et al.*, *Phys. Rev. Lett.* **75**, 3253, 1995). This represents a potentially convenient and flexible method for manipulating atoms. If an atom's de Broglie wavelength is comparable to the diameter of the core, the atom will propagate like a wave, as opposed to a particle. This could be used to make an atom fiber interferometer. The authors of the paper state that Rb atoms that have been cooled down to 290 nK and are launched into a fiber with a hollow core of 2 $\mu m$ in diameter would travel in a single transverse atomic mode. Verify this statement by computing the de Broglie wavelength of Rb atoms in this situation. The atomic mass number of Rb is 85.5.

**1.10** A certain state has a probability of 15% of being empty at 300K. What is its location with respect to the Fermi level?

**1.11** Planck's radiation law gives the frequency distribution of energy radiated by an ideal black body. It is given by:

$$E(\nu) = \frac{2\pi\nu^2}{c^2}\frac{h\nu}{\exp\frac{h\nu}{kT} - 1}$$

In this equation, $h$ is Planck constant, $k$ is Boltzmann constant, $c$ is the speed of light in vacuum, $T$ is absolute temperature, and $\nu$ is the frequency of radiation. This equation is given in the SI system of units where $E$ has units of $W \cdot s/m^2$.

Calculate the photon spectral density (photon flux per unit frequency) emitted by a blackbody at room temperature at an energy equal to the bandgap of Si at room temperature.