CHAPTER 7

Electrons in Solids: Electrical and Thermal Properties

7.1 Introduction

The electrical properties of materials are determined largely by the response of the electrons to external fields. The central theme of this chapter is electrical conductivity and how it varies across different classes of materials. The classical Drude model and how it attempts to explain conductivity and the Hall effect form the starting point. While providing an explanation of Ohm's law and Joule heating, Drude theory leaves open the question of why there is such a large difference in conductivity between different elements of the periodic table.

The Sommerfeld theory for the free-electron gas is then introduced. This treats the electrons quantum mechanically but ignores the presence of ions. It is applied to the calculation of the specific heat and the thermopower.

The focus then changes to the quantum theory of crystalline solids. Bloch's theorem is derived. A quantum-mechanical analysis is given in several limiting cases, such as the nearly free electron case and the tight-binding case. A qualitative discussion of the differences among metals, insulators, semiconductors, and semimetals is presented. Expressions for the density of electronic states are derived. The quantum theory of solids is then applied to a study of the temperature dependence of the resistivity of metals.

Following this is a section devoted to semiconductors. The valence and conduction energy bands are defined and the bandgap energy and effective-mass tensors are introduced. An application of band theory is made to calculation of the magnetoresistance of semiconductors.

Other phenomena occurring in semiconductors and insulators are considered. Variable-range hopping is discussed from the viewpoints of Mott and Efros-Shklovskii. Electrical conductivity in strong electrical fields is considered in the section on the Poole–Frenkel effect.

Some materials exhibit a radical change in conductivity as physical parameters are varied. In this chapter we discuss granular metals that are insulators until the concentration of the metallic inclusions is high enough to cause a percolation transition.

The remainder of the chapter concerns conduction in reduced-dimensional spaces. A discussion of conduction in carbon nanotubes based on the tight-binding method is presented. This is followed by the Landauer theory of one-dimensional conductance.

Additional topics covered at our Web site[†] include further material on the Onsager relations, the random tight-binding approximation, the Kronig–Penney model, the Hall effect in band theory, electron localization, and the evaluation of Fermi integrals.

CLASSICAL THEORY OF ELECTRICAL CONDUCTION

When a constant electric field is established in some materials, current flows. Assuming a local relationship, one writes $\mathbf{J}(\mathbf{r}) = \sigma\mathbf{E}(\mathbf{r})$, where $\mathbf{J}$ is the current per unit area (current density), σ the electrical conductivity and $\mathbf{E}$ the electric field. For an isotropic medium, σ is a scalar and the vectors $\mathbf{J}$ and $\mathbf{E}$ are parallel. For an anisotropic material σ becomes a tensor and these vectors need no longer be parallel. The microscopic form of Ohm's law states that σ is independent of the electric field. Its magnitude determines whether the material is a conductor (high σ), an insulator (low σ), or a semiconductor (intermediate σ). The difference in the value of σ among materials can be enormous, varying by more than 20 orders of magnitude in going from a good conductor to a good insulator. At first sight it is a mystery how, simply by changing the atomic number a few units in the periodic table, one can observe such a huge variation in a physical parameter. The answer to the mystery will involve understanding how tightly (relative to k_BT) the valence electrons are bound to the atoms and ions. We begin this section by studying metals and then proceed to semiconductors and insulators.

7.2 Drude Theory

An early attempt at a theory of Ohm's law was made by Drude, who applied Newtonian mechanics to study the motion of electrons through a metal. The electrons that are not tightly bound to atoms or ions (i.e., the "conduction" electrons), are accelerated by an applied electric field and are assumed to collide with "scatterers," which deflect them and randomize their velocities to a thermal distribution. Consider a sequence of collisions labeled by the index j. Let the velocity of an electron just after the jth collision be $\mathbf{v}_j$. Between collisions j and $j+1$ the velocity is

$$\mathbf{v}(t) = \mathbf{v}_j - \frac{e\mathbf{E}}{m}(t - t_j), \tag{7.1}$$

since the acceleration is given by $-e\mathbf{E}/m$. The probability of surviving to time t without making a collision and then making a collision between times t and $t + dt$ is

$$dP = e^{-(t-t_j)/\tau}\frac{dt}{\tau}, \tag{7.2}$$

where τ is the mean time between collisions. The mean velocity between times t_j and t_{j+1} is

$$\langle \mathbf{v}_j(t) \rangle = \int \mathbf{v}(t)dP = \int_{t_j}^{\infty} \left[\mathbf{v}_j - \frac{e\mathbf{E}}{m}(t - t_j)\right] e^{-(t-t_j)/\tau}\frac{dt}{\tau}. \tag{7.3}$$

[†] Supplementary material for this textbook is included on the Web at the resource site (ftp://ftp.wiley.com/public/sci_tech_med/materials). Cross-references to elements of the Web material are prefixed by "W."

The direction of $\mathbf{v}_j$ is random and so will average to zero for all the conduction electrons. The term $-e\mathbf{E}\tau/m$ is the *drift velocity* $\langle\mathbf{v}\rangle$. The average current density is

$$\mathbf{J} = -ne\langle\mathbf{v}(t)\rangle = \frac{ne^2\tau}{m}\mathbf{E} = ne\mu\mathbf{E} \tag{7.4}$$

where n is the number of conduction electrons per unit volume. The parameter $\mu = \langle v\rangle/E = e\tau/m$ is called the *mobility* of the electron. Thus the *Drude conductivity formula* is obtained:

$$\sigma = \frac{ne^2\tau}{m}. \tag{7.5}$$

The *Joule heating formula* may be derived similarly by examining the kinetic energy loss in a typical collision:

$$\Delta K = \frac{m}{2}\mathbf{v}_{j+1}^2 - \frac{m}{2}\left[\mathbf{v}_j - \frac{e}{m}\mathbf{E}(t_{j+1} - t_j)\right]^2. \tag{7.6}$$

On the average $\langle\mathbf{v}_{j+1}^2\rangle = \langle\mathbf{v}_j^2\rangle$ and $\langle\mathbf{v}_j\rangle = 0$, so

$$\langle\Delta K\rangle = -\frac{e^2\tau^2E^2}{m}. \tag{7.7}$$

Since the mean time between collisions is τ, the power produced per unit volume is

$$\mathscr{P} = -n\frac{\langle\Delta K\rangle}{\tau} = \sigma E^2. \tag{7.8}$$

Use has been made of the integral $\int t^2\,dP = 2\tau^2$, with dP given by Eq. (7.2). This power is dissipated as heat through the production of phonons or other excitations of the metal.

Representative values for the dc electrical conductivity of metals are given in Table 7.1. The electron density, n, is computed from the relation $n = N_A v\rho_m/A$, where

TABLE 7.1 Parameters of Some Metals at $T = 295$ K

Metal	Atomic Number A	Valence z	Mass Density ρ_m (10^3 kg/m^3)	Electron Density n (10^{29} m^{-3})	Conductivity σ $[10^6(\Omega\cdot\text{m})^{-1}]$	Collision Time τ (10^{-15} s)
Ag	107.9	1	10.5	0.585	62.1	37.6
Al	26.98	3	2.70	1.81	36.5	7.17
Ba	137.3	2	3.59	0.315	2.6	2.93
Be	9.012	2	1.82	2.43	30.8	4.50
Ca	40.08	2	1.53	0.460	27.8	21.5
Cd	112.4	2	8.65	0.927	13.8	5.29
Cs	132.3	1	2.00	0.091	5.0	19.5
Cu	63.55	1	8.93	0.846	58.8	24.7
In	114.8	3	7.29	1.15	11.4	3.53
K	39.10	1	0.91	0.140	13.9	35.2
Li	6.939	1	0.54	0.469	10.7	8.11
Mg	24.31	2	1.74	0.862	22.3	9.18

N_A is Avogadro's number, z the valence, ρ_m the mass density, and A the atomic weight. Note that the SI unit for conductivity is $(\Omega \cdot \mathrm{m})^{-1}$. The values of the collision times computed from Eq. (7.5) are displayed. The ac conductivity of metals is covered in Section 8.3.

Without collisions σ and τ are infinite and the metal is a perfect conductor. This state may actually be achieved in superconductors, the subject of Chapter 16.

7.3 Hall Effect in Metals

A conductor that carries an electric current in the presence of a transverse magnetic field develops a potential difference across the sample perpendicular to both the current and the magnetic field. This is the *Hall effect*, and the potential difference produced is the Hall voltage, V_H. By measuring V_H it is possible to measure directly the product nq, the mobile carrier density multiplied by the charge of the carrier.

Figure 7.1 depicts a situation in which a voltage V is impressed across a rectangular parallelipiped of metal of dimensions L by w by h (assume that $L \gg w$ and $L \gg h$). At steady state a current I exists and a longitudinal current density $J = I/hw$ is established in response to the electric field $E = V/L$. Ohm's law relates the two quantities: $J = \sigma E$.

From Eq. (7.4) it follows that the magnitude of the current density is proportional to the drift velocity of the carriers, $\langle v \rangle$ (i.e., $J = nq\langle v \rangle$). Figure 7.1 is drawn for the case of carriers of negative charge, (e.g., electrons). A magnetic force on the carrier, $F_B = q\langle v \rangle B$, acts vertically downward. If this were the only force acting, negative charge would accumulate on the lower face, and to preserve charge neutrality, positive charge would collect on the upper face. These charge sheets create the Hall electric field E_H, which is directed downward. Hence there is an upward electrostatic force of magnitude $F_H = qE_H$ which will continue to grow until F_H and F_B equilibrate, giving $E_H = \langle v \rangle B$. The resulting Hall voltage is $V_H = E_H h$. Thus, finally, an expression for nq in terms of macroscopically measurable quantities is obtained:

$$nq = \frac{IB}{wV_H} \equiv -\frac{1}{R_H}, \tag{7.9}$$

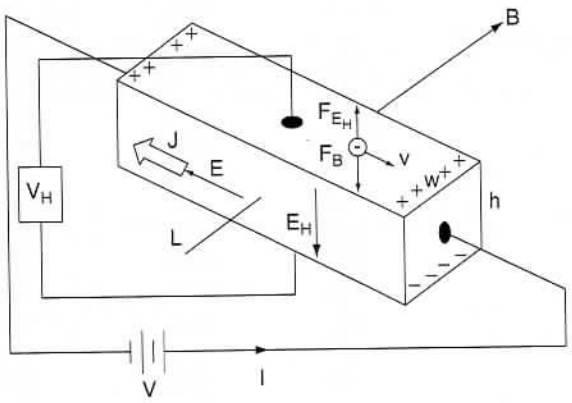


Figure 7.1. Geometry of a Hall effect measurement.

TABLE 7.2 Comparison of Measured Hall Coefficients with the Free-Electron Theory Prediction[a]

Metal	$-1/R_H ne$
Li	0.78
Na	0.99
K	1.00
Rb	1.08
Cs	1.11
Ag	1.19
Cu	1.37
Au	1.48
Al	−1.00/3
In	−1.00/3

[a]If the prediction were correct, $-1/R_H ne = 1$.

where R_H is called the *Hall coefficient*. Also, from the formulas $E = V/L$ and $I = V/R$, one obtains an expression for the resistance:

$$R = \frac{L}{\sigma h w}, \tag{7.10}$$

which is independent of the strength of the magnetic field.

A comparison of measured values of R_H with the free-electron parameters is given in Table 7.2. For the alkali metals the quantitative agreement is reasonable. It is poorer for other good conductors and disagrees substantially in magnitude (and sign!) for In and Al. It is curious that if one allowed In and Al to have valence −1 rather than +3, the agreement with the free-electron theory would be restored. The theory is further frustrated by the observation that both R_H and ρ are often found to depend on the magnetic field. A proper accounting of these anomalies must await the quantum treatment of metals.

FREE-ELECTRON GASES

7.4 Sommerfeld Theory

Sommerfeld proposed a simple model treating a metal as a free-electron gas confined by the surfaces of the solid. The electrons, however, are treated using quantum mechanics. In the bulk of the material, the surface is disregarded altogether and the electrons are completely free. The free-electron model works best for low-valence metals, such as the alkalis (Li, Na, K, Rb, Cs), although it is often employed for others as well, including such valence 3 metals as Al. The ion-core potential is strongly screened by the valence electrons and the ions are unable to bind the valence electrons. (Note the self-consistency.) The valence electrons are free to wander about the solid. Recalling

the fact that the average electric field in a conductor at equilibrium is zero, one can argue that the same screening effect acts to diminish the electron–electron interaction. Thus an independent-particle picture may be employed, and each electron is imagined to interact only with the constant background potential of the solid.

Quantum mechanics enters in two important ways. First, it defines the allowed eigenstates that the electrons can occupy. Second, since electrons obey Fermi–Dirac statistics, the Pauli exclusion principle applies (i.e., at most one electron may occupy a given eigenstate). These eigenstates are characterized by their wave vector and spin projection. The single-particle Hamiltonian is simply the free-electron kinetic energy operator, since the potential energy $V(\mathbf{r})$ is chosen by Sommerfeld to be zero,

$$H = \frac{p^2}{2m} = -\frac{\hbar^2}{2m}\nabla^2. \tag{7.11}$$

The eigenfunctions are plane waves multiplied by a two-element column vector specifying the spin state

$$\psi_{\mathbf{k},s}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}\chi_s. \tag{7.12}$$

The two independent spin states (up and down) are described by the column vectors (spinors)

$$\chi_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \chi_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{7.13}$$

The energy is independent of the spin and grows with increasing k:

$$E_{\mathbf{k},s} = \frac{\hbar^2 k^2}{2m}. \tag{7.14}$$

Suppose that there are N free electrons in a volume V. At $T = 0$ K, the ground state is obtained by filling the N lowest-lying energy levels. This implies a maximum value for occupied k's, called the *Fermi wave vector*, k_F. It is defined by

$$\begin{aligned} N &= \sum_{\mathbf{k},s} \Theta(k_F - k) = 2\int \frac{V d^3k}{(2\pi)^3}\Theta(k_F - k) \\ &= \frac{2V}{(2\pi)^3}\int_0^{k_F} 4\pi k^2 dk = V\frac{k_F^3}{3\pi^2}. \end{aligned} \tag{7.15}$$

Here the unit step function $\Theta(k_F - k)$ imposes the restriction $(k < k_F)$. Introducing the electron number density $n = N/V$, one finds that

$$k_F = (3\pi^2 n)^{1/3}. \tag{7.16}$$

The electrons have a range of energies extending from zero up to the Fermi energy, defined by

$$E_F = \frac{\hbar^2 k_F^2}{2m}. \tag{7.17}$$

The Fermi velocity is defined by $v_F = \hbar k_F/m$ and the Fermi temperature by $T_F = E_F/k_B$. The total electron energy is computed by integrating the kinetic energy over

the Fermi sphere:

$$U = \sum_{\mathbf{k},s} E_{\mathbf{k},s}\Theta(k_F - k) = 2\int \frac{Vd^3k}{(2\pi)^3}\frac{\hbar^2k^2}{2m}\Theta(k_F - k) = \frac{3}{5}NE_F, \tag{7.18}$$

from which it is seen that the average electron energy at $T = 0$ K is 60% of the Fermi energy. Some typical values of k_F, v_F, E_F, and T_F are presented in Table 7.3. Figure 7.2 depicts the Fermi sphere in **k** space.

At absolute zero the occupancy of any given state is either 1 or 0. At finite temperatures the occupancy is given by the Fermi–Dirac distribution $f(E_k, T) = [\exp(\beta(E_k - \mu)) + 1]^{-1}$ (see Appendix WB at our Web site) where $\beta = 1/k_BT$ and

TABLE 7.3 Free-Electron Parameters for Various Metals

Metal	k_F (10^{10} m^{-1})	v_F (10^6 m/s)	E_F (eV)	T_F (10^3 K)
Ag	1.20	1.39	5.49	63.7
Al	1.75	2.03	11.7	135
Ba	0.977	1.13	3.64	42.2
Be	1.93	2.23	14.2	165
Ca	1.11	1.28	4.68	54.3
Cd	1.40	1.62	7.47	86.6
Cs	0.646	0.748	1.59	18.4
Cu	1.36	1.57	7.03	81.5
In	1.50	1.74	8.63	100
K	0.745	0.863	2.12	24.6
Li	1.12	1.29	4.74	55.0
Mg	1.37	1.58	7.11	82.5
Na	0.922	1.07	3.24	37.6
Rb	0.698	0.808	1.86	21.5
Sr	1.02	1.18	3.94	45.7
Zn	1.57	1.82	9.40	109

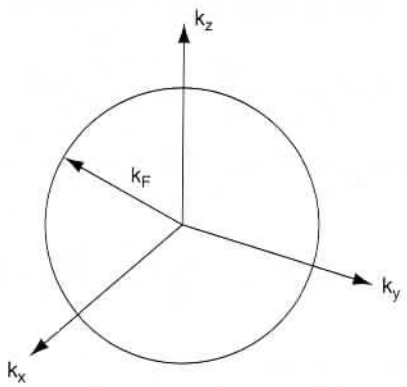


Figure 7.2. Fermi sphere in **k** space.

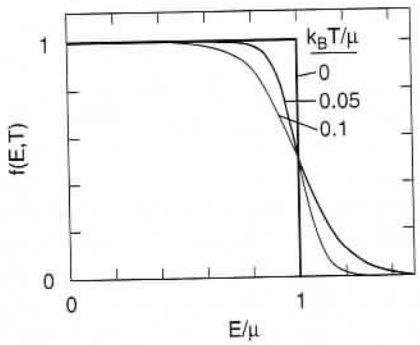


Figure 7.3. Fermi–Dirac distribution function $f(E, T)$ plotted for several values of T. Both $k_B T$ and E are given in units of μ, the chemical potential.

μ is the chemical potential of the electrons. A graph of this function is presented in Fig. 7.3 for several values of T. The size of the step width is seen to be on the order of $k_B T$. In place of Eq. (7.15), there is

$$N = \sum_{\mathbf{k},s} f(E_k, T) = 2 \int \frac{V d^3k}{(2\pi)^3} \frac{1}{e^{\beta(E_k - \mu)} + 1}$$
$$= \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty \frac{E^{1/2}}{e^{\beta(E-\mu)} + 1} dE. \tag{7.19}$$

In place of Eq. (7.18), the internal energy is now given by

$$U = \sum_{\mathbf{k},s} E_{\mathbf{k},s} f(E_k) = 2 \int \frac{V d^3k}{(2\pi)^3} \frac{\hbar^2 k^2}{2m} \frac{1}{e^{\beta(E_k - \mu)} + 1}$$
$$= \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty \frac{E^{3/2}}{e^{\beta(E-\mu)} + 1} dE. \tag{7.20}$$

One is interested in these formulas in the limit in which $E_F/k_B T \gg 1$, which is valid for metals, as may be seen from Table 7.3. The two required integrals are of the form

$$I_j(\beta, \beta\mu) = \int_0^\infty \frac{E^{j+1/2}}{e^{\beta(E-\mu)} + 1} dE, \tag{7.21}$$

with $j = 0$ and $j = 1$, respectively. One makes a power series in T (see Appendix W7A) and finds that

$$I_j(\beta, \beta\mu) = \frac{1}{(j + \frac{3}{2})\beta^{j+3/2}} \left[(\beta\mu)^{j+3/2} + \frac{\pi^2}{6} \left(j + \frac{3}{2}\right) \left(j + \frac{1}{2}\right) (\beta\mu)^{j-1/2} + \cdots \right]. \tag{7.22}$$

The Fermi energy may be written as

$$E_F = (\tfrac{3}{2} I_0)^{3/2}. \tag{7.23}$$

Solving for the chemical potential in terms of E_F and T gives

$$\mu = E_F \left[1 - \frac{\pi^2}{12}\frac{1}{(\beta E_F)^2} + \cdots\right]. \tag{7.24}$$

To order $(k_B T/E_F)^2$ the chemical potential and the Fermi energy are the same. The internal energy per unit volume, $u = U/V$, becomes

$$u = \frac{1}{5\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} E_F^{5/2}\left[1 + \frac{5\pi^2}{12}\left(\frac{k_B T}{E_F}\right)^2 + \cdots\right]. \tag{7.25}$$

The specific heat (at constant volume) per unit volume is

$$c_v = \frac{\partial u}{\partial T} = \frac{1}{6}\left(\frac{2m}{\hbar^2}\right)^{3/2} E_F^{1/2} k_B^2 T \equiv \frac{\pi^2 k_B^2}{3}\rho(E_F)T \equiv \gamma T. \tag{7.26}$$

[The significance of the quantity $\rho(E_F) = 3n/2E_F$ is discussed in Section 7.7.] Thus for metals there is an electronic contribution to the specific heat which is linear in T for $k_B T \ll E_F$. This is in addition to the lattice contribution discussed in Sections 5.5 and 5.6. The net result is that at low temperatures ($T < 10$ K) in metals,

$$c_v \approx \gamma T + AT^3. \tag{7.27}$$

A simple way of understanding the linear behavior is to recognize that most of the electrons in the Fermi sea cannot be thermally excited, since the states above them are occupied. Only those electrons within an energy band of width $\approx k_B T$ near the Fermi energy are capable of being excited to vacant states (see Fig. 7.3). Their number per unit volume is approximately $k_B T(\partial n/\partial E)|_{E_F} = 3nk_B T/2E_F$. The typical excitation of electrons from these states will also be of order $k_B T$, so the electronic contribution to the internal energy has a term varying as $(k_B T)^2$, and hence one obtains a linear specific heat.

An example of the low-T specific heat of a metal is presented in Fig. 7.4 for Ag. The measured value of γ is approximately 10% larger than the free-electron prediction.

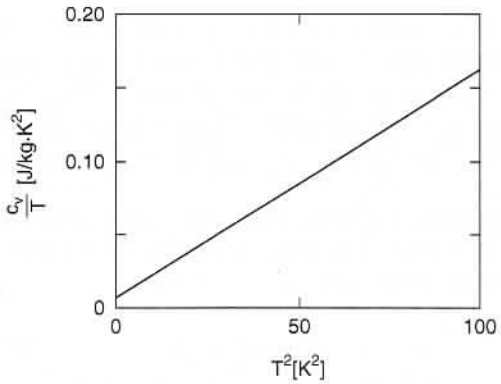


Figure 7.4. Ratio c_v/T plotted as a function of T^2 for Ag. (Data from CRC Handbook of Chemistry and Physics, 66th ed., R. C. Weast, ed., Boca Raton, Fla., 1985.)

This can be attributed to an enhancement of the free-electron mass due to interactions of the electron with the lattice and other electrons.

TRANSPORT THEORY

The description of conductivity in the Sommerfeld theory is facilitated by studying the motion of a collection of electrons in phase space, a six-dimensional space constructed from the three spatial and three momentum variables. The Boltzmann equation provides an alternative formulation for Newton's equations of motion used in Section 7.2. It also provides a framework in which other transport phenomena, such as thermal conductivity, may be studied. Details are presented in Section W7.1.

7.5 Onsager Relations

One may show that both an electric field $\mathbf{E}_0$ and a thermal gradient can produce an electric current density $\mathbf{J}$ and a heat current density $\mathbf{J}_Q$:

$$\mathbf{J} = \sigma\mathbf{E}_0 - \sigma S\nabla T, \tag{7.28}$$

$$\mathbf{J}_Q = \sigma ST\mathbf{E}_0 - \kappa'\nabla T. \tag{7.29}$$

These are the Onsager relations.

Integral expressions for the parameters are given in Section W7.1 and lead to

$$\sigma = \frac{ne^2\tau(E_F)}{m}, \tag{7.30}$$

$$S = -\frac{\pi^2}{3}\frac{k_B^2 T}{e}\frac{\partial}{\partial E}\ln[E^{3/2}\tau(E)]|_{E=\mu}, \tag{7.31}$$

where $\tau(E)$ is the collision time for electrons with energy E and

$$\kappa' = \frac{\pi^2}{3}\frac{k_B^2}{e^2}\sigma T. \tag{7.32}$$

The parameter σ is the same as encountered in the Drude theory for the electrical conductivity, except that $\tau(E)$ is evaluated at the Fermi energy. Indeed, when T is constant, Eq. (7.28) reduces to Ohm's law, $\mathbf{J} = \sigma\mathbf{E}_0$. The electron mean free path is defined by $\lambda = v_F\tau(E_F)$ indicating that carriers at the Fermi energy dominate the electrical conductivity. Typically, v_F is much larger than the drift velocity by many orders of magnitude.

The parameter S is called the *thermopower*. When a thermal gradient is established in a material and there is no flow of current, $\mathbf{J} = 0$, Eq. (7.28) predicts that a thermoelectric field is induced, given by

$$\mathbf{E}_0 = S\nabla T. \tag{7.33}$$

Equation (7.31) shows that the thermopower increases linearly in the absolute temperature. For metals typical values of S are on the order of 1 $\mu V/K$ at room temperature.

The thermopower can be an order of magnitude larger in transition metals, due to the contribution of d electrons to the conduction process. The sign and magnitude of S in a given metal are very sensitive to deviations from free-electron behavior.

The parameter κ' is related to the thermal conductivity, and Eq. (7.32) is known as the *Wiedemann–Franz law*. When no electrical current flows, Eqs. (7.28) and (7.29) may be combined to give

$$\mathbf{J_Q} = (\sigma S^2 T - \kappa')\nabla T \equiv -\kappa \nabla T \tag{7.34}$$

where κ is the thermal conductivity. To a good approximation, $\kappa' \approx \kappa$. Equation (7.34) is Fourier's heat conduction formula. Equation (7.32) shows that the thermal conductivity grows linearly with T at low T. It also demonstrates why good electrical conductors will also be good thermal conductors. The same electrons responsible for transporting electrical charge also tranport thermal energy. The ratio $L = \kappa/\sigma T = \pi^2 k_B^2/3e^2$ called the *Lorenz number*, depends only on fundamental constants. Its numerical value is 2.44×10^{-8} W · Ω/K^2. The experimental values of L at $T = 273$ K for (Li, Na, K, Rb, Al, Ag, Au) are $(2.22, 2.12, 2.23, 2.42, 2.14, 2.31, 2.32) \times 10^{-8}$ W · Ω/K^2, in reasonable agreement with theoretical prediction.

For metals the thermal conductivity grows linearly with T at low T, peaks at intermediate T, and falls off at high T. The falloff at high T is due largely to the shortened mean free path of the electrons due to electron–phonon scattering.

THE QUANTUM THEORY OF SOLIDS

In going beyond the Sommerfeld theory one attempts to describe all aspects of a solid in quantum-mechanical terms. The scattering from the lattice ions must be treated properly. The wave character of the electrons inevitably leads to interference effects, since typical wavelengths of the electrons are on the order of the Fermi wavelength, $\lambda_F = 2\pi/k_F$ and are comparable to interatomic spacings. In crystalline solids the ordered array of ions can produce Bragg diffraction effects, which block the propagation of electrons in certain directions and/or with certain ranges of energies. In the following sections the quantum theory of solids is developed.

7.6 Bloch's Theorem

The treatment of a many-particle system such as a solid requires that a number of simplifying approximations be made before a concise mathematical description is possible. One begins with the independent-electron approximation, where it is assumed that each electron interacts with a potential $V(\mathbf{r})$. This potential includes the effect of the ion cores as well as the other electrons in the solid. The result is that one need only solve a one-electron Schrödinger equation to find a set of eigenvalues and eigenfunctions.

In this section attention is focused on the solution of the Schrödinger equation for a periodic solid, where the potential has translation symmetry

$$V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r}), \tag{7.35}$$

where the set of vectors $\{\mathbf{R}\}$ defines the Bravais lattice. Fourier-analyzing $V(\mathbf{r})$ gives

$$V(\mathbf{r}) = \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{7.36}$$

where $\{\mathbf{G}\}$ are the reciprocal lattice vectors. The reality of $V(\mathbf{r})$ requires that the Fourier coefficients satisfy the "reality condition" $V_G^* = V_{-G}$, since

$$V^*(\mathbf{r}) = \sum_{\mathbf{G}} V_{\mathbf{G}}^* e^{-i\mathbf{G}\cdot\mathbf{r}} = \sum_{\mathbf{G}} V_{-\mathbf{G}}^* e^{i\mathbf{G}\cdot\mathbf{r}} = V(\mathbf{r}) = \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}. \tag{7.37}$$

The periodicity of $V(\mathbf{r})$ is obvious:

$$V(\mathbf{r}+\mathbf{R}) = \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\cdot(\mathbf{r}+\mathbf{R})} = \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}} = V(\mathbf{r}), \tag{7.38}$$

where $\exp(i\mathbf{G}\cdot\mathbf{R}) = 1$ has been used.

The Schrödinger equation is

$$[H - E]\psi(\mathbf{r}) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) - E\right]\psi(\mathbf{r}) = 0. \tag{7.39}$$

Note that H is unchanged if a translation through vector $\mathbf{R}$ is made. The wavefunction $\psi(\mathbf{r}+\mathbf{R})$ satisfies the same equation that $\psi(\mathbf{r})$ does, and so differs from it by at most a constant, that is,

$$\psi(\mathbf{r}+\mathbf{R}) = \tau_{\mathbf{R}}\psi(\mathbf{r}). \tag{7.40}$$

The quantity $\tau_{\mathbf{R}}$ must have magnitude 1. If it were greater than 1, repeated translations would make the wavefunction grow in magnitude in an exponential fashion, that is,

$$\psi(\mathbf{r}+N\mathbf{R}) = (\tau_{\mathbf{R}})^N\psi(\mathbf{r}), \tag{7.41}$$

and it would not be possible to normalize the wavefunction in the infinite solid limit. Similarly, if the modulus were less than 1, $\psi(\mathbf{r}-N\mathbf{R})$ would not be normalizable as $N \to \infty$. Thus $\tau_{\mathbf{R}} = \exp(i\theta_{\mathbf{R}})$ with $\theta_{\mathbf{R}}$ real. By compounding translations, one has

$$\tau_{\mathbf{R}_1}\tau_{\mathbf{R}_2} = \tau_{\mathbf{R}_1+\mathbf{R}_2}, \tag{7.42}$$

so

$$\theta_{\mathbf{R}_1} + \theta_{\mathbf{R}_2} = \theta_{\mathbf{R}_1+\mathbf{R}_2}, \tag{7.43}$$

which is satisfied by $\theta_{\mathbf{R}} = \mathbf{k}\cdot\mathbf{R}$ with $\mathbf{k}$ being a real vector. Thus

$$\psi(\mathbf{r}+\mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}}\psi(\mathbf{r}). \tag{7.44}$$

The function defined by

$$u(\mathbf{r}) = e^{-i\mathbf{k}\cdot\mathbf{r}}\psi(\mathbf{r}) \tag{7.45}$$

is seen to be a periodic function, since

$$u(\mathbf{r}+\mathbf{R}) = e^{-i\mathbf{k}\cdot(\mathbf{r}+\mathbf{R})}\psi(\mathbf{r}+\mathbf{R}) = e^{-i\mathbf{k}\cdot(\mathbf{r}+\mathbf{R})}e^{i\mathbf{k}\cdot\mathbf{R}}\psi(\mathbf{r}) = u(\mathbf{r}). \tag{7.46}$$

Enlarging the notation, Bloch's theorem states that the solution of the Schrödinger equation may be factored into a plane wave multiplied by a periodic function $u_k(\mathbf{r})$ with the Bravais lattice periodicity:

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}). \tag{7.47}$$

The acceptable values of $\mathbf{k}$ may be determined by imposing periodic boundary conditions (see Section 5.1). For a solid of size N_1 by N_2 by N_3 atoms, a translation through $N_1\mathbf{u}_1$ should leave the wavefunction unchanged. Thus

$$e^{i\mathbf{k}\cdot\mathbf{u}_1 N_1} = 1, \tag{7.48}$$

which has N_1 independent solutions of the form

$$\mathbf{k}_1 = \frac{j_1}{N_1}\mathbf{g}_1, \qquad j_1 = 0, 1, \ldots, N_1 - 1, \tag{7.49}$$

where $\mathbf{g}_1 \cdot \mathbf{u}_1 = 2\pi$. More generally,

$$\mathbf{k} = \frac{j_1}{N_1}\mathbf{g}_1 + \frac{j_2}{N_2}\mathbf{g}_2 + \frac{j_3}{N_3}\mathbf{g}_3, \qquad j_n = 0, 1, \ldots, N_n - 1, \tag{7.50}$$

where $\{\mathbf{g}_i\}$ are the primitive reciprocal lattice vectors [see Eq. (3.8).] Thus $\mathbf{k}$ is a point in the first Brillouin zone. The total number of such points is $N = N_1N_2N_3$, which is the number of lattice cells in the crystal.

Due to the periodicity of $u_{\mathbf{k}}(\mathbf{r})$ it may be expanded as a Fourier series:

$$u_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} u_{\mathbf{G}}(\mathbf{k})e^{i\mathbf{G}\cdot\mathbf{r}}. \tag{7.51}$$

Inserting this into the Schrödinger equation gives

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + \sum_{\mathbf{G}'} V_{\mathbf{G}'}e^{i\mathbf{G}'\cdot\mathbf{r}} - E\right]\sum_{\mathbf{G}} u_{\mathbf{G}}(\mathbf{k})e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}} = 0. \tag{7.52}$$

Using

$$\sum_{\mathbf{G},\mathbf{G}'} V_{\mathbf{G}'}u_{\mathbf{G}}e^{i(\mathbf{G}+\mathbf{G}')\cdot\mathbf{r}} = \sum_{\mathbf{G},\mathbf{G}'} V_{\mathbf{G}'}u_{\mathbf{G}-\mathbf{G}'}e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{7.53}$$

this may be simplified to

$$\left[\frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})^2 - E\right]u_{\mathbf{G}}(\mathbf{k}) + \sum_{\mathbf{G}'} V_{\mathbf{G}'}u_{\mathbf{G}-\mathbf{G}'}(\mathbf{k}) = 0, \tag{7.54}$$

where use has been made of the linear independence of the functions $\exp(i\mathbf{G}\cdot\mathbf{r})$ in removing the $\mathbf{G}$ sum. This infinite set of coupled linear algebraic equations for $\{u_{\mathbf{G}}(\mathbf{k})\}$ may have many solutions, so the notation is expanded to $\{u_{n,\mathbf{G}}(\mathbf{k})\}$, where $n = 1, 2, 3, \ldots$ is called the *band-index*. The energy eigenvalues will also be labeled by this index, $E_n(\mathbf{k})$. The relations in Eq. (7.54) are referred to as *Bloch's difference equations*.

The condition for nontrivial solutions to Eq. (7.54) to exist is the vanishing of the determinant (called the *Hill determinant*):

$$\left| \left[\frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})^2 - E \right] \delta_{\mathbf{G},\mathbf{G}''} + V_{\mathbf{G}-\mathbf{G}''} \right| = 0. \tag{7.55}$$

The roots of this equation determine the eigenvalue spectrum $\{E_n(\mathbf{k})\}$. This spectrum is invariant under the transformation $\mathbf{k} \to \mathbf{k} + \mathbf{K}$, where $\mathbf{K}$ is a reciprocal lattice vector. Making this substitution into the Hill determinant and letting $\mathbf{G}' = \mathbf{G} + \mathbf{K}$ and $\mathbf{G}'' = \mathbf{K}' - \mathbf{K}$ yields

$$\left| \left[\frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G}')^2 - E \right] \delta_{\mathbf{G}',\mathbf{K}'} + V_{\mathbf{G}'-\mathbf{K}'} \right| = 0. \tag{7.56}$$

The condition remains invariant under the transformation above. Hence the first Brillouin zone contains the entire energy spectrum. Other Brillouin zones simply contain replicas of this spectrum.

Truncating the Hill determinant by employing a finite set of $\mathbf{G}$ vectors and solving the resulting secular determinant for the roots $E = E_n(\mathbf{k})$ provides a method, in principle, for calculating band structures. Since the Fourier coefficients V_G fall off with increasing $\mathbf{G}$ for large $\mathbf{G}$, the energy eigenvalue spectrum will converge as the size of the determinant is increased. More practical methods exist, such as the Green's function [Korringa, Kohn, and Rostoker (KKR)] method, the augmented plane wave (APW) method, and the pseudopotential method, but these are beyond the scope of the present book. The books by Fletcher (1971) and Ashcroft and Mermin (1976) provide good introductions to modern band-structure computation techniques.

Figure 7.5 presents the results of an electronic band-structure calculation for diamond along various directions in the first Brillouin zone. The lowest band corresponds to $n = 1$, the next higher band to $n = 2$, and so on. One sees that there are two degenerate conduction bands and three degenerate valence bands at the Γ point. Diamond is an insulator. The Fermi level lies at *midgap*, so all states with $E < 0$ comprise the valence bands and are occupied. The conduction bands are empty. The bottom of the conduction band does not occur at the Γ point but rather, along the [100] direction. The bandgap is close to 6 eV, making diamond transparent to visible light.

7.7 Nearly Free Electron Approximation

In analyzing some metals the free-electron approximation forms a suitable starting point for explaining the band structure. In these materials the ion cores are largely screened by the valence electrons, leaving behind a weak potential that may be treated using perturbation theory. Examples of such materials are the alkali metals (Li, Na, K, Rb, Cs), Mg, and Al.

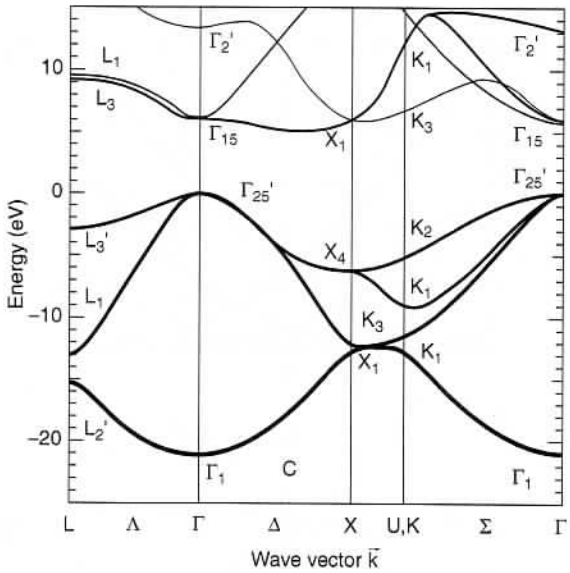


Figure 7.5. Electron band-structure calculation for diamond. [From J. R. Chelikowsky et al, *Phys. Rev. B*, **29**, 3470 (1984). Copyright 1984 by the American Physical Society.]

Begin with Bloch's difference equation [see Eq. (7.54)],

$$(\varepsilon_{\mathbf{k}+\mathbf{G}} - E)u_{\mathbf{G}}(\mathbf{k}) + \sum_{\mathbf{G}'} V_{\mathbf{G}'} u_{\mathbf{G}-\mathbf{G}'}(\mathbf{k}) = 0, \tag{7.57}$$

where a free-electron kinetic energy is defined by

$$\varepsilon_{\mathbf{k}+\mathbf{G}} = \frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})^2. \tag{7.58}$$

In the case where the $\{V_G\}$ all vanish, the solutions are

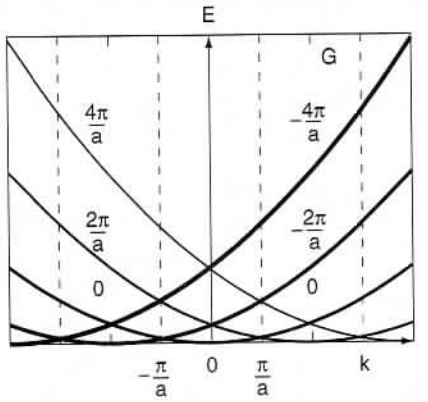
$$E = \varepsilon_{\mathbf{k}+\mathbf{G}}. \tag{7.59}$$

In the extended-zone scheme the energy dispersion curve is a single parabola of revolution. In the reduced-zone scheme the information gets compressed into the first Brillouin zone (e.g., $-\pi/a < k < \pi/a$ for the one-dimensional case). In the periodic-zone scheme the information is repeated periodically in each Brillouin zone. Five parabolas are illustrated in the respective schemes in Fig. 7.6.

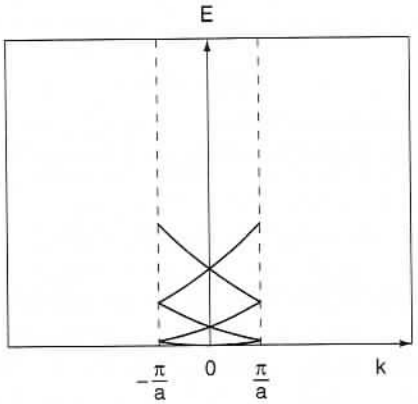
There is one case for which the free-electron approximation fails badly, and that is when energy bands intersect each other. At such degeneracies any perturbation will have a large effect. To see this in detail, suppose that the bands described by reciprocal lattice vectors $\mathbf{G}_1$ and $\mathbf{G}_2$ cross each other. An example of this is seen in the periodic-zone scheme in Fig. 7.6. The one-dimensional bands with $G = 0$ and $G = 2\pi/a$ intersect at $k = \pi/a$. Other couplings beside those between the two bands are neglected. Thus Eq. (7.57) becomes a pair of equations,

$$(\varepsilon_{\mathbf{k}+\mathbf{G}_1} + V_0 - E)u_{\mathbf{G}_1} + V_{\mathbf{G}_1-\mathbf{G}_2}u_{\mathbf{G}_2} = 0, \tag{7.60a}$$

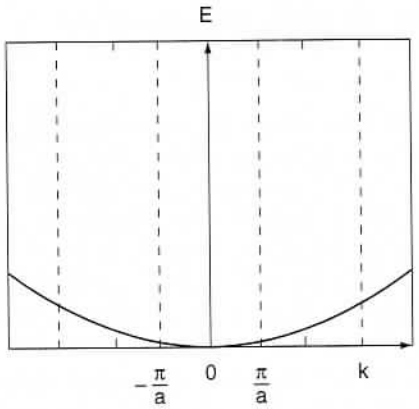
$$(\varepsilon_{\mathbf{k}+\mathbf{G}_2} + V_0 - E)u_{\mathbf{G}_2} + V_{\mathbf{G}_2-\mathbf{G}_1}u_{\mathbf{G}_1} = 0. \tag{7.60b}$$



(a)



(b)



(c)

Figure 7.6. Free-electron bands in the (*a*) periodic, (*b*) reduced, and (*c*) extended-zone schemes. The first Brillouin zone extends from $-\pi/a$ to π/a.

Here V_0 describes a uniform background potential. The condition for the existence of solutions for this pair of linear equations is the vanishing of the secular determinant

$$\begin{vmatrix} \varepsilon_{\mathbf{k}+\mathbf{G}_1} + V_0 - E & V_{\mathbf{G}_1-\mathbf{G}_2} \\ V^*_{\mathbf{G}_1-\mathbf{G}_2} & \varepsilon_{\mathbf{k}+\mathbf{G}_2} + V_0 - E \end{vmatrix} = 0, \tag{7.61}$$

yielding

$$E^2 - E(\varepsilon_{\mathbf{k}+\mathbf{G}_1} + \varepsilon_{\mathbf{k}+\mathbf{G}_2} + 2V_0) + (\varepsilon_{\mathbf{k}+\mathbf{G}_1} + V_0)(\varepsilon_{\mathbf{k}+\mathbf{G}_2} + V_0) - |V_{\mathbf{G}_1-\mathbf{G}_2}|^2 = 0, \tag{7.62}$$

where $V_{\mathbf{G}_2-\mathbf{G}_1} = V^*_{\mathbf{G}_1-\mathbf{G}_2}$ has been used. The two roots are

$$E_\pm = V_0 + \frac{\varepsilon_{\mathbf{k}+\mathbf{G}_1} + \varepsilon_{\mathbf{k}+\mathbf{G}_2}}{2} \pm \sqrt{\left(\frac{\varepsilon_{\mathbf{k}+\mathbf{G}_1} - \varepsilon_{\mathbf{k}+\mathbf{G}_2}}{2}\right)^2 + |V_{\mathbf{G}_1-\mathbf{G}_2}|^2}. \tag{7.63}$$

At the point of degeneracy the two $\varepsilon_{\mathbf{k}+\mathbf{G}}$ factors are equal and

$$E_\pm = V_0 + \varepsilon_{\mathbf{k}+\mathbf{G}_1} \pm |V_{\mathbf{G}_1-\mathbf{G}_2}|. \tag{7.64}$$

The condition for the equality of the free-electron energies reduces to

$$G_1^2 + 2\mathbf{k}\cdot\mathbf{G}_1 = G_2^2 + 2\mathbf{k}\cdot\mathbf{G}_2. \tag{7.65}$$

This is precisely the von Laue condition for Bragg scattering of an electron. An energy gap is opened up between the two branches of the dispersion curve equal to twice the strength of the Fourier coefficient of the potential at that reciprocal lattice vector:

$$E_g = \Delta E = E_+ - E_- = 2|V_{\mathbf{G}_1-\mathbf{G}_2}|. \tag{7.66}$$

Typical graphs for the two branches of the dispersion curves are presented in Fig. 7.7 in the reduced-zone scheme.

As in the case of phonons, the spectral properties of the electronic states of a solid may be represented by the density of states, $\rho(E)$, the number of states per unit volume

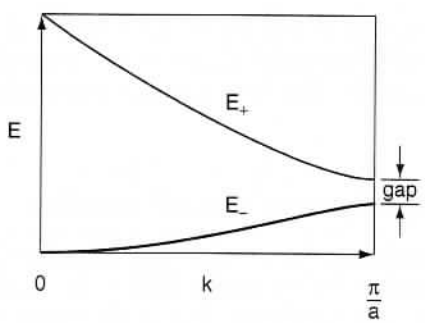


Figure 7.7. Typical dispersion curves for electrons in two interacting bands in the reduced-zone scheme.

ended-zone schemes.

per unit energy interval. It is defined by

$$\rho(E) = \frac{1}{V}\sum_{n,\mathbf{k},s} \delta(E - E_{n,\mathbf{k}}) = 2\sum_{n} \int \frac{d^d k}{(2\pi)^d} \delta(E - E_{n,\mathbf{k}}), \tag{7.67}$$

where d is the dimensionality of the solid and the sum extends over both electron spin projections, s, and the band index, n. For nonmagnetic materials the energy is independent of s. The $\mathbf{k}$ integration is limited to the first Brillouin zone. Note that $\rho(E)$ is a function of the form of the energy dispersion curves as well as the dimensionality of the solid. The quantity V is the crystal volume in three dimensions, area in two dimensions and length in one dimension.

For the one-dimensional free-electron gas one may use an extended-zone scheme and simply write

$$\rho(E) = 2\int_{-\infty}^{\infty} \frac{dk}{2\pi} \delta\left(E - \frac{\hbar^2 k^2}{2m}\right) = \frac{1}{\pi\hbar}\sqrt{\frac{2m}{E}}\Theta(E), \tag{7.68}$$

where $\Theta(E)$ is the unit step function. There is a threshold at $E = 0$ and an inverse-square-root behavior for $E > 0$. Note that for energies sufficiently far from the band edge, the limits of the $\mathbf{k}$ integration may be set to infinity. The corresponding result for a two-dimensional free-electron gas is

$$\rho(E) = 2\int_{-\infty}^{\infty} \frac{d^2 k}{(2\pi)^2} \delta\left(E - \frac{\hbar^2 k^2}{2m}\right) = \frac{m}{\pi\hbar^2}\Theta(E), \tag{7.69}$$

which has a threshold at $E = 0$ and is then constant for $E > 0$, For three dimensions the density of states is

$$\begin{aligned}\rho(E) &= 2\int \frac{d^3 k}{(2\pi)^3} \delta\left(E - \frac{\hbar^2 k^2}{2m}\right) \\ &= \frac{1}{\pi^2}\int_0^E \frac{dk}{dE} k^2 \delta\left(E - \frac{\hbar^2 k^2}{2m}\right) dE = \frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \sqrt{E}\Theta(E),\end{aligned} \tag{7.70}$$

and the density of states grows as $\sqrt{E}$ above the threshold at $E = 0$. Extending these results to the nearly free-electron model results in the introduction of van Hove singularities, which may occur at the zone boundaries or at any other extremal point of the energy band spectrum.

7.8 Tight-Binding Approximation in One Dimension

When an electron is bound in a deep-enough attractive potential well, the low-lying spectrum is a set of discrete energy levels. This is in contrast to the free-electron case, where the spectrum is a continuum. Just as the nearly free electron description of a solid in Section 7.7 was built around the free-electron model, it is possible to take as a starting point a solid treated as a collection of bound electrons. This is the

idea behind the tight-binding approximation. We begin this section by reviewing the Hückel approximation from molecular physics. The approximation is then extended to a periodic solid.

The interaction of two atoms to form a molecule is most simply described in terms of the Hückel approximation. Molecular orbitals are constructed as a linear combination of atomic orbitals. A model Hamiltonian is introduced in which the atomic properties and the interaction between any pairs of atoms are parametrized in terms of numbers. Thus, for a two-atom situation, with one relevant state on each atom, the electronic states are denoted by $|1\rangle$ and $|2\rangle$ respectively, and the molecular wavefunction is written as a linear combination of these states

$$|\psi\rangle = a_1|1\rangle + a_2|2\rangle. \tag{7.71}$$

The on-site matrix elements of the Hamiltonian are $\langle 1|H|1\rangle = \varepsilon_1$ and $\langle 2|H|2\rangle = \varepsilon_2$ These are the energies of the electronic states of the individual atoms. The off-diagonal matrix elements, or *tunneling* (or *hopping*) *matrix elements*, are $\langle 1|H|2\rangle = t$ and $\langle 2|H|1\rangle = t^*$. They come about because the Coulomb interaction has matrix elements connecting the states on different atoms. The Hamiltonian matrix is

$$H = \begin{bmatrix} \varepsilon_1 & t \\ t^* & \varepsilon_2 \end{bmatrix}. \tag{7.72}$$

Note that H defines a Hermitian matrix (i.e., one whose transpose and complex conjugate have identical matrix elements). It is convenient to neglect the direct overlap of states $|1\rangle$ and $|2\rangle$ and assume the orthonormality condition $\langle i|j\rangle = \delta_{ij}$. This implies the normalization condition $|a_1|^2 + |a_2|^2 = 1$. The eigenfunctions and eigenvalues are determined from the Schrödinger equation $(H - E)|\psi\rangle = 0$, which leads to the following pair of equations:

$$(\varepsilon_1 - E)a_1 + t^* a_2 = 0, \tag{7.73a}$$

$$t a_1 + (\varepsilon_2 - E)a_2 = 0. \tag{7.73b}$$

The secular determinant vanishes, so

$$E^2 - E(\varepsilon_1 + \varepsilon_2) + \varepsilon_1\varepsilon_2 - |t|^2 = 0, \tag{7.74}$$

with the solutions

$$E_\pm = \frac{\varepsilon_1 + \varepsilon_2}{2} \pm \sqrt{\left(\frac{\varepsilon_1 - \varepsilon_2}{2}\right)^2 + |t|^2}. \tag{7.75}$$

Here E_- is the energy of the bonding state and E_+ is that of the antibonding state of the molecule.

This result is readily generalized to a chain of N atoms. For finite N it may represent a linear polymer. In the limit of large N it may be thought of as a one-dimensional solid. Thus

$$|\psi\rangle = \sum_{j=1}^{N} a_j|j\rangle. \tag{7.76}$$

In the simplest case, one makes the assumption of NN interactions only. The nonvanishing matrix elements of H are then

$$\langle j|H|j\rangle = \varepsilon_j, \qquad \langle j+1|H|j\rangle = t_j, \qquad \langle j|H|j+1\rangle = t_j^*. \tag{7.77}$$

Periodic boundary conditions, introduced in Section 5.1, are imposed, so there is the subscript identification $N+1 \to 1$ and $0 \to N$. The Schrödinger equation leads to a set of N coupled linear algebraic equations:

$$(\varepsilon_j - E)a_j + t_{j-1}a_{j-1} + t_j^* a_{j+1} = 0. \tag{7.78}$$

In the case where all the atoms are identical, $\varepsilon_j = \varepsilon$ and $t_j = t$, for all j. Let $a_j = Ac^j$, so that

$$\varepsilon - E + \frac{t}{c} + t^* c = 0. \tag{7.79}$$

As in Section 7.6, it is found that the quantity c must be of modulus 1, so $c = \exp(i\theta)$. It is convenient to assume t to be real. Then

$$\varepsilon - E + 2t\cos\theta = 0. \tag{7.80}$$

Imposing the periodic boundary conditions $a_{N+1} = a_1$ implies that $\exp(iN\theta) = 1$, so that $\theta = 2\pi n/N$, where $n = 0, 1, 2, \ldots, N-1$. Letting $k_n = 2\pi n/Na$, where a is the lattice constant, and suppressing the subscript n leads to the dispersion curve:

$$E(k) = \varepsilon + 2t\cos ka. \tag{7.81}$$

Thus one finds a single allowed energy band of width $4|t|$. Depending on the sign of t, it could have a maximum or a minimum at $k = 0$. The allowed values of k extend over the first Brillouin zone, from $k = -\pi/\mathrm{a}$ to π/a. For core levels it is safe to assume that $t = 0$, so there is no overlap or interaction.

For $t < 0$ there is a minimum at $k = 0$, and for small k it follows that

$$E \approx \varepsilon - 2|t| + |t|a^2k^2 = E_0 + \frac{\hbar^2 k^2}{2m^*}, \tag{7.82}$$

where $E_0 = \varepsilon - 2|t|$ and m^* denotes the effective mass of an electron,

$$m^* = \frac{\hbar^2}{2|t|a^2}. \tag{7.83}$$

The inverse of m^* is proportional to the curvature of the energy dispersion curve at $k = 0$. One may define in a similar way the effective mass for the case in which there is a maximum at $k = 0$. In that case m^* will be negative. Equation (7.82) predicts that as $m^* \to \infty$, the energy E will be independent of k. The tight-binding bands are sketched in Fig. 7.8 for the cases $t > 0$ and $t < 0$.

The tight-binding approximation finds application in the description of semiconductors and the d bands of transition metals. It is also of use in finding the band

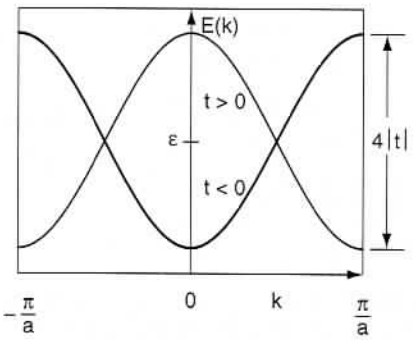


Figure 7.8. Tight-binding bands for the cases $t > 0$ and $t < 0$ in the reduced-zone scheme.

structure of polymers. It may be generalized to higher dimensions, to more than just NN interactions, and to the case where there is more than one state per site.

The density of states for the one-dimensional tight-binding model is obtained from Eq. (7.67) by integrating over the first Brillouin zone:

$$\rho(E) = 2\int_{-\pi/a}^{\pi/a} \frac{dk}{2\pi}\delta(E - \varepsilon - 2t\cos ka) = \frac{2\Theta[4t^2 - (E-\varepsilon)^2]}{\pi a\sqrt{4t^2 - (E-\varepsilon)^2}}. \tag{7.84}$$

This gives a diverging density of states at the upper and lower band edges, $E = \varepsilon \pm 2t$ at $k = 0$ and at $k = \pm\pi/a$. The density of states for a random one-dimensional solid is discussed briefly in Section W7.2.

7.9 Tight-Binding Approximation in Two Dimensions

The tight-binding approximation will now be applied to calculation of the band structure of a two-dimensional crystal: a CuO_2 plane. As discussed in Chapter 11, these planes of atoms play a central role in the structure of ceramic high-temperature superconductors. Assume that there is a square lattice with three atoms per unit cell: a copper atom, an oxygen to the right of it (in the x direction), and an oxygen above it (in the y direction). The dsp^2 orbitals from the copper atoms interact with the p_x and p_y orbitals of the oxygen atoms to form bonds (Fig. 7.9). These bonds have mixed ionic and covalent character.

Let the amplitude of the wavefunction on the copper atom in cell (m, n) be denoted by $A_{m,n}$ and the corresponding amplitudes of the two oxygen atoms of the cell by $R_{m,n}$ and $U_{m,n}$, respectively. The tight-binding equations are

$$(E_{Cu} - E)A_{m,n} + t(R_{m,n} + U_{m,n} + R_{m-1,n} + U_{m,n-1}) = 0, \tag{7.85a}$$

$$(E_O - E)R_{m,n} + t(A_{m+1,n} + A_{m,n}) = 0, \tag{7.85b}$$

$$(E_O - E)U_{m,n} + t(A_{m,n} + A_{m,n+1}) = 0, \tag{7.85c}$$

where E_{Cu} and E_O are the on-site electron energies and t is the tunneling matrix element between the Cu and O atoms. Inserting the expressions

$$A_{m,n} = Ae^{i(mk_x + nk_y)a}, \qquad R_{m,n} = Re^{i(mk_x + nk_y)a}, \qquad U_{m,n} = Ue^{i(mk_x + nk_y)a}, \tag{7.86}$$

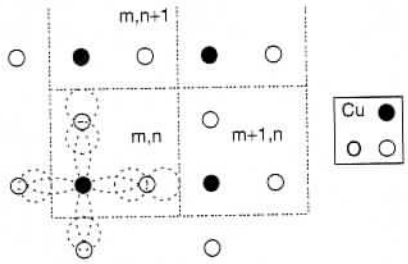


Figure 7.9. Copper atoms (solid circles) bonded to oxygen atoms (open circles) in the CuO_2 plane. The square unit cell is shown. The hybrid dsp^2-Cu orbitals combine with the p_x- and p_y-O orbitals to form σ-molecular orbitals.

where a is the lattice constant, leads to

$$(E_{Cu} - E)A + tR(1 + e^{-ik_x a}) + tU(1 + e^{-ik_y a}) = 0, \tag{7.87}$$

with

$$R = \frac{tA}{E - E_O}(1 + e^{ik_x a}), \tag{7.88}$$

$$U = \frac{tA}{E - E_O}(1 + e^{ik_y a}). \tag{7.89}$$

The resulting secular equation is

$$E^2 - (E_O + E_{Cu})E + E_O E_{Cu} - 4t^2\left(\cos^2\frac{k_x a}{2} + \cos^2\frac{k_y a}{2}\right) = 0. \tag{7.90}$$

The two solutions are

$$E_\pm(k_x, k_y) = \frac{E_{Cu} + E_O}{2} \pm \sqrt{\left(\frac{E_{Cu} - E_O}{2}\right)^2 + 4t^2\left(\cos^2\frac{k_x a}{2} + \cos^2\frac{k_y a}{2}\right)}. \tag{7.91}$$

The first Brillouin zone is a square occupying the space $|k_x| \leq \pi/a$, $|k_y| \leq \pi/a$. Along the line $k_x + k_y = \pi/a$, the energy is constant and has the value

$$E_\pm = \frac{E_{Cu} + E_O}{2} \pm \sqrt{\left(\frac{E_{Cu} - E_O}{2}\right)^2 + 4t^2}, \tag{7.92}$$

which is independent of **k**. The shape of the Fermi boundary is determined by the electron density in the plane, a factor that is often controlled by the nature of other atoms above or below the planes. For example, a half-filled energy band has all the states within the square defined by $|k_x + k_y| < \pi/a$ and $|k_x - k_y| < \pi/a$ occupied at $T = 0$ K. Several lines of constant energy are sketched in Fig. 7.10. Note that the free-electron model works well when k_x and k_y are near the zone center and the constant-energy contour is a circle.

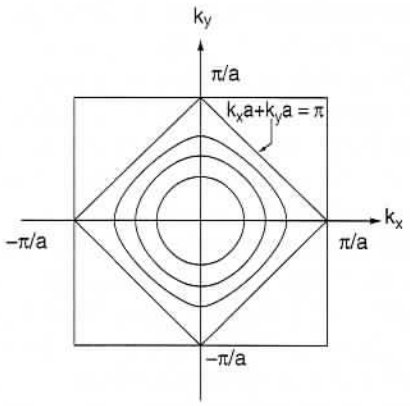


Figure 7.10. Lines of constant energy for the lower energy band in Eq. (7.91) for a typical case.

It is of some value to relate the bandwidth in the tight-binding approximation to the number of NNs in a lattice. The Schrödinger equation is of the form

$$(E_0 - E)A_0 + t\sum_{n=1}^{Z} A_n = 0, \tag{7.93}$$

where A_0 denotes the amplitude on a given site and A_n are the amplitudes on the Z neighboring sites. The latter amplitudes are related to A_0 through simple phase factors. At some points in the Brillouin zone all these phases are equal to 1 and the sum is maximized so that $E = E_0 + Zt$. On the other hand, at other points the sum equals -1 and is therefore minimized. Then $E = E_0 - Zt$. The bandwidth is therefore given by

$$B = 2Zt. \tag{7.94}$$

For fixed t, the larger the number of NNs, the larger the bandwidth.

7.10 Metals, Insulators, Semiconductors, and Semimetals

The energy-band picture provides a simple understanding for the wide variation in conductivity in going from material to material, at least for the case of crystalline solids. In a *metal* the Fermi level lies within an allowed energy band. There are unoccupied states with energies immediately above the Fermi energy and there are occupied states immediately below it. The application of even a weak electric field can elevate an electron in energy from an occupied state to a vacant state, where it can participate in the flow of a net current. In metals the concentration of charge carriers is essentially constant, independent of T, and determined by the atomic concentration and the valence of the ion. Referring to the Drude formula $\sigma = ne^2\tau/m$, the temperature variation of the conductivity may be understood. As the temperature of the material increases, σ decreases, due to the shortening of the collision time brought about by the emission, absorption, and scattering of phonons.

In an ideal *insulator* the Fermi level lies within an energy gap. All energy bands are either completely filled with electrons or completely vacant. (This is never true in real materials, due to defects, impurities, etc.). Since the number of states (with a given spin projection) in an energy band equals the number of unit cells in the crystal, a necessary condition for an insulator is that there be an even number of electrons per unit cell. This, however, is not a sufficient condition. For example, divalent Mg is a metal due to band overlap in different directions in **k** space. The occupied bands are called *valence bands*, and the unoccupied bands are called *conduction bands*. When a weak electric field is applied, there is no electric current in the filled bands since these electrons cannot be excited. The field is unable to provide the energy needed for making an interband transition to a conduction band. At finite temperatures, thermal excitation of electrons from the valence to the conduction bands is possible. The conductivity will be determined largely by the number of carriers produced. For wide-bandgap materials, this number will be very small. Free charge trapped in an insulator can remain there for very long periods of time without being conducted away, because of the very high electrical resistance. This forms the basis of solid-state CMOS memory devices, where a bit of information corresponds to a stored electric charge.

An intrinsic semiconductor is an insulator with a relatively small bandgap. Examples of elemental semiconductors include Ge and Si. For such materials thermal excitation of carriers is nonnegligible at room temperature. As the temperature rises the number of thermally generated carriers grows exponentially and the conductivity also increases exponentially. This effect more than offsets the shortening of the collision lifetime due to phonon interactions.

The conductivities of insulators and semiconductors are highly sensitive to the presence of impurities. Impure semiconductors are called *extrinsic semiconductors*: if the impurities are introduced in a controlled fashion, they are termed *doped semiconductors*. Semiconductors are discussed in detail in Section 7.12 and Chapter 11. The question of conduction in impure materials is considered later in this chapter.

Semimetals have a slight overlap between the valence and conduction bands. As a result, the valence band is nearly filled and the conduction band is nearly empty. The electrons in the conduction band can carry an electric current, as can the vacant electron states (holes) in the valence band. Examples of elemental semimetals are graphite, As, Bi, and Sb.

In Fig. 7.11 hypothetical band structures are sketched for a metal, an insulator (or semiconductor), and a semimetal.

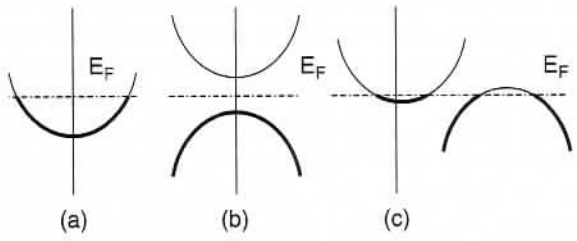


Figure 7.11. Hypothetical band structures for (*a*) a metal, (*b*) an insulator, and (*c*) a semimetal. The dashed line denotes the Fermi level. The heavy lines represent occupied states.

QUANTUM EFFECTS IN ELECTRICAL CONDUCTION

In the following section the temperature dependence of resistivity is analyzed and we will see that quantum-mechanical effects need to be introduced to describe the low-T resistivity. Then the Hall effect is reexamined from the vantage point of electron band theory in Section W7.4.

7.11 Temperature Dependence of Resistivity in Metals

Resistivity is defined as the inverse of conductivity: $\rho = 1/\sigma$. It has been seen that conductivity is determined by the average collision time, τ. A collision is any process that destroys the forward momentum of electrons. It may involve scattering from impurities or defects (time τ_i), or it may result from the emission or absorption of phonons or the scattering from phonons (time τ_{ph}). Since one may write the total scattering rate as the sum of the individual scattering rates, the resistivities will be additive:

$$\rho = \frac{m}{ne^2}\left(\frac{1}{\tau_i} + \frac{1}{\tau_{\text{ph}}}\right). \tag{7.95}$$

This separation, called *Matthiessen's rule*, is an approximation. The value of τ_i is controlled by the quality of the material and is essentially independent of the temperature, at least at low T. For a perfect material with no impurities (chemical or even isotopic) τ_i can be made arbitrarily long. The scattering time τ_{ph} depends on the temperature. In Fig. 7.12 data are presented on the variation of resistivity with temperature for aluminum.

Temperature dependence displays approximately linear behavior for $T > \Theta_D$. This may be understood simply in terms of the thermal occupancy of phonon modes in this temperature range. The Bose–Einstein distribution function for $\beta\hbar\omega \ll 1$ becomes

$$n(\omega, T) = \frac{1}{e^{\beta\hbar\omega} - 1} \approx \frac{k_B T}{\hbar\omega}, \tag{7.96}$$

so the number of phonons available (to be absorbed, to stimulate an emission process, or to scatter from) is linearly proportional to the temperature. These processes are

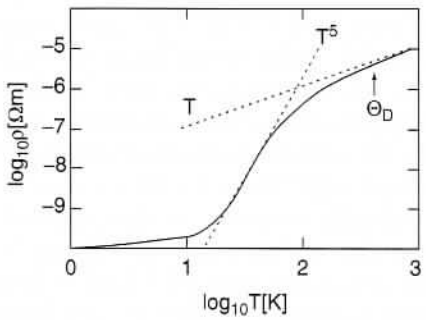


Figure 7.12. Plot of resistivity (in Ω·m) of Al versus temperature (in kelvin) on a log-log plot. The Debye temperature is denoted by $\Theta_D = 428$ K. Note that Al melts at $T = 934$ K. (Data from D. R. Lide, ed., *CRC Handbook of Chemistry and Physics*, 73rd ed., CRC Press, Boca Raton, Fla., 1991.)

likely to involve phonons with wave vectors anywhere in the Brillouin zone. The emission/absorption/scattering rate is proportional to the phonon mode occupancy, hence the linear behavior at high T.

At low temperatures, however, note that there is a range in which $\rho \propto T^5$. This comes about because of three effects. First, at low temperatures only low-energy phonons (with energies of approximately $k_B T$) can be absorbed. The phonon density of states grows quadratically with energy, so this introduces a factor proportional to T^2. Second, emission or absorption of these "soft" phonons are likely to result in only small deflections, and hence are not very effective in randomizing the electron's momentum direction. If θ is the scattering angle, the forward momentum will be reduced by the factor $1 - \cos\theta \sim \theta^2/2$. But $\theta \propto k/p$, where k and $p/\hbar$ are the wave vectors of the phonon and electron, respectively. This introduces an additional k^2 factor. Since ω and k are proportional for acoustic phonons and for a typical phonon $\omega \propto T$, one obtains an additional T^2 factor. The final factor of T arises from the behavior of the electron–phonon coupling constant. For example, in the deformation-potential interaction the phonon produces a transitory dilation or compression of the lattice. This, in turn, causes the energy bands to be shifted up or down locally and results in the conduction electron's wavefunction acquiring a phase shift. The result is that the electrons are scattered. The degree of dilation is measured by $\nabla \cdot \mathbf{u}$, $\mathbf{u}$ being the local displacement vector. Upon Fourier analysis this leads to a factor $\mathbf{k} \cdot \mathbf{u}$, and hence an additional T factor. The net result is the Bloch T^5 law for the resistivity. Other electron–phonon mechanisms produce a similar factor.

At very low temperatures (but above $T = 1.18$ K) one sees in Fig. 7.12 the residual effects of impurity scattering, and the resistivity tends to a constant value. Below $T = 1.18$ K the resistivity falls abruptly to zero. This is due to the onset of superconductivity, which is described in Chapter 16.

7.12 Semiconductors

Crystalline solids are characterized by the existence of allowed and forbidden electron energy bands. The distinction between insulators and semiconductors resides in the size of the bandgap and where the Fermi level sits relative to the band edges. In an insulator the Fermi level lies within a wide forbidden band, considerably removed from the filled valence band and the vacant conduction band. The gap is sufficiently wide that a typical applied electric field is unable to excite interband transitions, even if assisted by thermal fluctuations. (Very strong fields can lead to breakdown, however, due to interband tunneling.) In semiconductors there is generally a narrower gap. Two cases are distinguished: intrinsic and extrinsic. In the *intrinsic* case, the semiconductor is of high purity and has few defects. There are some thermally excited electrons that reside in the conduction band, and these leave behind holes in the valence band. These electrons and holes are each able to conduct an electrical current. In the *extrinsic* case, dopant atoms are added to the semiconductor, which may become thermally ionized and contribute either electrons to the conduction band or holes to the valence band.

The dynamics of carriers (electrons or holes) is often described in the semiclassical approximation by equations resembling the Hamilton equations of classical mechanics

$$\hbar \frac{d\mathbf{k}}{dt} = -e(\mathbf{E} + \mathbf{v}_n \times \mathbf{B}), \tag{7.97}$$

$$\mathbf{v}_n = \frac{1}{\hbar}\frac{\partial}{\partial \mathbf{k}} E_n(\mathbf{k}). \tag{7.98}$$

Each electron is assumed to reside in a particular band (fixed n), and interband transitions are neglected. This assumption puts an upper limit on the strength of the applied fields and the frequency at which they can change. Unlike the Hamilton equations, $\hbar\mathbf{k}$ is the crystal momentum (not the true momentum). Recalling that the energy bands are periodic functions in $\mathbf{k}$ space, $\mathbf{k}$ is effectively confined to the first Brillouin zone. Its value evolves under the influence of the Lorentz force.

Consider the case of an electron in a conduction band. Let E_c denote the minimum energy in the conduction band and assume that it occurs at the points $\mathbf{k} = \mathbf{k}^\nu$, $\nu = 1, \ldots, N_\nu$. If $N_\nu = 1$, there is a unique minimum, whereas if $N_\nu > 1$, there is a degeneracy and the corresponding valleys are probably linked by a crystal-symmetry operation. In the neighborhood of the minimum of the energy band at $\mathbf{k} = \mathbf{k}^\nu$ (where $\nabla_{\mathbf{k}} E_n(\mathbf{k}) = 0$), one may expand the energy surface

$$\begin{aligned} E_n(\mathbf{k}) &= E_c + \frac{1}{2}\sum_{ij}(k_i - k_i^\nu)(k_j - k_j^\nu)\frac{\partial^2 E_n}{\partial k_i \partial k_j} + \cdots \\ &\equiv E_c + \frac{\hbar^2}{2}(\mathbf{k} - \mathbf{k}^\nu)\cdot\frac{1}{\overleftrightarrow{m}_e^*}\cdot(\mathbf{k} - \mathbf{k}^\nu) + \cdots, \end{aligned} \tag{7.99}$$

where

$$\left(\frac{1}{\overleftrightarrow{m}_e^*}\right)_{ij} = \frac{1}{\hbar^2}\frac{\partial^2 E_n}{\partial k_i \partial k_j} \tag{7.100}$$

is called the *inverse* of the effective-mass tensor for the electrons.

In general, $\overleftrightarrow{m}_e^*$ is a symmetric matrix. It is positive definite since there is an absolute minimum in the band. The tensor may be diagonalized by an orthogonal matrix $\mathbf{R}$. This means that there exists a matrix $\mathbf{R}$ such that $\mathbf{R}\overleftrightarrow{m}_e^*\mathbf{R}^{-1}$ is a diagonal matrix. An orthogonal matrix is one whose transpose and inverse are the same. In the diagonal representation the mass eigenvalues are m_1^*, m_2^*, and m_3^*. In the case of twofold degeneracy, one sometimes writes $m_1^* = m_2^* = m_\perp^*$ (or m_t^*, the transverse mass) and $m_1^* = m_\parallel^*$ (or m_l^*, the longitudinal mass). The semiclassical equations for electrons may be written as

$$\overleftrightarrow{m}_e^* \cdot \frac{d\mathbf{v}_n}{dt} = -e(\mathbf{E} + \mathbf{v}_n \times \mathbf{B}). \tag{7.101}$$

In the case of the valence band, the maximum energy is denoted by E_v and the expansion of the energy surface around E_v leads to the introduction of a mass tensor defined with a negative sign:

$$\left(\frac{1}{\overleftrightarrow{m}_h^*}\right)_{ij} = -\frac{1}{\hbar^2}\frac{\partial^2 E_n}{\partial k_i \partial k_j}. \tag{7.102}$$

The equation for the holes in the valence band is

$$\overleftrightarrow{m}_h^* \cdot \frac{d\mathbf{v}_h}{dt} = +e(\mathbf{E} + \mathbf{v}_h \times \mathbf{B}). \tag{7.103}$$

Because of the downward curvature of the valence band, they behave effectively as if they had a positive charge.

The dynamical equations (7.101) and (7.103) for the velocity describe the motion of individual carriers in a band. They may be averaged over the band, however, and may then equally describe the motion of the average velocity of the carrier. One may introduce the effect of collisions phenomenologically, by means of a collision time matrix. The reason for introducing this quantity as a matrix is to take account of the possible anisotropy in the collision rates. Thus

$$\overleftrightarrow{m}_e^* \cdot \left(\frac{d\mathbf{v}_e}{dt} + \frac{1}{\overleftrightarrow{\tau}_e} \cdot \mathbf{v}_e \right) = -e(\mathbf{E} + \mathbf{v}_e \times \mathbf{B}), \tag{7.104a}$$

$$\overleftrightarrow{m}_h^* \cdot \left(\frac{d\mathbf{v}_h}{dt} + \frac{1}{\overleftrightarrow{\tau}_h} \cdot \mathbf{v}_h \right) = +e(\mathbf{E} + \mathbf{v}_h \times \mathbf{B}). \tag{7.104b}$$

In the absence of a magnetic field and for steady-state conditions, the electrical current may be written as

$$\mathbf{J} = \overleftrightarrow{\sigma} \cdot \mathbf{E}, \tag{7.105}$$

where the conductivity tensor is described in terms of contributions of electrons and holes by

$$\overleftrightarrow{\sigma} = e^2 \left(n_h \overleftrightarrow{\tau}_h \cdot \frac{1}{\overleftrightarrow{m}_h^*} + n_e \overleftrightarrow{\tau}_e \cdot \frac{1}{\overleftrightarrow{m}_e^*} \right). \tag{7.106}$$

Here n_h and n_e are the hole and electron concentrations, respectively.

7.13 Magnetoresistance

The Drude theory may be used to obtain a simple understanding of magnetoresistance (i.e., the effect of a magnetic field on the conductivity). Begin with the metallic case, where there is only one band, and then extend the analysis to semiconductors, where two or more bands may be involved. The equation of motion includes the magnetic force in addition to the electric force:

$$\frac{d\mathbf{v}}{dt} + \frac{\mathbf{v}}{\tau} = -\frac{e}{m^*}(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{7.107}$$

Take the magnetic field to be along the z direction and introduce the cyclotron frequency $\omega_c = eB/m^*$. In the absence of an electric field the electrons would move in cyclotron orbits, which are either planar circles or helices whose axes are along the direction of the magnetic field (z direction). The period around the circular orbit (or one turn of a helical orbit) is given by $2\pi/\omega_c$. In the presence of an electric field one may explore the solutions of Eq. (7.107) for the case in which the damping term dominates over the inertial term. This leads to the expressions

$$v_x + \omega_c \tau v_y = -\frac{e\tau}{m^*} E_x, \tag{7.108a}$$

$$v_y - \omega_c \tau v_x = -\frac{e\tau}{m^*} E_y, \tag{7.108b}$$

$$v_z = -\frac{e\tau}{m^*}E_z. \tag{7.108c}$$

Use $\mathbf{J} = ne\mathbf{v}$ to obtain

$$J_x = \sigma\frac{E_x - \omega_c\tau E_y}{1 + (\omega_c\tau)^2} = \sigma_{xx}E_x + \sigma_{xy}E_y, \tag{7.109a}$$

$$J_y = \sigma\frac{E_y + \omega_c\tau E_x}{1 + (\omega_c\tau)^2} = \sigma_{yx}E_x + \sigma_{yy}E_y, \tag{7.109b}$$

$$J_z = \sigma E_z = \sigma_{zz}E_z, \tag{7.109c}$$

where σ is the Drude conductivity. If $J_y = 0$, then $E_y = -\omega_c\tau E_x$ and $J_x = \sigma E_x$. There is no transverse magnetoresistance in this one-band model. As noted in Section 7.3, the Hall coefficient is given as $R_H = E_y/BJ_x = -1/ne$.

Next consider a semiconductor containing two bands, one with electrons and the other with holes. Let n_e and n_h denote the number of electrons and holes per unit volume, respectively. Their parameters are $(\tau_e, m_e^*, \sigma_e)$ and $(\tau_h, m_h^*, \sigma_h)$. The current density components are now

$$J_x = \sigma_e\frac{E_x - \omega_e\tau_e E_y}{1 + (\omega_e\tau_e)^2} + \sigma_h\frac{E_x + \omega_h\tau_h E_y}{1 + (\omega_h\tau_h)^2}, \tag{7.110a}$$

$$J_y = \sigma_e\frac{E_y + \omega_e\tau_e E_x}{1 + (\omega_e\tau_e)^2} + \sigma_h\frac{E_y - \omega_h\tau_h E_x}{1 + (\omega_h\tau_h)^2}, \tag{7.110b}$$

$$J_z = (\sigma_e + \sigma_h)E_z. \tag{7.110c}$$

For the case where $J_y = 0$ one now finds

$$E_y = \frac{\sigma_h Q_h/\Delta_h - \sigma_e Q_e/\Delta_e}{\sigma_h/\Delta_h + \sigma_e/\Delta_e}E_x, \tag{7.111}$$

where $Q = \omega_c\tau = \mu B$, where μ is the mobility and $\Delta = 1 + (\omega_c\tau)^2$. This leads to an expression for the magnetoconductivity:

$$J_x\left[\left(\frac{\sigma_e}{\Delta_e} + \frac{\sigma_h}{\Delta_h}\right) + \frac{(\sigma_e Q_e/\Delta_e - \sigma_h Q_h/\Delta_h)^2}{\sigma_e/\Delta_e + \sigma_h/\Delta_h}\right]E_x = \sigma_{xx}E_x. \tag{7.112}$$

In the limit of very large magnetic fields (i.e., $\mu_e B \gg 1$ and $\mu_h B \gg 1$), the first term in Eq. (7.112) becomes small and the B dependence of the second term drops out. Thus the magnetoconductivity is

$$\sigma_{xx} \approx \frac{(n_e - n_h)^2 e^2}{n_e m_e^*/\tau_e + n_h m_h^*/\tau_h}, \tag{7.113}$$

This, again, is independent of B (i.e., saturates). The high-field Hall coefficient approaches

$$R_H = -\frac{1}{(n_e - n_h)e}, \tag{7.114}$$

which displays the competition between electrons and holes in determining its sign.

In many metals, for example, the noble metals, things are more complicated. The starting assumption that the electrons are always characterized by a positive mass is not valid. The Fermi surface is such that it may contain necks connecting different Brillouin zones. The effective mass tensor then varies in magnitude and sign with location on the Fermi surface. A mass that varies in sign leads to a curvature of the orbit that also varies in sign. This leads to the possibility of open orbits instead of closed orbits perpendicular to the magnetic field. In such a situation the magnetoresistance can continue to grow with increasing magnetic field instead of saturating as in Eq. (7.113). The implications of open orbits are not considered further here.

CONDUCTION IN INSULATORS

A perfect insulator would, of course, block the flow of all charge. In practice, however, insulators do have some residual conductivity. There is always the possibility of activated conduction by transferring electrons to the conduction band or creating holes in the valence band. In the intrinsic case, this would require a thermal fluctuation to create an electron–hole pair, a process that has probability $\exp(-\beta E_g)$. For a wide-gap insulator, this would be negligible. For an insulator with impurity levels in the gap, the excitation energy need not be as large, but it may still be improbable. In this section it will be seen that one need not make transitions all the way to the conduction band for a carrier to move. All that is necessary is for a carrier from an occupied impurity level in the gap to find a vacant impurity level in the gap so that transitions can occur. The study of variable-range hopping explores this possibility. In addition, it will be shown that there is a possibility of nonohmic conduction in insulators, particularly at stronger electric fields, due to the Poole–Frenkel effect.

7.14 Variable-Range Hopping

Impurity levels in insulators or semiconductors may lie in the energy gap either below or above the Fermi level. At high temperatures these levels are likely to donate electrons to the conduction band (or accept electrons from the valence band) and the material will function as a doped semiconductor. At low temperatures, conduction occurs when an electron in an occupied impurity state is thermally excited and hops to a vacant state some distance R away (Fig. 7.13). Let the energy for this excitation be denoted by ΔE. The probability for being excited is $\exp(-\beta \Delta E)$. The probability that the electron will hop a distance R is also proportional to the overlap of the wavefunctions

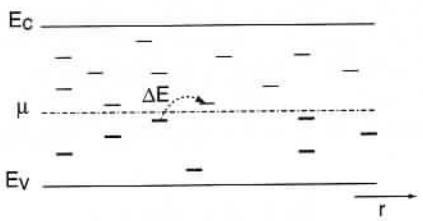


Figure 7.13. Set of impurity levels lying within the bandgap. Conduction occurs by hopping processes.

at the initial and final sites. This falls off exponentially with separation as $\exp(-2\alpha R)$. The parameter α is related to the binding energy E_B of the electron to the impurity ion (relative to the bottom of the conduction band) by $\alpha \approx \sqrt{2mE_B/\hbar^2}$. Thus the net hopping probability contains the factor $\exp(-\beta\Delta E - 2\alpha R)$.

An estimate for ΔE can be made by assuming an impurity concentration n_I, letting B denote the spread of energy values of impurity levels, and assuming a uniform spread of levels

$$\Delta E = \frac{B}{(4\pi R^3/3)n_I}. \tag{7.115}$$

Here the denominator represents the total number of impurities within a radius R of a given impurity. (More generally, one may replace n_I/B by the density of states at the Fermi level.) There is competition between the two terms in the exponent in the expression for the hopping probability. If $\beta\Delta E \gg 1$, electrons must hop long distances to find a state of the proper energy. If the hopping distance is to be small, the excitation energy can be large. This type of conduction is called *Mott variable-range hopping*.

The maximum hopping probability will be achieved when the magnitude of the exponent is minimized as a function of R:

$$\frac{\partial}{\partial R}\left(\frac{3\beta B}{4\pi n_I R^3} + 2\alpha R\right) = 0, \tag{7.116}$$

The most probable hopping distance is therefore

$$R_{\max} = \left(\frac{9\beta B}{8\pi n_I \alpha}\right)^{1/4}, \tag{7.117}$$

and the hopping probability and the conductivity are proportional to

$$\sigma = \sigma_0 \exp\left(-\frac{8}{3}\alpha R_{\max}\right) = \sigma_0 \exp\left[-\left(\frac{T_0}{T}\right)^{1/4}\right], \tag{7.118}$$

where $T_0 = 512B\alpha^3/9\pi n_I k_B$. Here σ_0 contains more slowly varying factors of temperature. Typical mobilities in a localized state are small.

Efros and Shklovskii introduced modifications to this formula arising from the Coulomb attraction between an electron that undergoes hopping and the hole it leaves behind. This changes the T dependence of the conductivity. A qualitative derivation of the effect is presented. The Coulomb interaction energy is $\Delta E_c = -e^2/4\pi\epsilon R$, where ϵ is the static dielectric permittivity. Thus there is a minimum distance that an electron must hop to overcome the Coulomb attraction. For an electron at the Fermi level to hop to a site a distance R away requires an energy input of amount ΔE such that $\Delta E + \Delta E_c > 0$, or else the electron will not reach a state above the Fermi energy. This leads to a connection between the minimum hopping distance and the energy:

$$R = \frac{e^2}{4\pi\epsilon\Delta E}. \tag{7.119}$$

The optimum hopping probability selects a value for ΔE given by

$$\frac{\partial}{\partial \Delta E}\left(\beta \Delta E + \frac{\alpha e^2}{2\pi\epsilon \Delta E}\right) = 0, \tag{7.120}$$

which leads to

$$\Delta E = \sqrt{\frac{\alpha e^2}{2\pi\epsilon\beta}}. \tag{7.121}$$

The conductivity therefore varies with temperature as

$$\sigma = \sigma_0 \exp\left(-\sqrt{\frac{T_{\mathrm{ES}}}{T}}\right), \tag{7.122}$$

where the characteristic temperature is $T_{\mathrm{ES}} = 2\alpha e^2/\pi\epsilon k_B$.

One expects the Efros–Shklovskii form for the temperature dependence to be valid at the lowest temperatures. At intermediate temperatures the Mott formula is more suitable. At high temperatures, if the energy is sufficient to excite the impurity electrons to the conduction band, the material will conduct as a doped semiconductor.

7.15 Poole–Frenkel Effect

The conductivity of an insulator is often not independent of the applied electric field and may be increased considerably by increasing the strength of the applied field. The long-range interaction of an electron with an impurity ion (with $z = 1$) in a host insulating or semiconducting crystal is given by the Coulomb interaction,

$$U_0(r) = -\frac{e^2}{4\pi\epsilon r}. \tag{7.123}$$

For ionization to occur, the electron must in principle be moved from the ion position to $r = \infty$, where U_0 vanishes. Let I_o denote the ionization potential of the impurity atom embedded in the host crystal. Now suppose that an electric field $\mathbf{E}$ is established in the crystal parallel to the z direction. The interaction becomes

$$U(r, \theta) = -\frac{e^2}{4\pi\epsilon r} - eEr\cos\theta, \tag{7.124}$$

where θ is the polar angle that $\mathbf{r}$ makes with the z direction. There is a saddle point in this function at the location

$$r = \sqrt{\frac{e}{4\pi\epsilon E}}, \qquad \theta = 0, \tag{7.125}$$

and the value of U there is

$$U = -\sqrt{\frac{e^3 E}{\pi\epsilon}}. \tag{7.126}$$

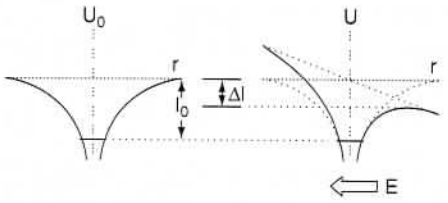


Figure 7.14. Potential energy without the electric field, U_0, and with the electric field, U.

For ionization to occur, one need only transport the electron to the location of the saddle point rather than all the way to infinity. The net result is that the ionization energy is lowered by an amount

$$\Delta I = -\sqrt{\frac{e^3 E}{\pi\epsilon}}. \tag{7.127}$$

The process of lowering the ionization barrier is illustrated in Fig. 7.14.

The ionization probability of the donor impurity, the number of ionized carriers, and the conductivity are all proportional to the Boltzmann factor $\exp(-\beta I_o)$. Hence the conductivity is enhanced and becomes

$$\sigma(E) = \sigma_0 \exp\left(\frac{1}{k_B T}\sqrt{\frac{e^3 E}{\pi\epsilon}}\right), \tag{7.128}$$

where σ_0 is the conductivity in the zero-field limit. This is called the *Poole–Frenkel effect*. It serves to show that nonohmic conductivity is likely to be important in describing charge transport in insulators at elevated field strengths. Dielectric breakdown is described in Chapter 15.

METAL–INSULATOR TRANSITION

There are materials that behave as metals for one range of parameters, but behave as insulators for another range. The separation between these two behaviors is often very sharp. Changing a physical parameter slightly can cause the material to undergo a metal–insulator transition. In this section several causes for this will be studied. The study begins with the phenomenon of percolation. One is concerned with an inhomogeneous system consisting of isolated conducting clusters (in three dimensions) or islands (in two dimensions). As their volume fraction increases, eventually they make contact with each other and the system can conduct electricity. This is followed by a discussion of the Mott transition, which is appropriate to homogeneous systems. In Section W7.5 the phenomenon of localization in solids is introduced, with attention being given to both weak and strong (Anderson) localization.

7.16 Percolation

Classical transport through a random medium may often be idealized by studying motion on a lattice. There are two conventional ways to introduce randomness: by

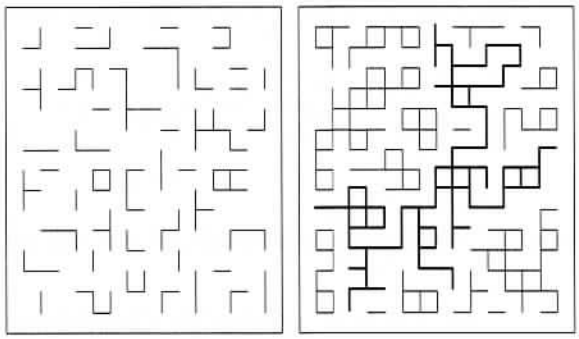


Figure 7.15. Random-bond model for a square lattice for the cases (a) $p < p_{cb}$ and (b) $p > p_{cb}$.

means of random-site or random-bond occupancy. In the *random-site* (RS) *model*, any site on the lattice is either occupied, with probability p, or vacant, with probability $q = 1 - p$. If NN sites are occupied, a path (bond) is drawn between these sites: otherwise, it is omitted. In the *random-bond* (RB) *model*, one distributes bonds randomly between all NN pairs of sites in the lattice, with probability p. In either case, connected clusters of bonds are formed. The distribution of cluster sizes is determined by the dimensionality of the lattice, d; the symmetry of the lattice; and the parameter p. If p is less than a critical probability (p_{cs} for RS, p_{cb} for RB), all clusters are finite in size. For $p > p_c$ (generic for p_{cs} or p_{cb}) there exists an infinite cluster spanning the lattice, in which case one says that *percolation* has occurred. Electrical conduction, for example, can occur only if there is percolation. In Fig. 7.15 a sketch is made of the RB model for prepercolation and percolation conditions on a square lattice. The percolated cluster is darkened in the figure.

Percolation theory is usually studied by means of simulations on a computer. Some crude insights into percolation theory can be obtained by studying a simple model that may be analyzed analytically. This is the tree structure called a *modified Bethe lattice* (Fig. 7.16). Each interior site (except the center) has Z nearest neighbors. There

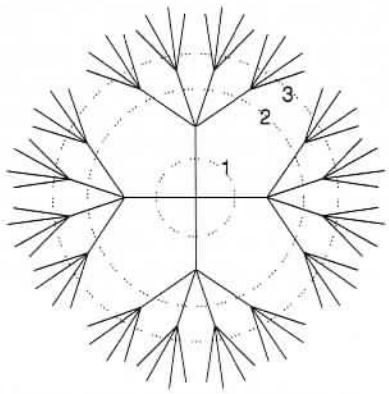


Figure 7.16. Modified Bethe lattice, illustrated for the case $Z = 5$.

are no closed loops in this structure. The dotted circles identify levels of increasing complexity in the lattice. The number of branches in zone n is $(Z-1)^n$. One wishes to know whether a particle starting at the center of the lattice will find its way to the outer perimeter by passing through occupied bonds. The main interest is in the limit as $n \to \infty$.

In the case $Z = 2$, the modified Bethe lattice is simply a single line passing through all the sites. The probability of reaching the nth level starting at the center is p^n. This will vanish as $n \to \infty$ for $p < 1$. For $Z = 2$, any break in the lattice is sufficient to prevent percolation.

The average number of paths that connect the center to the perimeter, N_p, is the product of the number of paths at level n, $(Z-1)^n$, multiplied by the probability of surviving the n steps to the perimeter, p^n:

$$N_p = [p(Z-1)]^n. \tag{7.129}$$

If $p(Z-1) < 1$, this will vanish in the limit $n \to \infty$. If $p(Z-1) > 1$, N_p will, in fact, be infinite. Therefore, $p_{cb} = 1/(Z-1)$ for the Bethe lattice.

For more familiar lattices in two and three dimensions, values for p_{cb} and p_{cs} are usually determined by computer simulations using Monte Carlo techniques. Values of these parameters are presented in Table 7.4. The critical cluster is found to be a fractal structure of noninteger dimension, D. An approximate formula for the RB critical probability is

$$p_{cb} \approx \frac{1}{Z}\frac{d}{d-1}. \tag{7.130}$$

One may think of the factor $d/(d-1)$ as a correction arising from the formation of closed loops in a finite-dimensional lattice.

A number of universal scaling relations pertain to the vicinity of the critical point, $p \approx p_c$. For example, the probability that a given bond belongs to the percolated cluster is $P \sim (p - p_c)^{\beta}$, with $\beta = (0.19, 0.41)$ for $d = (2, 3)$. The fraction of bonds that actually carry current (the *backbone*) is $P' \sim (p - p_c)^{\beta'}$, with $\beta' = (0.48, 1.05)$ for $d = (2, 3)$. For $p > p_c$, homogeneity of the lattice, on average, is maintained on distance scales larger than the *correlation length*, ξ. The correlation length obeys the scaling formula $\xi \sim |p - p_c|^{-\nu}$, with the critical exponent given by $\nu = (1.33, 0.88)$

TABLE 7.4 Percolation Thresholds for Some Lattices in Two and Three Dimensions

Lattic	Z	d	p_{cs}	D^a	p_{cb}
Honeycomb	3	2	0.653	91/48	0.696
Square	4	2	0.500	91/48	0.593
Triangular	6	2	0.347	91/48	0.500
Diamond	4	3	0.389	2.53	0.430
SC	6	3	0.249	2.53	0.312
BCC	8	3	0.180	2.53	0.246
FCC	12	3	0.198	2.53	0.119

Source: Data from M. Sahimi, *Applications of Percolation Theory*, Taylor & Francis, London, 1994.

[a]Fractal dimension of the critical percolation cluster.

for $d = (2, 3)$. For $p < p_c$, ξ also determines the size of a typical cluster. Note that the correlation length diverges at $p = p_c$.

Percolation is useful, for example, in analyzing the conductivity of a *cermet*, a ceramic material with embedded metallic clusters. For low cluster concentration the material is an insulator. As the concentration increases, one passes through the percolation threshold and the conductivity grows until it saturates at the value of appropriate to an amorphous metal.

7.17 Mott Metal–Insulator Transition

Imagine assembling a solid by starting with an array of atoms separated from each other by large distances and gradually decreasing the lattice constant until solid-state densities are achieved. When far apart, electrons are bound to their individual ions and the material is an insulator. When close together the interaction of an electron with its ion gets weakened by the presence of other electrons. If the other electrons were free to move around, for example, the ion potential will be shielded from its electron. For sufficiently strong shielding it is possible that the electron will no longer be bound to its ion and will delocalize over the solid. Therefore, a metal–insulator (M-I) transition may be expected to be seen. In more general situations a host insulator or semiconductor may be doped with impurity atoms and the conductivity studied as a function of increasing concentration of donors or acceptors.

The condition for the M-I transition density will be developed for the case of hydrogen atoms, although, strictly speaking, there is the need to consider atoms with an even number of electrons, so the bands have a possibility of being either full or empty. Begin by assuming that there is a concentration of free electrons, n, and look for the condition that the bound state is destroyed. In a metal the ion–electron interaction is given by the screened Coulomb interaction (again assuming that $Z = 1$)

$$V(r) = -\frac{e^2}{4\pi\epsilon r}\exp(-k_{\mathrm{TF}}r), \tag{7.131}$$

where $k_{\mathrm{TF}} = (3ne^2/2\epsilon E_F)^{1/2}$ is the inverse Thomas–Fermi screening length. Here ϵ is the electric permittivity of the host material and E_F is the Fermi energy.

The derivation of Eq. (7.131) proceeds from a semiclassical expression for the chemical potential of an electron

$$\frac{\hbar^2 k_F^2}{2m_e^*} - e\phi = \mu, \tag{7.132}$$

where $k_F = (3\pi^2 n)^{1/3}$ and ϕ is the potential due to the ion. Combining this with Poisson's equation,

$$\nabla^2\phi = \frac{e}{\epsilon}(n - n_0), \tag{7.133}$$

with n_0 being the background ion density, and linearizing about $n = n_0$ gives the Yukawa equation:

$$(\nabla^2 - k_{\mathrm{TF}}^2)\phi = -\frac{e}{\epsilon}\delta(\mathbf{r}), \tag{7.134}$$

whose solution is the screened Coulomb potential given in Eq. (7.131).

A variational estimate of the energy of a bound state may be made. Assume a wavefunction of the form $\psi(r) = N\exp(-\alpha r)$ with the normalization constant given by $N = \alpha^{3/2}/\sqrt{\pi}$. The expectation value of the energy in this state is

$$E = \left\langle \psi \left| \left(\frac{p^2}{2m_e^*} + V \right) \right| \psi \right\rangle = \frac{\hbar^2\alpha^2}{2m_e^*} - \frac{e^2\alpha^3}{\pi\epsilon(2\alpha + k_{\mathrm{TF}})^2}. \tag{7.135}$$

For small n, where k_{TF} is also small, there exists an absolute minimum in the E versus α curve for $\alpha > 0$ and the material is an insulator. For small α a minimum occurs at $\alpha = 0$ and the electron is delocalized over the solid so that the material is a metal. The metal–insulator transition occurs when a critical value for α is reached such that $E = 0$ and $\partial E/\partial\alpha = 0$ occur simultaneously, that is, when

$$\frac{4\pi\hbar^2\epsilon k_{\mathrm{TF}}}{m_e^* e^2} = k_{\mathrm{TF}} a_d = 1, \tag{7.136}$$

where $a_d = 4\pi\hbar^2\epsilon/m_e^* e^2$ is the Bohr radius for donors in the material.

CONDUCTIVITY OF REDUCED-DIMENSIONAL SYSTEMS

As integrated circuits become smaller and smaller, one begins to probe the intrinsic limits on the conduction process set by nature. Quantum wires refer to one-dimensional conductors whose length is comparable to or less than the elastic scattering length of the electrons passing through them. In this section two aspects of quantum wires are studied. First, a newly discovered system presented by nature, the carbon nanotube, is examined. Then the Landauer theory for conductivity will be developed and it will be shown that the resistance of a carbon nanotube is quantized.

7.18 Carbon Nanotubes

In graphite each carbon atom sits at a vertex of a planar hexagonal honeycomb lattice separated from its nearest neighbors by a bond distance $d = 0.142$ nm. The bonds are composed of sigmalike molecular orbitals arising from the interaction of sp^2 hybrid orbitals on the NN C atoms. There also delocalized π-like molecular orbitals. In the carbon nanotube most of the graphite sheet is wrapped into a seamless cylinder of finite diameter rather than extending over a plane. This requires an admixture of sp^3 hybrid orbitals, with the amount increasing as the radius of curvature decreases. The nanotubes currently fabricated in a carbon arc at a temperature of $T = 4000$ K are typically 1 μm in length, with diameters ranging from 1 to 20 nm. The ends of the tubes are often capped by an array of six carbon pentagons. In some cases one finds nested nanotubes with an intertubular separation of 0.34 nm, which is close to the distance between adjacent planes in graphite. Often, the nanotubes are bundled together to form filamentary fibers. Even closed rings have been found. The interest in these nanotubes stems from their high mechanical strength, light weight, and their potential for use as microscopic wires. When subjected to bending stresses, the nanotubes will deform, but they return to their original shape promptly when the stress is removed as long as no rupture occurs.

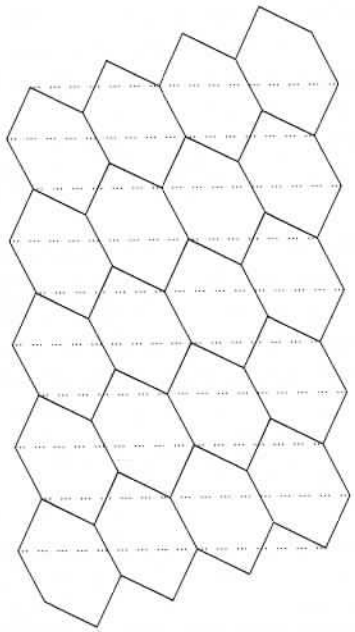


Figure 7.17. Section of graphite about to be wrapped into a chiral nanotube. The dotted lines denote points to be joined.

The nanotubes are observed to have varying degrees of chirality (i.e., one may think of rows of C atoms as if they were wire wound around a cylinder with various degrees of pitch). Consider the primitive honeycomb lattice with a row of parallel hexagons. Adjacent to this is another row of hexagons offset from the first row (Fig. 7.17), with the pattern repeated periodically. The hexagons are located by specifying the primitive vectors $\mathbf{u}_1 = \hat{i}d\sqrt{3}$, $\mathbf{u}_2 = d(\hat{i}\sqrt{3} + 3\hat{j})/2$, and writing $\mathbf{R}_{mn} = m\mathbf{u}_1 + n\mathbf{u}_2$, where m and n are integers. The chirality comes about during the folding into a cylinder when the hexagon at $\mathbf{R}_{00}$ is made to overlap the hexagon at $\mathbf{R}_{mn}$. The circumference of the cylinder is therefore $C = |m\mathbf{u}_1 + n\mathbf{u}_2|$ or $C = d\sqrt{3(m^2 + n^2 + mn)}$, and the diameter is $D = C/\pi$. This is illustrated in Fig. 7.17, where $m = 4$ and $n = 1$. In the limit where m and n become infinite, the nanotube becomes equivalent to a graphite sheet.

Depending on the numbers (m, n) determining chirality, the fiber can be either a semiconductor or a metal. This comes about as a consequence of the tight-binding model for graphite with NN interaction t. The honeycomb lattice has two inequivalent sites, with the π-electron wavefunction amplitudes on these sites being $A_{m,n}$ and $B_{m,n}$, in the unit cell designated by $\mathbf{R}_{m,n}$. The coupled tight-binding equations determining the energy eigenvalue, E, are

$$t(B_{m,n} + B_{m-1,n} + B_{m,n-1}) = EA_{m,n}, \tag{7.137a}$$

$$t(A_{m,n} + A_{m-1,n} + A_{m,n-1}) = EB_{m,n}. \tag{7.137b}$$

(see Sections 7.8 and 7.9). Letting

$$A_{m,n} = A\exp(i\mathbf{k}\cdot\mathbf{R}_{mn}) \tag{7.138a}$$

and

$$B_{m,n} = B \exp\left[i\mathbf{k} \cdot \mathbf{R}_{\mathrm{mn}} + i\mathbf{k} \cdot d\left(\hat{i}\frac{\sqrt{3}}{2} + \hat{j}\frac{3}{2} \right) \right], \tag{7.138b}$$

one finds that

$$E^2(k) = t^2 |1 + e^{i\mathbf{k}\cdot\mathbf{u}_1} + e^{i\mathbf{k}\cdot\mathbf{u}_2}|^2. \tag{7.139}$$

The bonding and antibonding energy bands for the delocalized π-electron system are given by

$$E_{\pm} = \pm t \sqrt{1 + 4\cos\frac{3k_y d}{2}\cos\frac{k_x d\sqrt{3}}{2} + 4\cos^2\frac{k_x d\sqrt{3}}{2}}. \tag{7.140}$$

The value of t is 2.66 eV, as determined by a fit to the graphite band structure.

Wrapping the structure into a chiral cylinder restricts the solutions to a one-dimensional band where the values of $\mathbf{k}$ are constrained by the periodicity condition $\mathbf{R}_{\mathrm{mn}} \cdot \mathbf{k} = 2\pi j$, where j is an integer. Instead of having a two-dimensional band structure extending over the hexagonal first Brillouin zone, the band structure is defined on a series of parallel lines in the Brillouin zone given by this formula. In Fig. 7.18, electronic band structures are presented for the cases $(m, n) = (8, 0)$ and $(m, n) = (7, 1)$. Note that in the former case there is a gap between the bonding and antibonding states, so the nanotube is a semiconductor. In the latter case there is no gap and the material is metallic.

The bonding and antibonding bands will be degenerate at some point in $\mathbf{k}$ space when

$$1 + e^{i\mathbf{k}\cdot\mathbf{u}_1} + e^{i\mathbf{k}\cdot\mathbf{u}_2} = 0. \tag{7.141}$$

If $\mathbf{k} \cdot \mathbf{u}_1 = 2\pi/3$ and $\mathbf{k} \cdot \mathbf{u}_2 = 4\pi/3$, this equation will be satisfied. (Another possibility is $\mathbf{k} \cdot \mathbf{u}_1 = 4\pi/3$ and $\mathbf{k} \cdot \mathbf{u}_2 = 2\pi/3$.) Coupling this with the periodicity condition, it is found that

$$\mathbf{k} \cdot \mathbf{R}_{\mathrm{mn}} = \frac{2\pi m}{3} + \frac{4\pi n}{3} = 2\pi j. \tag{7.142}$$

This will be satisfied for $m + 2n = 3j$. (In the other case it is $n + 2m = 3j$.) When this condition is obeyed, there is no gap between the bonding and antibonding bands, and the nanotube is a metal. When $m + 2n = 3j + 1$ or $m + 2n = 3j + 2$, there is a gap and it is a semiconductor.

It is possible to introduce defects into the nanotube in such a way that the number of edges, vertices, and faces is not changed. One way of doing this is to replace a pair of hexagons by a pentagon and a heptagon. Several other ways of accomplishing this are possible. When one introduces these defects, a bend is introduced, the diameter and helicity of the tubule changes, and the value of $m + 2n$ changes, since $m \to m \pm 1$ and $n \to n \mp 1$. It is thus possible to change metals to semiconductors and to fabricate metal–semiconductor junctions simply by inserting a pentagon–heptagon pair in place of a pair of hexagons. More generally, it should be possible to fabricate nanotube heterojunctions by controlling the defect locations. For example, the (7,1) tube is metallic, whereas the (8,0) tube is a semiconductor with a 1.2 eV energy gap. A sketch of the (8,0)/(7,1) junction is given in Fig. 7.19. Some additional features of carbon nanotubes may be found in Section W7.6.

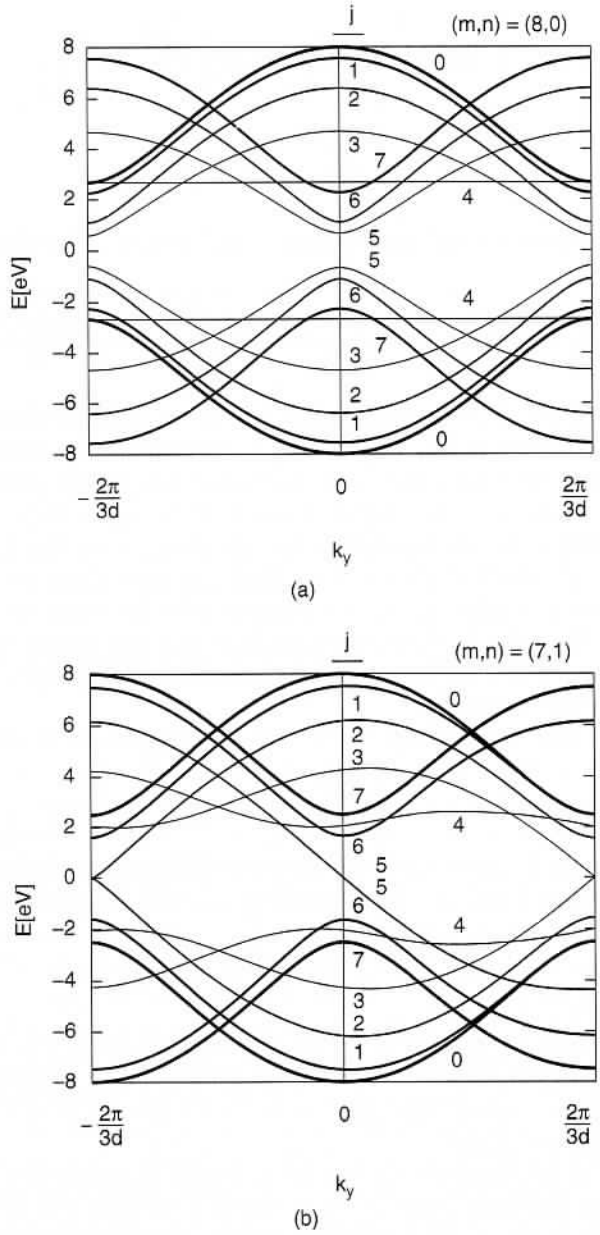


Figure 7.18. Band structures for carbon nanotubes for (*a*) $(m, n) = (8, 0)$ and (*b*) $(m, n) = (7, 1)$. The various curves correspond to values of j in the range 0 to 7.

There is evidence for *weak localization* in the electric conduction of nanotubes at low temperatures (see Section W7.5). This comes about because the probability that an electron entering a nanotube will backscatter is increased because of constructive interference between a given electron path and a time-reversed path. If backscattering is enhanced, it comes at the expense of transmission through the nanotube

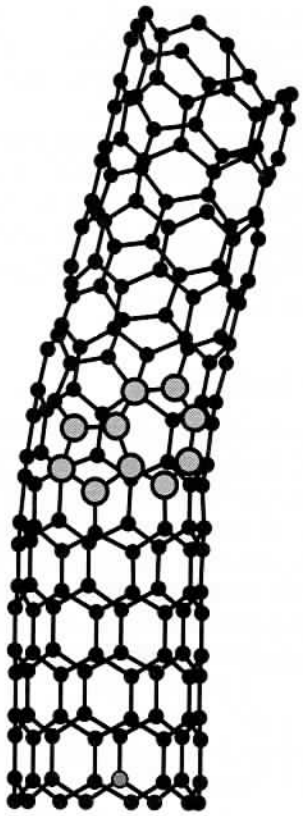


Figure 7.19. The (8,0)/(7,1) junction between two carbon nanotubes. [From L. Chico et al, *Phys. Rev. Lett.*, **76**, 971 (1996). Copyright 1996 by the American Physical Society.]

(i.e., conduction). Four-probe resistance measurements have shown a marked variation of the resistance of a nanotube with diameter and helical pitch. Resistances per unit length are typically in the range 10^4 to 10^5 Ω/m for diameters in the range 5 to 20 nm. Current densities as large as 6×10^{10} A/m^2 could be passed through a nanotube without causing irreversible damage. The resistance grows considerably when there are curves or kinks in the tubes.

There is also evidence[†] that carbon nanotubes can serve as quantum-mechanical wires in the sense that conduction occurs by having electrons tunnel through delocalized quantum states. When idealized as a strictly one-dimensional system of length L, the states of the wire are given by $E_n = (\hbar n \pi)^2/2mL^2$, with n a positive integer. The energy separation between states, $\Delta E = \hbar \pi v_F/L$, can be substantial (i.e., on the order of 0.6 meV for a length $L = 3$ μm). Coherence lengths in excess of 140 nm and perhaps as large as L have been inferred. For a Fermi velocity of 0.8×10^6 m/s this implies a coherence time of several tenths of a picosecond.

It is possible both to fill and coat carbon nanotubes with molten materials, and thus have them serve as templates. This occurs if the surface tension of the adsorbate

[†] S. J. Tans et al., *Nature*, **386**, 474 (1997).

is sufficiently low ($\gamma < 0.2$ N/m). For example, such materials as Pb, Bi, V_2O_5, Se, S, Cs, and Rb as well as oxides of Ni, Co, Fe, and U have been taken up by the nanotubes. The formation of metallic nanowires seems to be related to the presence of an incomplete electronic shell of the metal ions.

Recently, it was shown that arrays of carbon nanotubes can function as field-emission electron sources and may therefore play a future role in the construction of flat-panel displays. The large enhancement of the local electric field due to the sharpness of the tube (lightning-rod effect) allows them to function with very high efficiency, despite the nearly 5 eV work function of graphite. Additional properties of carbon nanotubes are discussed Chapter W21.

7.19 Landauer Theory of Conduction

A quantum-mechanical description of the conduction process in one-dimensional wires was developed by Landauer[†]. Rather than attempt to derive the general formula, in this section attention is focused on an idealized model for the single-walled carbon nanotube, and a formula for its conductance is developed. It will be seen that this conductance is quantized (i.e., comes in multiples of a basic conductance).

The carbon nanotube will be treated as a hollow cylinder described by the coordinates (z, ϕ) and the electron is assumed to move freely on the surface of the tube. The Schrödinger equation governing the electrons is

$$\left[-\frac{\hbar^2}{2m^*}\left(\frac{\partial^2}{\partial z^2}+\frac{1}{a^2}\frac{\partial^2}{\partial \phi^2}\right)-E\right]\psi(z,\phi)=0, \tag{7.143}$$

where a is the radius of the tube. Here m^* is the effective mass of an electron in the conduction band. One writes the wavefunction in the form $\psi = \exp[i(kz + \mu\phi)]$, where μ must be an integer to maintain single valuedness in the azimuthal direction, and

$$k=\sqrt{\frac{2m^*E}{\hbar^2}-\frac{\mu^2}{a^2}}. \tag{7.144}$$

Note that there are two values for μ, leading to the same k, as long as $\mu \neq 0$. Also note that for a given value of E, there is a maximum magnitude that μ can have, leading to a real value for k. Landauer viewed quantum currents as transmission processes. Let τ denote the transmission amplitude. The current is given by

$$I=-e\sum_s\sum_\mu\int_{-\infty}^{\infty}\frac{dk}{2\pi}v_z\Theta(k)f(E,T)[1-f(E-eV,T)]|\tau|^2, \tag{7.145}$$

where the unit step function $\Theta(k)$ ensures that there is only a forward current. The s-sum is over the two spin projections and the μ-sum is over allowable integers. Since the nanotube has a very long elastic-scattering length, it will be assumed that the

[†] R. Landauer, *Philos Mag.*, **21**, 863 (1970). See also D. S. Fisher and P. A. Lee, *Phys. Rev. B*, **23**, 6851 (1981).

transmission is perfect and $|\tau|^2 = 1$. One says that the conduction is ballistic. The voltage drop along the tube is denoted by V. For low-enough T, as in the case of metals, one may approximate the Fermi factor product as

$$\begin{aligned} f(E,T)[1 - f(E - eV, T)] &\approx f(E,T)\Theta(E - eV - E_F) \\ &\approx f(E,T)[\Theta(E - E_F) - eV\delta(E - E_F)] \\ &\approx -\frac{eV}{2}\delta(E - E_F). \end{aligned} \tag{7.146}$$

The current becomes

$$\begin{aligned} I &= \frac{e^2 V}{2h} \sum_s \sum_\mu \int_0^\infty \delta\left(E_\mu + \frac{\hbar^2\mu^2}{2m^*a^2} - E_F\right) dE_\mu \\ &= \frac{e^2 V}{2h} \sum_s \sum_\mu \Theta\left(E_F - \frac{\hbar^2\mu^2}{2m^*a^2}\right) \equiv GV, \end{aligned} \tag{7.147}$$

where G is the conductance. Thus

$$G = \frac{2e^2}{h}\left(n + \frac{1}{2}\right). \tag{7.148}$$

Here n is the number of subbands, defined by $|\mu|$, that can conduct.

Experiments on carbon nanotube quantum resistors[†] find $G = 2e^2/h = (12.9\ \text{k}\Omega)^{-1}$, consistent with $n = \frac{1}{2}$. Although this is not yet fully understood, it may be that the Coulomb blockade effect prevents a larger number of electrons from traversing the quantum wire.

REFERENCES

Ashcroft, N. W., and N. D. Mermin, *Solid State Physics*, Holt, Rinehart and Winston, New York, 1976.

Fletcher, G. C., *The Electron Band Theory of Solids*, North-Holland, Amsterdam, 1971.

Rossiter, P. L., *The Electrical Resistivity of Metals and Alloys*, Cambridge University Press, New York, 1991.

Sahimi, M., *Applications of Percolation Theory*, Taylor & Francis, London, 1994.

PROBLEMS

7.1 Graphite is a hexagonal crystal with a four-atom unit cell. In one plane there is a honeycomb lattice with lattice constant a = 0.246 nm. The adjacent layer has atoms sitting over the centers of the hexagons of the honeycomb lattice. The interplanar spacing is large (0.337 nm) (see Fig. 1.27). This allows the interplanar

[†] S. Frank et al, *Science*, **280**, 1744 (1998).

coupling to be neglected as a first approximation. Use the two-dimensional tight-binding approximation to obtain the band structure for graphite.

7.2 Use the tight-binding approximation to obtain the band structure for linear polyacetylene. You may assume that there is one mobile π-electron per carbon atom (refer to Fig. W14.7).

7.3 Electrons are confined to a two-dimensional sheet with areal density N. If a positive ion is placed on the sheet analyze the role played by the electrons in possibly screening the potential of the ion.

7.4 Derive the equations for J_x, J_y, and J_z in Section 7.13 for the magnetoresistance.

7.5 Show that the electronic density of states at E_F according to the free-electron model is $3z/2E_F$ per atom, $3N/4E_F$ per spin, and $3n/2E_F$ per unit volume (z = valence, N = number of electrons, n = density of electrons).

7.6 For energies E and temperatures T at which the Fermi–Dirac distribution $f(E, T) \ll 1$, show that $f(E, T)$ reduces to the classical Maxwell–Boltzmann distribution $\exp[-\beta(\mathbf{E} - \mu)]$. For Cu at $T = 300$ K, to what electron energies $\mathbf{E} - \mu$ does this correspond? What fraction of the electrons in Cu at $T = 300$ K are excited to these, or higher, energies?

7.7 Calculate the mean free path and mobility of electrons in Cu at $T = 300$ K using the electrical conductivity and the free-electron theory.

7.8 Calculate the drift velocity, $\langle v \rangle$, of electrons in Cu at $T = 300$ K in an electric field of 100 V/m. Compare this with the Fermi velocity and explain the difference.

7.9 Compute the isothermal compressibility $K = -(1/V)(\partial V/\partial P)_T$ of a free-electron gas.

7.10 Use a computer to calculate the π-electron energy spectrum of a fullerene molecule, C_{60}, assuming a tight-binding model with nearest-neighbor interactions. The molecule is in the shape of a regular dodecahedron (see Fig. 4.8).

7.11 Show that in the tight-binding approximation [see Eq. (7.93)] $\langle (E_0 - E)^2 \rangle = Z|t|^2$, where the average is taken over the band.

7.12 Calculate the enhancement of the conductivity $\sigma(E)/\sigma_0$ at $T = 300$ K according to the Poole–Frenkel effect [Eq. (7.128)] for electric field strengths $E = 10^6$, 10^7, and 10^8 V/m in α-SiO_2 and in NaCl.

7.13 Consider the high-field, $\omega_c \tau \gg 1$ limit of Eq. (7.109). Show that $\sigma_{xx} \to 0$ and $\sigma_{xy} = -ne/B = -\sigma_{yx}$ in this limit. The quantity $\sigma_{xy} = 1/R_H B$ is known as the *Hall conductivity*.