2. Electrical Conduction in Metals and Semiconductors

Safa Kasap, Cyril Koughia, Harry E. Ruda

Electrical transport through materials is a large and complex field, and in this chapter we cover only a few aspects that are relevant to practical applications. We start with a review of the semi-classical approach that leads to the concepts of drift velocity, mobility and conductivity, from which Matthiessen’s Rule is derived. A more general approach based on the Boltzmann transport equation is also discussed. We review the conductivity of metals and include a useful collection of experimental data. The conductivity of nonuniform materials such as alloys, polycrystalline materials, composites and thin films is discussed in the context of Nordheim’s rule for alloys, effective medium theories for inhomogeneous materials, and theories of scattering for thin films. We also discuss some interesting aspects of conduction in the presence of a magnetic field (the Hall effect). We present a simplified analysis of charge transport in semiconductors in a high electric field, including a modern avalanche theory (the theory of lucky drift). The properties of low-dimensional systems are briefly reviewed, including the quantum Hall effect.

A good understanding of charge carrier transport and electrical conduction is essential for selecting or developing electronic materials for device applications. Of particular importance are the drift mobility of charge carriers in semiconductors and the conductivity of conductors and insulators. Carrier transport is a broad field that encompasses both traditional bulk processes and, increasingly, transport in low dimensional or quantized structures. In other chapters of this handbook, Baranovskii describes hopping transport in low mobility solids such as insulators, Morigaki deals with the electrical properties of amorphous semiconductors and Gould discusses in detail conduction in thin films. In this chapter, we outline a semi-quantitative theory of charge transport suitable for a wide range of solids of use to materials researchers and engineers. We introduce theories of bulk transport followed by processes pertinent to ultra-fast transport and quantized transport in lower dimensional systems. The latter covers such phenomena as the Quantum Hall Effect, and Quantized Conductance and Ballistic Transport in Quantum Wires that has potential use in new kinds of devices. There are many more rigorous treatments of charge transport; those by Rossiter [2.1] and Dugdale [2.2] on metals, and Nag [2.3] and Blatt [2.4] on semiconductors are highly recommended.
2.1 Fundamentals: Drift Velocity, Mobility and Conductivity

Basic to the theory of the electronic structure of solids are the solutions to the quantum mechanical problem of an electron in a periodic potential known as Bloch waves. These wavefunctions are traveling waves and provide the physical basis for conduction. In the semi-classical approach to conduction in materials, an electron wavepacket made up of a superposition of Bloch waves can in principle travel unheaded in an ideal crystal. No crystal is ideal, however, and the imperfections cause scattering of the wavepacket. Since the interaction of the electron with the potential of the ions is incorporated in the Bloch waves, one can concentrate on the relatively rare scattering events which greatly simplifies the theory. The motion of the electrons between scattering events is essentially free (with certain provisos such as no interband transitions) subject only to external forces, usually applied electric or magnetic fields. A theory can then be developed that relates macroscopic and measurable quantities such as conductivity or mobility to the microscopic scattering processes. Principle in such a theory is the concept of mean free time \( \tau \) which is the average time between scattering events. \( \tau \) is also known as the conductivity relaxation time because it represents the time scale for the momentum gained from an external field to relax. Equivalently, \( 1/\tau \) is the average probability per unit time that an electron is scattered.

There are two important velocity quantities that must be distinguished. The first is the mean speed \( v \) or thermal velocity \( v_t \) which, as the name implies, is the average speed of the electrons. \( v \) is quite large being on the order of \( \sqrt{3k_B T/m_e^*} \approx 10^5 \text{ m/s} \) for electrons in a nondegenerate semiconductor and \( \sqrt{2E_F/m_e^*} \approx 10^6 \text{ m/s} \) for electrons in a metal, where \( k_B \) is Boltzmann’s constant, \( T \) is the temperature, \( E_F \) is the Fermi energy of the metal, and \( m_e^* \) is the electron effective mass. The distance an electron travels between scattering events is called the free path. It is straightforward to show that the average or mean free path for an electron is simply \( \ell = \mu v_t \). The second velocity is the mean or drift velocity \( v_d \) (variables in boldface are vectors) which is simply the vector average over the velocities of all \( N \) electrons,

\[
v_d = \frac{1}{N} \sum_{i=1}^{N} v_i.
\]

With no external forces applied to the solid, the electron motion is random and thus the drift velocity is zero. When subject to external forces like an electric field, the electrons acquire a net drift velocity. Normally, the magnitude of the drift velocity is much smaller than \( u \) so that the mean speed of the electron is not affected to any practical extent by the external forces. An exception is charge transport in semiconductors in high electric fields, where \( |v_d| \) becomes comparable to \( u \).

The drift velocity gives rise to an electric current. If the density of electrons is \( n \) then the current density \( J_e \) is

\[
J_e = -env_d
\]

(2.2)

where \( e \) is the fundamental unit of electric charge. For the important case of an applied electric field \( E \), the solutions of the semi-classical equations give a drift velocity that is proportional to \( E \). The proportionality constant is the drift mobility \( \mu_e \)

\[
v_d = -\mu_e E.
\]

(2.3)

The drift mobility might be a constant or it might depend on the applied field (usually only if the field is large). Ohm’s Law defines the conductivity \( \sigma \) of a material \( J = \sigma E \) resulting in a simple relation to the drift mobility

\[
\sigma = en\mu_e.
\]

(2.4)

Any further progress requires some physical theory of scattering. A useful model results from the simple assumption that the scattering randomizes the electron’s velocity (taking into proper account the distribution of electrons and the Pauli Exclusion Principle). The equation of motion for the drift velocity then reduces to a simple form

\[
\frac{dv_d}{dt} = F(t)/m_e^* - v_d/\tau,
\]

(2.5)

where \( F(t) \) is the sum of all external forces acting on the electrons. The effect of the scattering is to introduce a frictional term into what otherwise would be just Newton’s Law. Solutions of (2.5) depend on \( F(t) \). In the simplest case of a constant applied electric field, the steady-state solution is trivial,

\[
v_d = -eE\tau/m_e^*.
\]

(2.6)

The conductivity and drift mobility can now be related to the scattering time [2.5],

\[
\mu_e = e\tau/m_e^* \quad \text{and} \quad \sigma = ne^2\tau/m_e^*.
\]

(2.7)
More sophisticated scattering models lead to more accurate but more complicated solutions.

The simple expression (2.7) can be used to explain qualitative features of conduction in materials once a physical origin for the scattering is supplied. For any scattering site, the effective area is the scattering cross-section $S$ as depicted in Fig. 2.1. The scattering cross-section is related to the mean free path since the volume $S\ell$ must contain one scattering center. If there are $N_s$ scattering centers per unit volume then

$$\ell = \frac{1}{SN_s},$$

or, rewriting in terms of $\tau$,

$$\tau = \frac{1}{SN_s\ell}.$$

Once the cross-section for each physically relevant scattering mechanism is known then the effect on the scattering time and conductivity is readily calculated.

An overly restrictive assumption in the above analysis is that the electron’s velocity is completely randomized every time it is scattered. On the other hand, suppose that $v$ collisions are required to completely destroy the directional velocity information. That is only after an average of $v$ collisions do all traces of correlation between the initial and the final velocities disappear. The effective mean free path $\ell_{\text{eff}}$ traversed by the electron until its velocity is randomized will now be larger than $\ell$; to first order $\ell_{\text{eff}} = v\ell$. $\ell_{\text{eff}}$ is termed the effective or the conduction mean free path. The corresponding effective scattering cross section is

$$S_{\text{eff}} = \frac{1}{N_s\ell_{\text{eff}}}.$$

The expressions for mobility and conductivity become

$$\mu = \frac{e\nu\tau}{m_e^* u} = \frac{e\nu\ell}{m_e^* u} = \frac{e\ell_{\text{eff}}}{m_e^* u} = \frac{e}{m_e^* uN_sS_{\text{eff}}}$$

and

$$\sigma = \frac{e^2\nu\ell}{m_e^* u} = \frac{e^2\nu\ell_{\text{eff}}}{m_e^* u} = \frac{e^2\ell_{\text{eff}}}{m_e^* uN_sS_{\text{eff}}}.$$

Suppose that in a collision the electron is scattered at an angle $\theta$ to its original direction of travel as shown in Fig. 2.2. It is convenient to introduce a quantity $S_\theta(\theta)$, called the differential scattering cross section, defined so that $2\pi \sin \theta S_\theta(\theta) d\theta$ represents the probability of scattering at an angle between $\theta$ and $\theta + d\theta$ with respect to the original direction. If the magnitude of the velocity is not changed then the fractional change in component of the velocity along the original direction is $1 - \cos \theta$.

The average number of collisions $\nu$ required to randomize the velocity is then

$$\nu = \frac{1}{(1 - \cos \theta)},$$

where the average is given by

$$(1 - \cos \theta) = \frac{\int_0^{\pi} (1 - \cos \theta) S_\theta(\theta) \sin \theta d\theta}{\int_0^{\pi} S_\theta(\theta) \sin \theta d\theta}.$$

The effective cross sectional area $S_{\text{eff}}$ is then

$$S_{\text{eff}} = 2\pi \int_{\theta_m}^{\pi} (1 - \cos \theta) S_\theta(\theta) \sin \theta d\theta.$$

**Fig. 2.1** Scattering of an electron from a scattering center. The electron travels a mean distance $\ell = u\tau$ between collisions.

**Fig. 2.2** An electron moving in the $x$-direction becomes scattered though an angle $\theta$ with respect to the original direction.
As an example, consider conduction electrons scattering from charged impurities in a nondegenerate semiconductor where small angle deviations predominate. The differential cross section for coulombic scattering from a charged impurity center with charge $+Ze$ is

$$S_\theta (\theta) = \frac{16k^2}{u^2} \times \frac{1}{\theta^2}; \quad k = \frac{Ze^2}{4\pi\varepsilon_0\varepsilon_r m_e},$$

(2.16)

where $k = Ze^2/4\pi\varepsilon_0\varepsilon_r m_e^*\varepsilon_r$ and $\varepsilon_r$ is the relative permittivity of the semiconductor. Let $r_m$ be the maximum effective radius of action for the impurity at which the minimum scattering angle $\theta_m$ occurs. Then the integral in (2.15) evaluates to

$$S_{\text{eff}} = \frac{2\pi^2 k^2}{u^4} \ln \left(1 + \frac{r_m^2 u^2}{k^2}\right).$$

(2.17)

In a nondegenerate semiconductor, the equipartition theorem links velocity to temperature, $m_e^* u^2/2 = 3kT/2$, so that $u \propto T^{1/2}$. Thus, $S_{\text{eff}} = A/\bar{T}^2 \ln \left(1 + BT^2\right)$, (2.18)

where $A$ and $B$ are constants. The drift mobility due to scattering from ionized impurities becomes

$$\mu_1 = \frac{e}{m_e^* u_{\text{eff}}} \frac{n_s}{N_i} \propto \frac{T^2}{\bar{T}^{1/2}} \frac{A}{N_i} \ln(1 + BT^2),$$

(2.19)

where $N_i$ is the density of ionized impurities (it represents $N_s$), and $C$ is a new constant. At low temperatures where lattice scattering is insignificant, we expect $\mu_1 \propto T^{1/2}/N_i$ for nondegenerate semiconductors.

The above semiquantitative description is sufficient to understand the basic principles of conduction. A more rigorous approach involves solving the Boltzmann charge transport equation and is addressed in Sect. 2.6.

### 2.2 Matthiessen’s Rule

In general, the conduction electron whether in a metal or in a semiconductor can be scattered by a number of mechanisms, such as lattice vibrations, impurities, lattice defects such as dislocations, grain boundaries, vacancies, surfaces, or any other deviation from a perfectly periodic lattice. All these scattering processes increase the overall resistivity of the substance by reducing the mean scattering time. The relation between the types of scattering and the total scattering time can be obtained by considering scattering from lattice vibrations and impurities as shown in Fig. 2.3. We define two mean free times $\tau_L$ and $\tau_I; \tau_L$ is the mean free time considering only scattering from lattice vibrations (phonons) and $\tau_I$ is the mean free time considering only collisions with impurities. In a small unit of time $dt$, the total probability of scattering ($d\tau/d\tau$) is simply the sum of the probability for phonon scattering ($d\tau/d\tau_L$) and the probability for impurity scattering ($d\tau/d\tau_I$), and thus

$$\frac{1}{\tau} = \frac{1}{\tau_L} + \frac{1}{\tau_I}.$$  

(2.20)

We have assumed that neither $\tau_L$ nor $\tau_I$ is affected by the presence of the other scattering mechanism, that is each type of scattering is independent. The above expression can be generalized to include all types of independent scattering mechanisms yielding

$$\frac{1}{\tau} = \sum_i \frac{1}{\tau_i},$$

(2.21)

where $\tau_i$ is the mean scattering time considering the $i$th scattering process alone. Since the drift mobility $\mu_d$...
Electrical Conduction in Metals and Semiconductors | 2.3 Resistivity of Metals

2.3 Resistivity of Metals

23

2.3.1 General Characteristics

The effective resistivity of a metal, by virtue of Matthiessen’s rule, is normally written as

\[ \rho = \rho_I + \rho_R . \tag{2.24} \]

where \( \rho_R \) is called the residual resistivity and is due to the scattering of electrons by impurities, dislocations, interstitial atoms, vacancies, grain boundaries and so on. The residual resistivity shows very little temperature dependence whereas \( \rho_I \) is nearly linear in absolute temperature. \( \rho_I \) will be the main resistivity term for many high-quality, nonmagnetic pure, crystalline metals. In classical terms, we can take the thermal vibrations of a lattice atom with mass \( M \) as having a mean kinetic energy \( KE \) of \((1/2)Ma^2\omega^2\), where \( a \) and \( \omega \) are the amplitude and frequency of the vibrations. This mean KE must be of the order of \( kT \) so that the amplitude \( a \propto T^{1/2} \). Thus the electron scattering cross section \( S \approx \pi a^2 T \). Since the mean speed of conduction electron in a metal is the Fermi speed and is temperature insensitive, \( \mu \propto T \propto 1/S \propto T^{-1} \), and hence the resistivity \( \rho \propto T \). Most nonmagnetic pure metals obey this relationship except at very low temperatures. Figure 2.4 shows the resistivity of Cu as a function of temperature where above \( \approx 100 \mathrm{K} \), \( \rho \propto T \).

The \( \rho \propto T \) behavior can also be derived more rigorously by noting that at high temperatures, the phonon concentration \( n_{ph} \) increases as \( T \). The mean scattering time \( \tau \) is inversely proportional to \( n_{ph} \) so that \( \rho \propto T \). Frequently, the resistivity versus temperature behavior of pure metals can be empirically represented by a power law of the form

\[ \rho = \rho_0 \left( \frac{T}{T_0} \right)^n . \tag{2.25} \]

where \( \rho_0 \) is the resistivity at the reference temperature, \( T_0 \), and \( n \) is a characteristic index that best fits the data. For the nonmagnetic pure metals, \( n \) is close to unity whereas it is close to 2 for the magnetic metals Fe and Ni \([2.5]\). Figure 2.5 shows \( \rho \) versus \( T \) for various metals. Table 2.1 summarizes the values \( \rho_0 \) and \( n \) for various metals.

![Fig. 2.4](image-url) The resistivity of copper from low to high temperatures (near its melting temperature, 1358 K) on a log–log plot. Above about 100 K, \( \rho \propto T \), whereas at low temperatures, \( \rho \propto T^5 \), and at the lowest temperatures \( \rho \) approaches the residual resistivity \( \rho_R \). The inset shows the \( \rho \) versus \( T \) behavior below 100 K on a linear plot. \( \rho_R \) is too small to see on this scale.) (After [2.5])

### Table 2.1

<table>
<thead>
<tr>
<th>Metal</th>
<th>( \rho_0 ) (( \Omega \mathrm{m} ))</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ni</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ag</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Au</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
as a function of temperature above 0°C. Tin melts at 505 K, whereas nickel and iron go through a magnetic-to-nonmagnetic (Curie) transformation at about 627 K and 1043 K, respectively. The theoretical behavior ($\rho \propto T$) is shown for reference. (After [2.5])

As apparent from Fig. 2.4, below $\approx 100$ K, the $\rho \propto T$ behavior fails, and $\rho \propto T^2$. The reason is that, as the temperature is lowered, the scattering by phonons becomes less efficient, and it takes many more collisions to fully randomize the initial velocity of the electron. The mean number of collisions $\nu$ required to randomize the velocity scales with $T^{-2}$ [2.5], and at low temperatures, the concentration $n_{\text{phonon}}$ of phonons increases as $T^3$. Thus,

$$\sigma \propto \nu T \propto \frac{\nu}{n_{\text{phonon}}} \propto \frac{T^2}{T^3} \propto T^{-1}$$  \hspace{1cm} (2.26)

which explains the low-temperature $\rho \propto T$ behavior in Fig. 2.4. The low temperature $\rho \propto T^2$ and high temperature $\rho \propto T$ regimes are roughly separated by the Debye temperature $\theta_D$. For $T > \theta_D$ we expect $\rho \propto T$, and for $T < \theta_D$ we expect $\rho \propto T^2$.

In the case of metals with impurities and for alloys, we need to include the $\rho_R$ contribution to the overall resistivity. For $T > \theta_D$, the overall resistivity is

$$\rho \approx AT + \rho_R$$  \hspace{1cm} (2.27)

where $A$ is a constant, and the $AT$ term in (2.27) arises from scattering from lattice vibrations. Normally, $\rho_R$ has very little temperature dependence, and hence very roughly $\rho$ versus $T$ curves shift to higher values as $\rho_R$ is increases due to the addition of impurities, alloying or cold working the sample (mechanical deformation that generates dislocations) as illustrated for Cu–Ni alloys in Fig. 2.6.

### Table 2.1 Typical resistivities at 273 K (0°C) $\rho_0$ and thermal coefficients of resistivity $\alpha_0$ at 0°C for various metals above 0°C and below melting temperature. The resistivity index $n$ in $\rho = \rho_0(T/\theta_D)^n$ is also shown. Note that $n$ is fit to resistivity data above 273 K but below the melting temperature of the metal. Data selectively combined from various sources, including [2.6, 7].

<table>
<thead>
<tr>
<th>Metal</th>
<th>$\rho_0$ (nΩ m)</th>
<th>$n$</th>
<th>$\alpha_0 \times 10^{-3}$ (K$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminium, Al</td>
<td>24.2</td>
<td>1.20</td>
<td>4</td>
</tr>
<tr>
<td>Antimony, Sb</td>
<td>302</td>
<td>1.27</td>
<td>8</td>
</tr>
<tr>
<td>Beryllium, Be</td>
<td>30.2</td>
<td>1.73</td>
<td>8</td>
</tr>
<tr>
<td>Bismuth, Bi</td>
<td>1070</td>
<td>–</td>
<td>4.5</td>
</tr>
<tr>
<td>Cadmium, Cd</td>
<td>68</td>
<td>–</td>
<td>4.3</td>
</tr>
<tr>
<td>Calcium, Ca</td>
<td>31.1</td>
<td>1.09</td>
<td>4.0</td>
</tr>
<tr>
<td>Cerium, Ce</td>
<td>730</td>
<td>–</td>
<td>0.9</td>
</tr>
<tr>
<td>Cesium, Cs</td>
<td>187</td>
<td>1.23</td>
<td>4.8</td>
</tr>
<tr>
<td>Chromium, Cr</td>
<td>118</td>
<td>1.01</td>
<td>2.9</td>
</tr>
<tr>
<td>Cobalt, Co</td>
<td>56</td>
<td>–</td>
<td>6.6</td>
</tr>
<tr>
<td>Copper, Cu</td>
<td>15.4</td>
<td>1.16</td>
<td>4.3</td>
</tr>
<tr>
<td>Gold, Au</td>
<td>20.5</td>
<td>1.13</td>
<td>4.0</td>
</tr>
<tr>
<td>Hafnium, Ha</td>
<td>304</td>
<td>1.21</td>
<td>4.4</td>
</tr>
<tr>
<td>Indium, In</td>
<td>80</td>
<td>1.31</td>
<td>5.0</td>
</tr>
<tr>
<td>Iridium, Ir</td>
<td>47</td>
<td>–</td>
<td>4.5</td>
</tr>
<tr>
<td>Iron, Fe</td>
<td>85.7</td>
<td>1.73</td>
<td>6.3</td>
</tr>
<tr>
<td>Lead, Pb</td>
<td>192</td>
<td>1.14</td>
<td>4.2</td>
</tr>
<tr>
<td>Magnesium, Mg</td>
<td>40.5</td>
<td>1.07</td>
<td>4.2</td>
</tr>
<tr>
<td>Molybdenum, Mo</td>
<td>48.5</td>
<td>1.21</td>
<td>5.0</td>
</tr>
<tr>
<td>Nickel, Ni</td>
<td>61.6</td>
<td>1.76</td>
<td>6.5</td>
</tr>
<tr>
<td>Niobium, Nb</td>
<td>152</td>
<td>–</td>
<td>2.3</td>
</tr>
<tr>
<td>Palladium, Pd</td>
<td>97.8</td>
<td>0.94</td>
<td>0.39</td>
</tr>
<tr>
<td>Platinum, Pt</td>
<td>98.4</td>
<td>1.01</td>
<td>3.9</td>
</tr>
<tr>
<td>Rhodium, Rh</td>
<td>43</td>
<td>–</td>
<td>4.4</td>
</tr>
<tr>
<td>Ruthenium, Ru</td>
<td>71</td>
<td>–</td>
<td>4.1</td>
</tr>
<tr>
<td>Silver, Ag</td>
<td>14.7</td>
<td>1.13</td>
<td>4.1</td>
</tr>
<tr>
<td>Strontium, Sr</td>
<td>123</td>
<td>0.99</td>
<td>3.6</td>
</tr>
<tr>
<td>Tantalum, Ta</td>
<td>122</td>
<td>0.93</td>
<td>3.6</td>
</tr>
<tr>
<td>Tin, Sn</td>
<td>115</td>
<td>1.1</td>
<td>4.0</td>
</tr>
<tr>
<td>Titanium, Ti</td>
<td>390</td>
<td>1.01</td>
<td>4.8</td>
</tr>
<tr>
<td>Tungsten, W</td>
<td>48.2</td>
<td>1.24</td>
<td>4.8</td>
</tr>
<tr>
<td>Vanadium, V</td>
<td>181</td>
<td>1.02</td>
<td>3.9</td>
</tr>
<tr>
<td>Zinc, Zn</td>
<td>54.6</td>
<td>1.14</td>
<td>4.2</td>
</tr>
<tr>
<td>Zirconium, Zr</td>
<td>388</td>
<td>1.00</td>
<td>4.2</td>
</tr>
</tbody>
</table>

Fig. 2.5 The resistivities of various metals as a function of temperature above 0°C. Tin melts at 505 K, whereas nickel and iron go through a magnetic-to-nonmagnetic (Curie) transformation at about 627 K and 1043 K, respectively. The theoretical behavior ($\rho \propto T$) is shown for reference. (After [2.5])

In the case of metals with impurities and for alloys, we need to include the $\rho_R$ contribution to the overall resistivity.
Resistivity versus temperature behavior of nearly all metals is characterized by the temperature coefficient of resistivity (TCR) $\alpha_0$, which is defined as the fractional change in the resistivity per unit temperature increase at the reference temperature $T_0$, i.e.,

$$\alpha_0 = \frac{1}{\rho_0} \left( \frac{d\rho}{dT} \right)_{T=T_0},$$  \hspace{1cm} (2.28)

where $\rho_0$ is the resistivity at the reference temperature $T_0$, usually at 273 K (0°C) or 293 K (20°C), $d\rho = \rho - \rho_0$ is the change in the resistivity due to a small increase, $dT = T - T_0$, in temperature. Assuming that $\alpha_0$ is temperature independent over a small range from $T_0$ to $T$, we can integrate (2.28), which leads to the well known equation,

$$\rho = \rho_0[1 + \alpha_0(T - T_0)].$$  \hspace{1cm} (2.29)

Equation (2.29) is actually only valid when $\alpha_0$ is constant over the temperature range of interest which requires (2.27) to hold. Over a limited temperature range this will usually be the case. Although it is not obvious from (2.28), we should, nonetheless, note that $\alpha_0$ depends on the reference temperature, $T_0$, by virtue of $\rho_0$ depending on $T_0$.

It is instructive to mention that if $\rho \approx AT$ as we expect for an ideal pure metal, then $\alpha_0 = T_0^{-1}$. If we take the reference temperature $T_0$ as 273 K (0°C), then $\alpha_0$ should ideally be 1/(273 K) or 3.66 $\times$ 10$^{-3}$ K$^{-1}$. Examination of $\alpha_0$ for various metals shows that $\rho \propto T$ is not a bad approximation for some of the familiar pure metals used as conductors, e.g., Cu, Al, Au, but fails badly for others, such as indium, antimony and, in particular, the magnetic metals, e.g., iron and nickel.

Frequently we are given $\alpha_0$ at a temperature $T_0$, and we wish to use some other reference temperature, say $T'_0$, that is, we wish to use $\rho_0'$ and $\alpha_0'$ for $\rho_0$ and $\alpha_0$ respectively in (2.29) by changing the reference from $T_0$ to $T'_0$. Then we can find $\alpha_1$ from $\alpha_0$.

$$\alpha_0' = \frac{\alpha_0}{1 + \alpha_0(T'_0 - T_0)}$$  \hspace{1cm} (2.30)

$$\rho = \rho_0'[1 + \alpha_0'(T - T'_0)].$$

For example, for Cu $\alpha_0 = 4.31 \times 10^{-3}$ K$^{-1}$ at $T_0 = 0$ °C, but it is $\alpha_0 = 3.96 \times 10^{-3}$ K$^{-1}$ at $T_0 = 20$ °C. Table 2.1 summarizes $\alpha_0$ for various metals. Note that $\alpha_0$ in Table 2.1 is at 0 °C and can be converted to $\alpha_0'$ at 20°C using (2.30).

### 2.3.2 Fermi Electrons

The electrical properties of metals depend on the behavior of the electrons at the Fermi surface. The electron states at energies more than a few $kT$ below $E_F$ are almost fully occupied. The Pauli exclusion principle requires that an electron can only be scattered into an empty state, and thus scattering of deep electrons is highly suppressed by the scarcity of empty states (scattering where the energy changes by more than a few $kT$ is unlikely). Only the electrons near $E_F$ undergo scattering. Likewise, under the action of an external field, only the electron occupation near $E_F$ is altered. As a result, the density of states (DOS) near the Fermi level is most important for the metal electrical properties, and only those electrons in a small range $\Delta E$ around $E_F$ actually contribute to electrical conduction. The density of these electrons is approximately $g(E_F)\Delta E$ where $g(E_F)$ is the DOS at the Fermi energy. From simple arguments, the overall conductivity can be shown to be [2.5]

$$\sigma = \frac{1}{3}e^2 v_F^2 \tau g(E_F).$$  \hspace{1cm} (2.31)

where $v_F$ is the Fermi speed and $\tau$ is the scattering time of these Fermi electrons. According to (2.31), what is important is the density of states at the Fermi energy, $g(E_F)$. For example, Cu and Mg are metals with valencies I and II. Classically, Cu and Mg atoms each contribute 1 and 2 conduction electrons respectively into the crystal. Thus, we would expect Mg to have higher conductivity. However, the Fermi level in Mg is where the top tail of the 3p-band overlaps the bottom tail of the 3s band where the density of states is small.
In Cu, on the other hand, $E_F$ is nearly in the middle of the $4s$ band where the density of states is high. Thus, Mg has a lower conductivity than Cu.

The scattering time $\tau$ in (2.31) assumes that the scattered electrons at $E_F$ remain in the same energy band as, for example, in Cu, whose simplified energy band diagram around $E_F$ is shown in Fig. 2.7a. In certain metals, there are two different energy bands that overlap at $E_F$. For example, in Ni, 3d and 4s bands overlap at $E_F$ as shown in Fig. 2.7b. An electron can be scattered from the 4s to the 3d band and vice versa. Electrons in the 3d band have very low drift mobilities and effectively do not contribute to conduction so that only the 4s band electrons are scattered from the 4s to the 3d band and vice versa. This scattering time $\tau$ for the 4s band electrons is shortened and hence $\sigma$ from (2.31) is smaller. Thus, Ni has poorer conductivity than Cu.

Equation (2.31) does not assume a particular density of states model. If we now apply the free electron model for $g(E_F)$, and also relate $E_F$ to the total number of conduction electrons per unit volume $n$, we would find that the conductivity is the same as that in the Drude model, that is

$$\sigma = \frac{e^2 n \tau}{m_e}.$$  \hfill (2.32)

### 2.4 Solid Solutions and Nordheim’s Rule

In an isomorphous alloy of two metals, i.e., a binary substitutional solid solution, an additional mechanism of scattering appears, the scattering off solute phase atoms. This scattering contributes to lattice scattering, and therefore increases the overall resistivity. An important semi-empirical equation which can be used to predict the resistivity of an alloy is Nordheim’s rule. It relates the impurity part of the resistivity $\rho_I$ to the atomic fraction $X$ of solute atoms in a solid solution via

$$\rho_I = CX(1-X),$$  \hfill (2.33)

where the constant $C$ is termed the Nordheim coefficient and represents the effectiveness of the solute atom in increasing the resistivity. Nordheim’s rule was originally derived for crystals. Combining Nordheim’s rule with Matthiessen’s rule (2.23), the resistivity of an alloy of composition $X$ should follow

$$\rho = \rho_{\text{matrix}} + CX(1-X),$$  \hfill (2.34a)

where $\rho_{\text{matrix}}$ is the resistivity of the matrix due to scattering from thermal vibrations and from other defects, in the absence of alloying elements.

Nordheim’s rule assumes that the solid solution has the solute atoms randomly distributed in the lattice. For sufficiently small amounts of impurity, experiments show that the increase in the resistivity $\rho_I$ is nearly always simply proportional to the impurity concentration $X$, that is, $\rho_I \propto X$. For dilute solutions, Nordheim’s rule predicts the same linear behavior, that is, $\rho_I = CX$ for $X \ll 1$.

Originally the theoretical model for $\rho_I$ was developed by Nordheim [2.8] by assuming that the solute atoms simply perturb the periodic potential and thereby increase the probability of scattering. Quantum mechanical calculations for electron scattering within a single band, such as the s-band, at $E_F$ show that

$$\rho_I \propto g(E_F) V_{\text{scatter}}^2 X(1-X),$$  \hfill (2.35)

where $g(E_F)$ is the DOS at $E_F$, and $V_{\text{scatter}}$ is matrix element for scattering from one wavefunction to another at the Fermi surface in the same band, which for an s-band is

$$V_{\text{scatter}} = \langle \psi^*_s | \Delta V | \psi_s \rangle,$$  \hfill (2.36)

where $\Delta V$ is the difference between the potentials associated with solvent and solute atoms, and $\psi_s$ is the...
Table 2.2 Nordheim coefficients (at 20 °C) for dilute alloys obtained from $\rho_1 = CX$ and $X < 1$ at.%. Note: For many isomorphous alloys, $C$ may be different at higher concentrations; that is, it may depend on the composition of the alloy [2.7, 9]. Maximum solubility data from [2.10]

<table>
<thead>
<tr>
<th>Solute in solvent (element in matrix)</th>
<th>Nordheim coefficient ($n\Omega m$)</th>
<th>Maximum solubility at 25°C (at.-%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Au in Cu matrix</td>
<td>5500</td>
<td>100</td>
</tr>
<tr>
<td>Mn in Cu matrix</td>
<td>2900</td>
<td>24</td>
</tr>
<tr>
<td>Ni in Cu matrix</td>
<td>1250</td>
<td>100</td>
</tr>
<tr>
<td>Sn in Cu matrix</td>
<td>2900</td>
<td>0.6</td>
</tr>
<tr>
<td>Zn in Cu matrix</td>
<td>300</td>
<td>30</td>
</tr>
<tr>
<td>Cu in Au matrix</td>
<td>450</td>
<td>100</td>
</tr>
<tr>
<td>Mn in Au matrix</td>
<td>2410</td>
<td>25</td>
</tr>
<tr>
<td>Ni in Au matrix</td>
<td>790</td>
<td>100</td>
</tr>
<tr>
<td>Sn in Au matrix</td>
<td>3360</td>
<td>5</td>
</tr>
<tr>
<td>Zn in Au matrix</td>
<td>950</td>
<td>15</td>
</tr>
</tbody>
</table>

wavefunction of an electron in the s-band at $E_F$. It is clear that $C$ is only independent of $X$ if $g(E_F)$ and $V_{\text{scatter}}$ remain the same for various $X$ which may not be true. For example, if the effective number of free electrons increases with $X$, $E_F$ will be shifted higher, and $C$ will not be constant.

Table 2.2 lists some typical Nordheim coefficients for various additions to copper and gold. The value of the Nordheim coefficient depends on the type of solute and the solvent. A solute atom that is drastically different in size to the solvent atom will result in a bigger increase in $\rho_1$ and will therefore lead to a larger $C$. An important assumption in Nordheim’s rule in (2.33) is that the alloying does not significantly vary the number of conduction electrons per atom in the alloy. Although this will be true for alloys with the same valency, that is, from the same column in the Periodic Table (e.g., Cu–Au, Ag–Au), it will not be true for alloys of different valency, such as Cu and Zn. In pure copper, there is just one conduction electron per atom, whereas each Zn atom can donate two conduction electrons. As the Zn content in brass is increased, more conduction electrons become available per atom. Consequently, the resistivity predicted by (2.34) at high Zn contents is greater than the actual value because $C$ refers to dilute alloys. To get the correct resistivity from (2.34) we have to lower $C$, which is equivalent to using an effective Nordheim coefficient $C_{\text{eff}}$ that decreases as the Zn content increases. In other cases, for example, in Cu–Ni alloys, we have to increase $C$ at high Ni concentrations to account for additional electron scattering mechanisms that develop with Ni addition. Nonetheless, the Nordheim rule is still useful for predicting the resistivities of dilute alloys, particularly in the low-concentration region.

In some solid solutions, at some concentrations of certain binary alloys, such as 75% Cu–25% Au and 50% Cu–50% Au, the annealed solid has an orderly structure; that is, the Cu and Au atoms are not randomly mixed, but occupy regular sites. In fact, these compositions can be viewed as a pure compound – like the solids Cu$_3$Au and CuAu. The resistivities of Cu$_3$Au and CuAu will therefore be less than the same composition random alloy that has been quenched from the melt. As a consequence, the resistivity $\rho$ versus composition $X$ curve does not follow the dashed parabolic curve throughout; rather, it exhibits sharp falls at these special compositions, as illustrated in Fig. 2.8. The effective media approximation may be used as an effective tool to estimate the resistivities of inhomogeneous media.
2.5 Carrier Scattering in Semiconductors

At low electric fields, ionized impurity scattering and phonon scattering predominate. Other types of scattering include carrier-carrier scattering, inter-valley scattering, and neutral impurity scattering; these may generally be ignored to a first approximation.

For phonon scattering, both polar and non-polar phonon scattering should be considered. In polar scattering, short wavelength oscillations of atoms on different sub-lattices vibrating out of phase produce an effective dipole moment proportional to the bond polarity. Since such vibrational modes are optically active (since this dipole moment can interact with an incident electromagnetic field), this type of lattice scattering is usually referred to as polar optical phonon scattering. Since a sub-lattice is necessary for optical modes, this scattering mechanism is not present in elemental semiconductors such as Si, Ge, or diamond.

Non-polar phonon scattering comes from long wavelength oscillations in the crystal, involving small displacements of tens to thousands of atoms. The wavelength depends on the material and its elastic properties. Such modes are very similar to sound vibrations and are thus referred to as acoustic modes. The associated atomic displacements correspond to an effective built-in strain, with local change in the lattice potential, causing carrier scattering known as deformation potential acoustic phonon scattering. Since the change in potential is relatively small, the scattering efficiency is relatively low as compared with polar optical phonon scattering.

Each scattering process contributes to the drift mobility according to Matthiessen’s rule

$$\frac{1}{\mu_e} = \sum \frac{1}{\mu_i}$$

as discussed in Sect. 2.3 above. Figure 2.9 shows the contributions of each scattering process for n-type ZnSe – a material that has applications in optoelectronics. At room temperature, polar optical phonon scattering and ionized impurity scattering dominate. These processes depend on the carrier concentration. The curve for ionized impurity scattering decreases markedly with increasing carrier concentration owing to the increasing concentration of ionized donors that supply these carriers. The ionized impurity scattering mobility is roughly inversely proportional to the concentration of ionized impurities. However, as the carrier concentration increases, as in a degenerate semiconductor, the average energy per carrier also increases (i.e., carriers move faster on average) and thus carriers are less susceptible to being scattered from the ionized impurity centers. In contrast, the polar phonon scattering rate is determined by the number of participating phonons which depends on the thermal energy available to create a given quantized vibrational mode.

The temperature dependence of the mobility may be estimated by considering the effect of temperature on ionized impurity and phonon scattering and combining these mechanisms using Matthiessen’s rule. Phonon
scattering increases strongly with increasing temperature $T$ due to the increase in the number of phonons resulting in $T^{-3/2}$ dependence for the polar phonon mobility. For ionized impurity scattering, increasing the temperature increases the average carrier velocity and hence increases the carrier mobility for a set concentration of ionized impurities. Once the system is reached such that impurity ionization is complete, the ionized impurity based carrier mobility can be shown to increase with temperature $T$ as approximately, $T^{+3/2}$. At low temperatures, the mobility is basically determined by ionized impurity scattering and at high temperatures by phonon scattering leading to a peaked curve. Invoking the previous discussions for the dependence of the total mobility on carrier concentration, it is clear that the peak mobility will depend on the doping level, and the peak location will shift to higher temperatures with increased doping as shown in Fig. 2.10.

### 2.6 The Boltzmann Transport Equation

A more rigorous treatment of charge transport requires a discussion of the Boltzmann Transport Equation. The electronic system is described by a distribution function $f(k, r, t)$ defined in such a way that the number of electrons in a six-dimensional volume element $d^3r$ at time $t$ is given by $\frac{1}{2}\pi^3 f(k, r, t) d^3r$. In equilibrium, $f(k, r, t)$ depends only on energy and reduces to the Fermi distribution $f_0$ where the probability of occupation of states with momenta $+k$ equals that for states with $-k$, and $f_0(k)$ is symmetrical about $k = 0$, giving no net charge transport. If an external field acts on the system (i.e., non-equilibrium), the occupation function $f(k)$ will become asymmetric in $k$-space. If this non-equilibrium distribution function $f(k)$ is completely specified and appropriate boundary conditions supplied, the electronic transport properties can be completely determined by solving the steady state Boltzmann transport equation [2.12]

$$\mathbf{v} \cdot \nabla_f + \mathbf{k} \cdot \nabla f = \left( \frac{\partial f}{\partial t} \right)_c$$

(2.38)

where:

1. $\mathbf{v} \cdot \nabla f$ represents diffusion through a volume element $d^3r$ about the point $r$ in phase space due to a gradient $\nabla f$.
2. $\mathbf{k} \cdot \nabla f$ represents drift through a volume element $d^3k$ about the point $k$ in phase space due to a gradient $\nabla f$ (for example, $\hbar \mathbf{k} = e(E + \frac{1}{2} \mathbf{v} \times \mathbf{B})$ in the presence of electric and magnetic fields).
3. $\frac{\partial f}{\partial t}$ is the collision term and accounts for the scattering of electrons from a point $k$ (for example, this may be due to lattice or ionized impurity scattering).

Equation (2.38) may be simplified by using the relaxation time approximation

$$\left( \frac{\partial f}{\partial t} \right)_c = \frac{f(f - f_0)}{\tau}$$

(2.39)

which is based on the assumption that for small changes in $f$ carriers return to equilibrium in a characteristic time $\tau$, dependent on the dominant scattering mechanisms. Further simplifications of (2.38) using (2.39) applicable for low electric fields lead to a simple equation connecting the mobility $\mu$ to the average scattering time $\langle \tau \rangle$.

$$\mu \approx e \langle \tau \rangle / m^*$$

(2.40)

The details of calculations may be found in various advanced textbooks, for example Bube [2.13], Blatt [2.4]. The average scattering time may be calculated assuming a Maxwell–Boltzmann distribution function and a parabolic band

$$\langle \tau \rangle = \frac{2}{3k_B T} \int_0^\infty f(E) E^{3/2} e^{-E/k_B T} dE$$

(2.41)

Quantum mechanical perturbation theory can be used to calculate the carrier scattering rate for different processes $i$, giving,

$$\tau_i(E) = a E^{-\alpha}$$

(2.42)

where $a$ and $\alpha$ are constants and $E$ is the electron energy. Substituting (2.42) into (2.41) gives

$$< \tau > = \frac{4a \Gamma (5/2 - \alpha)}{3\pi^{1/2}(k_B T)^{\alpha}}$$

(2.43)

in terms of the gamma function $\Gamma$. Using this approach, the expressions for the mobility for the case of lattice and impurity scattering may be easily found

$$\mu_L \propto \frac{4e}{m^* \sqrt{3\pi k_B T}} T^{-\frac{1}{2}}$$

(2.44)

$$\mu_I \propto \frac{8ek_B^{3/2}}{N_i m^* \sqrt{\pi}} T^{+\frac{1}{2}}$$

(2.45)

where $N_i$ is the concentration of ionized impurities.
2.7 Resistivity of Thin Polycrystalline Films

Two new dominant scattering mechanisms must be considered in polycrystalline thin films – scattering by grain boundaries and scattering at the surface. The scattering by grain boundaries is schematically shown in Fig. 2.11. As a first approximation, the conduction electron may be considered free within a grain, but becomes scattered at the grain boundary. Its mean free path $\ell_{\text{grains}}$ is therefore roughly equal to the average grain size $d$. If $\lambda = \ell_{\text{crystal}}$ is the mean free path of the conduction electrons in the single crystal, then according to Matthiessen’s rule

$$\frac{1}{\ell} = \frac{1}{\ell_{\text{crystal}}} + \frac{1}{\ell_{\text{grains}}} = \frac{1}{\lambda} + \frac{1}{d}.$$  (2.46)

The resistivity is inversely proportional to the mean free path which means the resistivity of the single crystal $\rho_{\text{crystal}} \propto 1/\lambda$ and resistivity of the polycrystalline sample $\rho \propto 1/\ell$. Thus,

$$\frac{\rho}{\rho_{\text{crystal}}} = 1 + \left(\frac{\lambda}{d}\right).$$  (2.47)

Figure 2.12 clearly demonstrates that even simple considerations agree well with experimental data. However, in a more rigorous theory we have to consider a number of effects. It may take more than one scattering at a grain boundary to totally randomize the velocity so that we need to calculate the effective mean free path that accounts for how many collisions are needed to randomize the velocity. There is a possibility that the electron may be totally reflected back at a grain boundary (bounce back). Let $\sigma$ be the conductivity of the polycrystalline (grainy) material and $\sigma_{\text{crystal}}$ be the conductivity of the bulk single crystal. Suppose that the probability of reflection at a grain boundary is $R$. If $d$ is the average grain size (diameter) then the two conductivities may be related using the Mayadas and Shatzkes formula [2.15]

$$\frac{\sigma}{\sigma_{\text{crystal}}} = 1 - \frac{3}{2} \beta + 3 \beta^2 - 3 \beta^3 \ln\left(1 + \frac{1}{\beta}\right)$$

where $\beta = \frac{\lambda}{d} \left(\frac{R}{1-R}\right)$.  (2.48)

The $\beta$ in (2.48) represents the $\lambda/d$ ratio adjusted for the reflection to transmission ratio of the electron at the grain boundary. When the grain size is large, $\beta$ is small.
the conductivity accordance with Fuchs-Sondheimer equation \[2.17, 18\]

Equation (2.48) for a Cu film with \(R\) values are \(0\):

\[
\frac{\rho}{\rho_{\text{crystal}}} \approx 1 + \frac{3}{2} \beta .
\]  

(2.49)

For highly polycrystalline films, the grain size would be small and \(\beta \gg 1\)

\[
\frac{\rho}{\rho_{\text{crystal}}} \approx \frac{4}{3} \beta .
\]  

(2.50)

Equation (2.49) implies that the Matthiessen’s rule is reasonably well obeyed when the grains are larger than the mean free path. For copper, typically \(R\) values are 0.24–0.40, and \(R\) is somewhat smaller for Al. Equation (2.48) for a Cu film with \(R \approx 0.3\) predicts \(\rho/\rho_{\text{crystal}} \approx 1.21\) for roughly \(d \approx 3\lambda\) or a grain size \(d \approx 120\) nm since in the bulk crystal \(\lambda \approx 40\) nm. Tellier et al. [2.16] have given extensive discussions of grain boundary scattering limited resistivity of thin films.

Scattering from the film surfaces must also be included in any resistivity calculation. It is generally assumed that the scattering from a surface is partially inelastic, that is the electron loses some of the velocity gained from the field. The inelastic scattering is also called nonspecular. (If the electron is elastically reflected from the surface just like a rubber ball bouncing from a wall, then there is no increase in the resistivity.) If the parameter \(p\) is the fraction of surface collisions which are specular (elastic) and if the thickness of film \(D\) is greater than \(\lambda\), and \(\sigma_{\text{bulk}} = 1/\rho_{\text{bulk}}\), then in accordance with Fuchs-Sondheimer equation [2.17, 18] the conductivity \(\sigma\) of the film is

\[
\frac{\sigma}{\sigma_{\text{bulk}}} = 1 - \frac{3\lambda}{8D}(1-p) , \quad (\lambda/D > 1) .
\]  

(2.51)

If \(D\) is much shorter than \(\lambda\),

\[
\frac{\sigma}{\sigma_{\text{bulk}}} = \frac{3D}{4\lambda} \ln \left[ \frac{\lambda}{D} + 0.423 \right]
\times (1 + 2p) , \quad (\lambda/D \ll 1) .
\]  

(2.52)

For purely nonspecular (inelastic) scattering, an approximate estimate can be obtained from

\[
\frac{\rho}{\rho_{\text{bulk}}} \approx 1 + \frac{3}{8} \left( \frac{\lambda}{D} \right) .
\]  

(2.53)

Figure 2.13 shows the resistivity of Cu films as a function of film thickness \(D\). The Cu thin films in this case are single crystal layers grown on the (001) surface of a single crystal of MgO (which is the substrate). As the film is a single crystal, there is no grain boundary scattering, and the observed increase in the resistivity with decreasing film thickness is due to the scattering of the electrons from the film surfaces. The experiments in Fig. 2.13 can be explained by the simplified Fuchs-Sondheimer equation with an average \(p = 0.20\).

For elastic or specular scattering \(p = 1\) and there is no change in the conductivity. The value of \(p\) depends on the film preparation method (e.g., sputtering, epitaxial growth etc.) and the substrate on which the film has been deposited. Table 2.3 summarizes the resistivities of thin Cu and Au gold films deposited by various preparation techniques. Notice the large difference between the Au films deposited on a noncrystalline glass substrate and on a crystalline mica substrate. Such differences between films are typically attributed to different values of \(p\). The \(p\)-value can also change (increase) when the film is annealed. Obviously, the polycrystallinity of the film will also affect the resistivity as discussed above. Typically, most epitaxial thin films, unless very thin (\(D \ll \ell\), deposited onto heated crystalline substrates exhibit highly specular scattering with \(p = 0.8 - 1\).

It is generally very difficult to separate the effects of surface and grain boundary scattering in thin polycrystalline films; the contribution from grain boundary scattering is likely to exceed that from the surfaces. In any event, both contributions, by Matthiessen’s general rule, increase the overall resistivity. Figure 2.12 shows an example in which the resistivity \(\rho_{\text{film}}\) of thin Cu polycrystalline films is due to grain boundary scattering, and thickness has no effect (\(D\) was 250–900 nm and much greater than \(\lambda\)). The resistivity \(\rho_{\text{film}}\) is plotted against the reciprocal mean grain size \(1/D\), which then follows the...
2.8 Inhomogeneous Media: Effective Media Approximation

The effective media approximation (EMA) attempts to estimate the properties of inhomogeneous mixture of two or more components using the known physical properties of each component. The general idea of any EMA is to substitute for the original inhomogeneous mixture some imaginary homogeneous substance – the effective medium (EM) – whose response to an external excitation is the same as that of the original mixture. The EMA is widely used for investigations of non-uniform objects in a variety of applications such as composite materials [2.28, 29], microcrystalline and amorphous semiconductors [2.30–33], light scattering [2.34], conductivity of dispersed ionic semiconductors [2.35] and many others.

Calculations of the conductivity and dielectric constant of two component mixtures are reviewed by Reynolds and Hough [2.36] and summarized by Rossiter [2.1]. For such a mixture we assume that the two components \( \alpha \) and \( \beta \) are randomly distributed in space with volume fractions of \( \chi_\alpha \) and \( \chi_\beta = 1 - \chi_\alpha \).

The dielectric properties are described by an effective permittivity \( \varepsilon_{\text{eff}} \) given by the ratio

\[
\varepsilon_{\text{eff}} = \frac{\langle D \rangle}{\langle E \rangle},
\]

where \( \langle E \rangle \) is the average electric field and \( \langle D \rangle \) is the average displacement field. The displacement field averaged over a large volume may be calculated from

\[
\langle D \rangle = \frac{1}{V} \int D \, dv = \frac{1}{V} \left( \int_{V_\alpha} D \, dv + \int_{V_\beta} D \, dv \right) = \chi_\alpha \langle D_\alpha \rangle + \chi_\beta \langle D_\beta \rangle,
\]

where \( \langle D_\alpha \rangle \) and \( \langle D_\beta \rangle \) are the average displacement fields inside regions of the respective components and \( V_\alpha \) and \( V_\beta \) are their volumes. Likewise the electric field...
is given by

\[
\langle E \rangle = \chi_\alpha \langle E_\alpha \rangle + \chi_\beta \langle E_\beta \rangle.
\]

From (2.54) one gets

\[
\varepsilon_{\text{eff}} = \varepsilon_\beta + (\varepsilon_\alpha - \varepsilon_\beta) \chi_\alpha f_\alpha
\]

or

\[
(\varepsilon_{\text{eff}} - \varepsilon_\alpha) \chi_\alpha f_\alpha + (\varepsilon_{\text{eff}} - \varepsilon_\beta) \chi_\beta f_\beta = 0
\]

where \(\varepsilon_\alpha\) and \(\varepsilon_\beta\) are the permittivities of the components and \(f_\alpha = \langle E_\alpha \rangle / \langle E \rangle\) and \(f_\beta = \langle E_\beta \rangle / \langle E \rangle\) are so-called field factors. The choice between (2.57) and (2.58) depends on particle geometry. Equation (2.57) is better when the particles of component are dispersed in a continuous medium \(\beta\). Equation (2.58) is preferred when the particle size of the two components is of the same order of magnitude.

The field factors can be calculated analytically only for phase regions with special specific geometries. The field factor for ellipsoids is given by (Stratton [2.48])

\[
f_\alpha = \frac{\cos^2 \alpha_i}{\sum_{i=1}^{3} \left( 1 + A_i \left( \frac{\alpha_i}{\varepsilon_\alpha} - 1 \right) \right)},
\]

where \(\alpha_i\) are the angles between the ellipsoid axes and the applied field and the constants \(A_i\) depend on the axial ratios of the ellipsoids; they must satisfy

\[
\sum_{i=1}^{3} A_i = 1.
\]

For a sphere, \(A_2 = A_1 = A\) and \(A_1 = 1 - 2A\). For a random orientation of spheroids, \(\cos^2 \alpha_1 = \cos^2 \alpha_2 = \cos^2 \alpha_3 = \frac{1}{3}\). In the case of long particles with aligned axes, \(\cos^2 \alpha_1 = \cos^2 \alpha_2 = \frac{1}{2}\) and \(\cos^2 \alpha_3 = 0\). The values of parameters entering (2.59) can be found in Table 2.4 which shows a set of mixture rules, i.e., a set of formulae allowing one to calculate \(\varepsilon_{\text{eff}}\) for some specific cases (such as spheres, rods, lamellae, etc.). The presence of some degree of orientation somewhat simplifies the calculations as shown in the Table 2.5.

The same formulae can be used to calculate the conductivity of mixtures by substituting the appropriate conductivity \(\sigma\) for \(\varepsilon\). For some special cases, the mixture rules of Table 2.5 lead to very simple formul\(\text{ae}\) which allows one to calculate the conductivity of inhomogeneous alloys with those specific geometries. These mixture rules are summarized in Table 2.6.

The most general approach to calculating the effective dielectric permittivity comes from

\[
\varepsilon_{\text{eff}} = \varepsilon_\beta \left( 1 - \frac{1}{\varepsilon_\alpha} \int_0^L \frac{G(L)}{1 - L} dL \right),
\]

where \(G(L)\) is given by

\[
G(L) = \sum_{i=1}^{3} \left( 1 + A_i \left( \frac{\alpha_i}{\varepsilon_\alpha} - 1 \right) \right).
\]

### Table 2.4: Mixture rules for randomly oriented particles. (Compiled from Reynolds and Hough [2.36])

<table>
<thead>
<tr>
<th>Particle shape</th>
<th>Mixture rule</th>
<th>Factors in (2.59)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spheres</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_\beta + \frac{\varepsilon_\alpha - \varepsilon_\beta}{\varepsilon_\alpha + 2\varepsilon_\beta} )</td>
<td>(A) (\varepsilon_\alpha)</td>
<td>[2.37–41]</td>
</tr>
<tr>
<td>Spheres</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_\beta + \frac{3}{\varepsilon_\alpha + 2\varepsilon_\beta} )</td>
<td>(A) (\varepsilon_\alpha)</td>
<td>[2.42]</td>
</tr>
<tr>
<td>Spheres</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_\beta + \frac{\varepsilon_\alpha - \varepsilon_\beta}{\varepsilon_\alpha + 2\varepsilon_\beta} )</td>
<td>(A) (\varepsilon_\alpha)</td>
<td>[2.43]</td>
</tr>
<tr>
<td>Spheres</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_\beta + \frac{\varepsilon_\alpha - \varepsilon_\beta}{\varepsilon_\alpha + 2\varepsilon_\beta} )</td>
<td>(A) (\varepsilon_\alpha)</td>
<td>[2.44]</td>
</tr>
<tr>
<td>Spheres</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_\beta + \frac{\varepsilon_\alpha - \varepsilon_\beta}{\varepsilon_\alpha + 2\varepsilon_\beta} )</td>
<td>(A) (\varepsilon_\alpha)</td>
<td>[2.45]</td>
</tr>
<tr>
<td>Spheres</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_\beta + \frac{1}{3} \sum_{i=1}^{3} \frac{\varepsilon_\alpha - \varepsilon_\beta}{\varepsilon_\alpha + 2\varepsilon_\beta} )</td>
<td>(A) (\varepsilon_\alpha)</td>
<td>[2.46]</td>
</tr>
<tr>
<td>Lamellae</td>
<td>( \varepsilon_{\text{eff}} = \frac{1}{2} \frac{(\varepsilon_\alpha \varepsilon_\beta + \varepsilon_\alpha \varepsilon_\beta - \varepsilon_\alpha)}{\varepsilon_\alpha + \varepsilon_\beta} )</td>
<td>(A) (\varepsilon_\alpha)</td>
<td>[2.43]</td>
</tr>
<tr>
<td>Rods</td>
<td>( \varepsilon_{\text{eff}} = \frac{1}{2} \frac{(5 \varepsilon_\alpha - 4 \varepsilon_\beta)}{\varepsilon_\alpha + \varepsilon_\beta} )</td>
<td>(A) (\varepsilon_\alpha)</td>
<td>[2.47]</td>
</tr>
</tbody>
</table>
Table 2.5  Mixture rules for partially oriented particles. (Compiled from Reynolds and Hough [2.36])

<table>
<thead>
<tr>
<th>Particle shape</th>
<th>Formula</th>
<th>Factors in (2.59) A</th>
<th>( \rho ) *</th>
<th>( \cos \alpha_1 ) = ( \cos \alpha_2 ) = ( \cos \alpha_3 )</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel cylinders ( \varepsilon )</td>
<td>( \varepsilon_{\text{eff}} - \varepsilon_{\beta} ) = ( \frac{x_\alpha - \varepsilon_{\text{eff}}}{x_\alpha + \varepsilon_{\beta}} + \frac{x_\beta - \varepsilon_{\text{eff}}}{x_\beta + \varepsilon_{\text{eff}}} ) ( \frac{1}{2} )</td>
<td>( \rho_{\text{eff}} = \rho_{\alpha} + L_{\beta} \rho_{\beta} )</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>[2.40, 41]</td>
</tr>
<tr>
<td>Parallel cylinders ( \varepsilon )</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_{\alpha} \frac{x_\alpha - \varepsilon_{\alpha}}{x_\alpha + \varepsilon_{\alpha}} + \varepsilon_{\beta} \frac{x_\beta - \varepsilon_{\beta}}{x_\beta + \varepsilon_{\beta}} ) ( \frac{1}{2} )</td>
<td>( \rho_{\text{eff}} = \rho_{\alpha} + L_{\beta} \rho_{\beta} )</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>[2.43]</td>
</tr>
<tr>
<td>Parallel lamellae ( \varepsilon )</td>
<td>( \varepsilon_{\text{eff}} = \frac{\varepsilon_{\alpha} x_{\alpha} + \varepsilon_{\beta} x_{\beta}}{x_{\alpha} + \varepsilon_{\alpha} + \varepsilon_{\beta}} ) ( \frac{1}{2} )</td>
<td>( \rho_{\text{eff}} = \rho_{\alpha} + L_{\beta} \rho_{\beta} )</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>[2.43]</td>
</tr>
<tr>
<td>Lamellae with all axes aligned ( \varepsilon )</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_{\alpha} + \varepsilon_{\beta} ) ( \frac{1}{2} )</td>
<td>( \rho_{\text{eff}} = \rho_{\alpha} + L_{\beta} \rho_{\beta} )</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>[2.49]</td>
</tr>
<tr>
<td>Lamellae with all axes aligned ( \varepsilon )</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_{\alpha} x_{\alpha} + \varepsilon_{\beta} x_{\beta} ) ( \frac{1}{2} )</td>
<td>( \rho_{\text{eff}} = \rho_{\alpha} + L_{\beta} \rho_{\beta} )</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>[2.49, 50]</td>
</tr>
<tr>
<td>Spheroids with all axes aligned ( \varepsilon )</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_{\alpha} + \varepsilon_{\beta} ) ( \frac{1}{2} )</td>
<td>( \rho_{\text{eff}} = \rho_{\alpha} + L_{\beta} \rho_{\beta} )</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>[2.51]</td>
</tr>
<tr>
<td>Spheroids with all axes aligned ( \varepsilon )</td>
<td>( \varepsilon_{\text{eff}} = \varepsilon_{\alpha} + \varepsilon_{\beta} ) ( \frac{1}{2} )</td>
<td>( \rho_{\text{eff}} = \rho_{\alpha} + L_{\beta} \rho_{\beta} )</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>[2.52]</td>
</tr>
</tbody>
</table>

Table 2.6  Conductivity/resistivity mixture rules. (After [2.5])

<table>
<thead>
<tr>
<th>Particle shape</th>
<th>Formula</th>
<th>Commentary</th>
<th>( \rho_{\text{eff}} = \rho_{\alpha} + L_{\beta} \rho_{\beta} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lamellae with all axes aligned ( \varepsilon )</td>
<td>( \rho_{\text{eff}} = \rho_{\alpha} + L_{\beta} \rho_{\beta} )</td>
<td>Resistivity mixture rule: ( \rho_{\alpha} ) and ( \rho_{\beta} ) are the resistivities of two phases and ( \rho_{\text{eff}} ) is the effective resistivity of mixture</td>
<td></td>
</tr>
<tr>
<td>Lamellae with all axes aligned ( \varepsilon )</td>
<td>( \sigma_{\text{eff}} = \sigma_{\alpha} + L_{\beta} \sigma_{\beta} )</td>
<td>Conductivity mixture rule: ( \sigma_{\alpha} ) and ( \sigma_{\beta} ) are the conductivities of two phases and ( \sigma_{\text{eff}} ) is the effective conductivity of mixture</td>
<td></td>
</tr>
<tr>
<td>Small spheroids ( \varepsilon ) ( \beta )-phase</td>
<td>( \rho_{\text{eff}} = \frac{1 + 2x_{\alpha}}{1 - x_{\alpha}} \rho_{\beta} )</td>
<td>( \rho_{\alpha} &gt; 10\rho_{\beta} )</td>
<td></td>
</tr>
<tr>
<td>Small spheroids ( \varepsilon ) ( \beta )-phase</td>
<td>( \rho_{\text{eff}} = \rho_{\beta} \frac{1 - x_{\alpha}}{1 + 2x_{\alpha}} )</td>
<td>( \rho_{\alpha} &lt; 0.1\rho_{\beta} )</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.7  Mixture rules and corresponding spectral functions \( G(L) \)

Mixture rule by Bruggemann [2.43]:

\[
G(L) = \frac{3x_{\alpha} - 1}{2x_{\alpha}} 2^{(L - 1)/2} \theta(3x_{\alpha} - 1) + \frac{3}{4\pi x_{\alpha}} 2^{(L - 1)/2} \left( \sqrt{(L - L^-)} (L + L^-) \right) \theta(L - L^-) \theta(L^+ - L) \theta(L^- - L) = 0
\]

where \( L^+/- = \frac{1}{3} \left( 1 + x_{\alpha} \pm 2\sqrt{(2x_{\alpha} - 2)^2} \right) \)

Mixture rule by Maxwell-Garnett [2.53]:

\[
G(L) = \frac{x_{\alpha} - x_{\beta}}{x_{\alpha} + 2x_{\beta}} = \frac{x_{\alpha} - x_{\beta}}{x_{\alpha} + 2x_{\beta}}
\]

\[
G(L) = \delta \left( L - \frac{1}{3} \frac{x_{\alpha}}{3} \right)
\]
The advantage of the spectral representation is that it distinguishes between the influence of geometrical quantities and that of the dielectric properties of the components on the effective behavior of the system. Although the spectral function $G(L)$ is generally unknown for an arbitrary two-phase composite, it’s analytically known or can be numerically derived for any existing mixture rule. Examples of spectral functions and corresponding solutions are shown in Table 2.7.

### 2.9 The Hall Effect

The Hall effect is closely related to the phenomenon of conductivity and is observed as the occurrence of a voltage appearing across a conductor carrying an electric current in a magnetic field. The schematic of the experiment is shown in the Fig. 2.14. The effect is often characterized by the Hall coefficient

$$R_H = \frac{E_y}{J_x B_z}, \quad (2.61)$$

where $E_y$ is the Hall effect electric field in the $y$-direction, $J_x$ is the current density in the $x$-direction and $B_z$ is magnetic field in the $z$-direction.

The Hall effect for ambipolar conduction in a sample where there are both negative and positive charge carriers, e.g., electrons and holes in a semiconductor, involves not only the concentrations of electrons and holes, $n$ and $p$ respectively, but also the electron and hole drift mobilities, $\mu_e$ and $\mu_h$. In the first approximation, the Hall coefficient can be calculated in the following way. Both charge carriers experience a Lorentz force in the same direction since they are drifting in the opposite directions but of course have opposite signs as illustrated in Fig. 2.14. Thus, both

![Fig. 2.14 Hall effect for ambipolar conduction.](image)
holes and electrons accumulate near the bottom surface. The magnitude of the Lorentz force, however, will be different since the drift mobilities and hence drift velocities will be different. Once equilibrium is reached, there should be no current flowing in the y-direction as we have an open circuit. The latter physical arguments lead to the following Hall coefficient [2.5],

$$R_H = \frac{p\mu_b^2 - n\mu_e^2}{e(p\mu_b + n\mu_e)^2}$$

or

$$R_H = \frac{p - nb^2}{e(p + nb)^2},$$

(2.62)

where \(b = \mu_e / \mu_b\).

It is clear that the Hall coefficient depends on both the drift mobility ratio and the concentrations of holes and electrons. For \(p > nb^2\), \(R_H\) will be positive and for \(p < nb^2\), it will be negative. Note that the carrier concentration is not zero when the Hall coefficient is zero but rather \(n/p = (\mu_b/\mu_e)^2\). As an example, Fig. 2.15 shows the dependence of Hall coefficient versus electron concentration for a single crystal silicon. The calculations are based on (2.62) and the law of mass action

$$np = n_t^2,$$

(2.63)

where \(n_t\) is the electron concentration in the intrinsic semiconductor.

In the case of monopolar conduction, e.g., conduction in metals or in doped semiconductors, (2.62) reduces to

$$R_H = -\frac{1}{en}, \quad \text{(for} n \gg p \text{)}$$

(2.64)

or

$$R_H = \frac{1}{ep}, \quad \text{(for} p \gg n \text{)}$$

(2.65)

Therefore, (considering as an example a n-type semiconductor where \(\sigma = ne\mu_e\)) one can write

$$\mu_H = \frac{\sigma}{ne} = -\sigma R_H$$

(2.66)

which provides a simple expression for determining the electron mobility known as the Hall mobility. Note, however, that the Hall mobility may differ from the drift mobility discussed in the previous sections. The difference arises from the carriers in a semiconductor having

![Fig. 2.15 Normalized Hall coefficient versus normalized electron concentration plot for single-crystal silicon. The values 0.17, 1.14 and 0.33 shown are the \(n/n_t\) values when the magnitude \(R\) reaches maxima and zero respectively. In single-crystal silicon, \(n_t = 1.5 \times 10^{10} \text{ cm}^{-3}, \mu_e = 1350 \text{ cm}^2 \text{V}^{-1}\text{s}^{-1}\) and \(\mu_b = 450 \text{ cm}^2 \text{V}^{-1}\text{s}^{-1}\) a distribution of energies. An average is used to describe the effect of carriers occupying different allowed states. (This is distinct from the earlier discussions of mobility where it was assumed that all the carriers have the same mean free time between collisions.) To include this, it is necessary to use a formal analysis based on the Boltzmann transport equation, as discussed in Sect. 2.6. If we express the averaging of the carriers with energy \(E\) and distribution function \(f(E)\), we can write the energy-averaged \(\tau(E)\) (i.e., \(\langle \tau \rangle \)) as

$$\langle \tau \rangle = \frac{\int \tau(E)f(E)dE}{\int f(E)dE}$$

(2.67)

and the energy-averaged \(\tau^2\) (i.e., \(\langle \tau^2 \rangle \)) as

$$\langle \tau^2 \rangle = \frac{\int \tau^2(E)f(E)dE}{\int f(E)dE}.$$ 

(2.68)

The rigorous analysis [2.12] shows that the Hall mobility \(\mu_H\) in terms of the drift mobility \(\mu_d\) is

$$\mu_H = \mu_d \times \mu_d,$$

(2.69)

where \(\mu_d\) is the Hall factor given by the ratio

$$\mu_H = \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2}.$$ 

(2.70)

The magnitude of the Hall factor will depend on the magnitude of the scattering mechanisms that contribute to \(\tau\), but usually at low magnetic fields, \(r_H \approx 1\) and the two mobilities are identical.
2.10 High Electric Field Transport

The previous sections focused on carrier transport in weak electric fields where the energy gained by carriers from the field is lost to the lattice through collisions with phonons or ionized impurities. However, at higher fields, the efficiency of collision mechanisms diminishes and the carrier system contains more energy than the lattice. The carriers are then called hot with effective temperatures $T_e$ for electrons and $T_h$ for holes. In this case, the drift velocity no longer obeys Ohm’s law, and becomes non-linear in the applied field with a clear tendency to saturation due to the appearance of a new dissipation mechanism involving optical phonon generation. Figure 2.16 shows the drift velocity saturation for both electrons and holes in Si and electrons in GaAs; GaAs shows a region of electron velocity overshoot and then negative differential resistivity due to inter-valley scattering, i.e., the transfer of electrons from the $NUL$ minimum to the $L$ conduction band minimum.

Solving the Boltzmann transport equation by analogy with (2.6) and (2.7) the electron (or hole) mobility may be defined as

$$
\mu = \frac{e \{ \tau(T_e) \}}{m^*},
$$

where $e$ is elementary charge and $m^*$ is the electron (or hole) effective mass and $\{ \tau(T_e) \}$ is the mean free time which now strongly depends on $T_e$. Therefore, the high-field mobility is related to the low field mobility $\mu_0$ by

$$
\mu = \mu_0 \left( \frac{T_e}{T} \right)^\beta,
$$

where $\beta$ depends on electric field and scattering mechanisms. Thus, in order to determine the dependence of the mobility on electric field, the dependence of the effective carrier temperature on field is required. This may be found by using the time-dependent Boltzmann transport equation. Suppose that $F$ is the field, then for non-degenerate conditions with $T_e \gg T$

$$
\frac{T_e}{T} \propto F^{-\frac{1}{2\beta}}
$$

and hence

$$
\mu(E) \propto \mu_0 F^{-\frac{1}{2\beta}}.
$$

For acoustic phonon scattering, $\beta = -3/2$ and the drift mobility shows no saturation, increasing with field as $F^{1/4}$. Saturation in the drift velocity may be achieved only when $\beta \to \infty$ due to optical phonon scattering where large energy changes are involved. The saturation velocity $v_s$ (related to the saturation mobility by $v_s = F \mu_s$) may be calculated using the energy and momentum rate equations for optical phonon scattering

$$
\frac{d \langle E \rangle}{dt} = eFv_s - E_{op} \tau_0
$$

and

$$
\frac{d \langle m^* v_s \rangle}{dt} = eF - m^* v_s \tau_m
$$

where $E_{op}$ is the optical phonon energy, $\tau_0$ and $\tau_m$ are the energy and momentum relaxation times, respectively. At steady state and for not extremely high fields, one may assume that $\tau_e \approx \tau_m$. (It is worth noting that at the highest electric fields $\tau_e > \tau_m$ and may lead to the appearance of avalanche, as discussed in Sect. 2.11). Therefore, the solution of (2.75) and (2.76) is

$$
v_s = \left( \frac{E_{op}}{m^*} \right)^{1/2}.
$$

in agreement with the experimental values shown in Fig. 2.16.
2.11 Impact Ionization

At very high electric field (in the range 2 × 10^4 V/cm or larger) a new possibility appears: a carrier may have kinetic energy in excess of the binding energy of a valence electron to its parent atom. In colliding with an atom, such a carrier can rupture the covalent bond and produce an electron-hole pair. This process is called impact ionization and is characterized by the impact ionization coefficient α for electrons and α_h for holes. The released electrons and holes may, in turn, impact ionize more atoms producing new electron-hole pairs. This process is called avalanche and is characterized by the multiplication factor, M, which is the ratio of number of collected carriers to the number of initially injected carriers.

The field dependence of the impact ionization coefficient, at least over the limited fields where avalanche is observed, is usually modeled by experimentalists by using

$$\alpha = A \exp \left[ -\left( \frac{B}{F} \right)^n \right], \quad (2.78)$$

where $F$ is the field $A$, $B$, $n$ are constants that depend on the semiconductor material properties such as the $E$-$k$ electronic structure, phonon energies and spectra, scattering mechanisms and so on. The constant $B$ has been semi-quantitatively argued to depend on $E_k/\tau$ where $E_k$ is the bandgap and $1/\tau$ is the phonon scattering rate; higher bandgap semiconductors tend to have steeper slopes on log $\alpha$ versus $1/F$ plots and the log $\alpha$ versus $1/F$ curve tends to shift to higher fields. Typically, it has been found that $n \approx 1$ at low fields and $n \approx 2$ for high fields. Figure 2.17 shows experimental data for a variety of materials over a wide range of electric fields.

Avalanche multiplication in crystalline semiconductors has been well studied over several decades and is now well developed and used in various commercial applications, such as avalanche photodiodes in optical communications, high sensitivity radiation detectors, and x-ray and nuclear detectors [2.69, 70]. In the case of amorphous semiconductors, the present work points to only a-Se exhibiting avalanche multiplication in the bulk for reasons explained in reference [2.71], with reports dating back to the 1980s [2.61, 62, 72]. Avalanche in a-Se has also been commercialized in ultrasensitive video tubes [2.73], with potential application in medical imaging [2.74].

The origin of (2.78) in its simple $n = 1$ form lies in Shockley’s [2.75] lucky electron model. When a carrier moves a distance $\zeta$ down-stream (along the field) without being scattered, it gains an energy $eF\zeta$. An unlucky carrier is scattered so frequently that its $eF\zeta$ never reaches the threshold ionization energy $E_I$ for impact ionization. On the other hand, a lucky electron is a ballistic electron that avoids scattering for a substantial distance, and hence is able to build its $eF\zeta$ to reach $E_I$ and thereby cause impact ionization. If $\lambda$ is the mean free path of collisions, then Shockley’s model gives

$$\alpha = \frac{eF}{E_I} \exp \left( -\frac{E_I}{eF\lambda} \right). \quad (2.79)$$

The main problem with the Shockley model is that there are just not enough ballistic electrons to cause sufficient impact ionizations to explain the experiments. A better model was developed by Baraff [2.76] who numerically solved the Boltzmann transport equation for a simple parabolic band and energy independent mean free path to provide a relationship between $\alpha$ and $F$ in terms of four parameters, that is, threshold energy for impact ionization, mean free path associated with ionizations, optical phonon energy, and mean free path for optical phonon scattering. Baraff’s theory served experimental-
exists quite well in terms of comparing their results even though the model was not intuitive and was limited in terms of its assumptions and applicability to real semiconductors. Impact ionization theory in crystalline solids only reached an acceptable level of confidence and understanding in the 1980s and 1990s with the development of the lucky-drift model by Ridley [2.77] and its extension by Burt [2.78], and Mackenzie and Burt [2.79]. The latter major advancement in the theory appeared as the lucky drift (LD) model, and it was based on the realization that at high fields, hot electrons do not relax momentum and energy at the same rates. Momentum relaxation rate is much faster than the energy relaxation rate. An electron can drift, being scattered by phonons, and have its momentum relaxed, which controls the drift velocity, but it can still gain energy during this drift. Stated differently, the mean free path \( \lambda_E \) for energy relaxation is much longer than the mean free path \( \lambda \) for momentum relaxation.

In the [2.79] version of the Mackenzie and Burt LD model, the probability \( P(E) \) that a carrier attains an energy \( E \) is given by

\[
P(E) = \exp \left( - \int_0^E \frac{dE'}{eF\lambda(E')} \right) \int_0^E \frac{dE_1}{eF\lambda(E_1)} \times \exp \left( - \int_0^{E_1} \frac{dE'}{eF\lambda(E')} \right) \times \exp \left( - \int_{E_1}^E \frac{dE'}{eF\lambda(E')} \right),
\]

(2.80)

where as mentioned above \( \lambda \) is the mean free path associated with momentum relaxing collisions and \( \lambda_E \) is the mean energy relaxation length associated with the energy relaxing collisions. The first term is the Shockley lucky electron probability, i.e., the electron moves ballistically to energy \( E \). The second term is the lucky drift probability term which is composed of the following: the electron first moves ballistically to some intermediate energy \( E_1 (0 < E_1 < E) \) from where it begins its lucky drift to energy \( E_1 \); hence the integration over all possible \( E_1 \). The impact ionization coefficient can then readily be evaluated from

\[
\alpha = \frac{eF\lambda}{\int_0^{E_1} P(E)dE}.
\]

(2.81)

The model above is based on a hard threshold ionization energy \( E_1 \), that is, when a carrier attains the energy \( E_1 \), ionization ensues. The model has been further refined by the inclusion of soft threshold energies which represent the fact the ionization does not occur immediately when the carrier attains the energy \( E_1 \), and the carrier drifts further to gain more energy than \( E_1 \) before impact ionization [2.80–84].

Assuming \( \lambda \) and \( \lambda_E \) are energy independent, which would be the case for a single parabolic band in the crystalline state, (2.80) and (2.81) can be solved analytically to obtain

\[
\alpha = \frac{1}{\lambda} \times \frac{\lambda_E}{\lambda} \exp \left( \frac{-E_1}{eF\lambda} \right) + \left( \frac{\lambda_E}{\lambda} \right)^2 \exp \left( \frac{-E_1}{eF\lambda} \right)
\]

(2.82)

For \( \lambda_E > \lambda \), and in the low field region, where typically \( (\alpha \lambda) < 10^{-1} \), or \( x = E_1/eF\lambda > 10 \), (2.82) leads to a simple expression for \( \alpha \),

\[
\alpha = \left( \frac{1}{\lambda_E} \right) \exp \left( - \frac{E_1}{eF\lambda_E} \right).
\]

(2.83)

For crystalline semiconductors, one typically also assumes that \( \lambda_E \) depends on the field \( F \), \( \lambda \) and the optical phonon energy \( \hbar \omega \) as

\[
\lambda_E = \frac{eF\lambda^2}{2\hbar \omega} \coth \left( \frac{\hbar \omega}{2kT} \right).
\]

(2.84)

As the field increases, \( \lambda_E \) eventually exceeds \( \lambda \), and allows lucky drift to operate and the LD carriers to reach the ionization energy.

It is worth noting that the model of lucky drift is successfully used not only for crystalline semiconductors but to amorphous semiconductors [2.83].
2.12 Two-Dimensional Electron Gas

Heterostructures offer the ability to spatially engineer the potential in which carriers move. In such structures having layers deposited in the $z$-direction, when the width of a region with confining potential $t_c < \lambda_{\text{DB}}$, the de Broglie electron wavelength, electron states become stationary states in that direction, retaining Bloch wave character in the other two directions (i.e., $x$- and $y$-directions), and is hence termed a 2-D electron gas (2DEG). These structures are notable for their extremely high carrier mobility.

High mobility structures are formed by selectively doping the wide bandgap material behind an initially undoped spacer region of width $d$ as shown in Fig. 2.18a. Ionization and charge transfer leads to carrier build-up in the low potential region of narrow bandgap material adjacent to the heterointerface. The equilibrium band bending (i.e., through Poisson’s equation) in the well region, as shown in Fig. 2.18b, results from the equalization of ionized donor concentration in the wide bandgap material and 2DEG concentration adjacent to the heterointerface. When the associated interfacial field is sufficiently strong, carriers are confined within $\lambda_{\text{DB}}$ and electron states are quantized into sub-bands (i.e., $E_0$ and $E_1$), as shown in Fig. 2.18b.

Figure 2.19 shows the contributions of component scattering mechanisms to the low temperature mobility of a 2DEG formed at a Ga$_{0.7}$Al$_{0.3}$As-GaAs heterointerface, as a function of the electron gas density. As for bulk samples, the most important mechanism limiting the low temperature mobility is ionized impurity
scattering, except at high electron densities, where so-called *alloy disorder scattering* is significant. Ionized impurity scattering may be further broken down into scattering from ionized impurities that are within the GaAs quantum well, known as *background impurities*, those beyond the spacer region, termed *remote impurity scattering*. For high purity growth, the unintentional background impurity concentration can be kept to very low limits and impurity scattering based mobility values are then dictated by remote impurity scattering. Since carriers in the well are only weakly scattered by the tail field of these remote Coulomb centers, the mobility of such 2DEG systems can be orders of magnitude higher than bulk samples. The temperature dependence of the electron mobility for such a system is shown in Fig. 2.20. Notice how, similar to bulk sample, increasing temperature increases the phonon population such that for the example shown, above about 100 K, polar optical phonon scattering controls the mobility.

Figure 2.21 shows the dependence of the 2DEG mobility on the spacer width. Two competing factors are active – for narrow spacer widths, the transfer efficiency of carriers to the GaAs well is high and so a lower remote doping concentration is sufficient to provide for a given constant 2DEG concentration, but since the Coulomb scatters are so close to the 2DEG, they scatter very efficiently and limit the mobility. On the other hand, for large spacer widths, carrier transfer efficiency is quite poor requiring higher remote doping to supply the given 2DEG concentration; however, being relatively far away each of the scattering centers are less effective at lowering the mobility, but given their high concentration, the net effect is still a decrease in the mobility at large spacer widths as seen in the figure.
2.13 One-Dimensional Conductance

In the case where carriers are confined within regions of width \( L_z \), \( L_y < \lambda_{\text{dB}} \) in two directions \( x \) and \( y \), respectively, the electron energy in those two directions become quantized with quantum numbers \( n_x \) and \( n_y \), respectively. In the third direction \( z \), electrons travel as Bloch waves with energy that may be approximated by \( h^2 k^2 / 2m^* \) giving an expression for the total energy \( E \) of the so-called 1-D electron system or quantum wire as

\[
E(n_x, n_y, n_z) = \frac{h^2}{2m^*} \left( \frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{h^2 k_z^2}{2m^*} \right). \tag{2.85}
\]

The associated electron wavefunctions are

\[
\Psi(x, y, z) = \frac{1}{2\sqrt{L_x L_y L_z}} \sin \left( \frac{n_x \pi x}{L_x} \right) \times \sin \left( \frac{n_y \pi y}{L_y} \right) e^{ik_z z}. \tag{2.86}
\]

Using these equations, one can readily derive an expression for the density of states per unit energy range

\[
\text{DOS} = 2 \times 2 \left( \frac{L_x}{2\pi} \right) (\nabla_k E)^{-1} = \frac{2L_x}{h} \sqrt{\frac{m^*}{2(E - E_{n_x, n_y})}}. \tag{2.87}
\]

In order to evaluate the conductance of this quantum wire, consider the influence of a weak applied potential \( V \). Similar to the case for bulk transport the applied field displaces the Fermi surface and results in a change in the electron wave-vector from \( k_0 \) (i.e., with no applied potential) to \( k_f \) (i.e., when the potential is applied). When \( V \) is small compared with the electron energy

\[
k_0 = \sqrt{\frac{2m^* (E - E_{n_x, n_y})}{h^2}}, \quad k_f = k_0 \sqrt{1 + \frac{eV}{E - E_{n_x, n_y}}}
\approx k_0 \left( 1 + \frac{1}{2} \frac{eV}{E - E_{n_x, n_y}} \right). \tag{2.88}
\]

This leads to establishing a current density \( J \) in the wire

\[
J = \frac{2e^2 \text{DOS} \sqrt{(E_f - E_{n_x, n_y})}}{\sqrt{2m^*}}. \tag{2.90}
\]

Which may be simplified to the following expressions for \( J \) and the current flowing in the wire for a given quantum state \( E(n_x, n_y, n_z) \),

\[
J_{n_x, n_y} = \frac{2e^2 V_{x,y}}{h} \quad \text{and} \quad I_{n_x, n_y} = \frac{2e^2 V}{h}. \tag{2.91}
\]

The expression for the conductance through one channel corresponding to a given quantum state \( \{n_x, n_y\} \) is then given by

\[
G_{n_x, n_y} = \frac{I_{n_x, n_y}}{V} = \frac{2e^2}{h}. \tag{2.92}
\]

Notice how the conductance is quantized in units of \( e^2 / h \) with each populated channel contributing equally to the conductance – moreover, this is a fundamental result, being independent of the material considered. In practice, deviations from this equation can occur (although generally less than 1%) owing to the finiteness of real nanowires and impurities in or near the channel, influencing the conductivity and even resulting in weak localization. Generally, unlike both bulk and 2DEG systems, ionized impurity scattering is suppressed in nanowires. The main reason for this is that an incident electron in a quantum state \( \{n_x, n_y\} \) traveling along the wire with wave-vector \( k_f \{n_x, n_y\} \), can not be elastically scattered into any states except those in a small region of \( k \)-space in the vicinity of \( -k \{n_x, n_y\} \). Such a scattering event involves a large change in momentum of \( \approx 2k \{n_x, n_y\} \) and thus, the probability of such events is very small. As a result, the mean free path and mobility of carriers in such quantum wires are substantially increased.

The nature of carrier transport in quantum wires depends on the wire dimensions (i.e., length \( L_{\text{Wire}} \) and diameter \( d_{\text{Wire}} \)) as compared with the carrier mean free path, \( l_{\text{Carrier}} \). When \( l_{\text{Carrier}} \gg L_{\text{Wire}} \), \( d_{\text{Wire}} \) the only potential seen by the carriers is that associated with the wire walls, and carriers exhibit waveguide behavior, being guided through the wire as if it were a waveguide without any internal scattering. Conversely, if \( d_{\text{Wire}} \ll \lambda_{\text{DeBroglie}} \), only a few energy states in the wire are active, and in the limit of an extremely small waveguide, only one state or channel is active, analogous to a single mode waveguide cavity – this case is termed quantum ballistic transport. In the limit, \( l_{\text{Carrier}} \ll L_{\text{Wire}} \), \( d_{\text{Wire}} \), scattering dominates transport throughout the wire – with numerous scattering events occurring before a carrier can traverse the wire or move far along its length. In such a case the transport is said to be diffusive.
As discussed previously, ionized impurity and lattice scattering contribute to \( \rho_{\text{Car}} \), with \( \rho_{\text{Car}} \) decreasing with increasing temperature due to phonon scattering. For strong impurity scattering, this may not occur until relatively high temperatures. In the intermediate case of \( L_{\text{Wir}} \gg l_{\text{Car}} \gg d_{\text{Wir}} \) and where \( d_{\text{Wir}} \ll \lambda_{\text{DeBroglie}} \), scattering is termed mixed mode and is often called quasi-ballistic.

### 2.14 The Quantum Hall Effect

The observation of, and first explanation for the Hall Effect in a 2DEG by von Klitzing et al. [2.85], won them a Nobel Prize. As shown in Fig. 2.22 the Hall resistivity exhibits plateaus for integer values of \( h/e^2 \), independent of any material dependent parameters. This discovery was later shown to be correct to a precision of at least one part in \( 10^7 \), enabling extremely accurate determinations of the fine structure constant (i.e., \( \alpha = (\mu_0 c e^2)/(2\hbar) \approx 1/137 \) and a fundamental resistance standard to be established.

For the six point Hall geometry used (as shown in the insert to Fig. 2.18), one can define the Hall resistivity \( \rho_{xy} = g_1 V_{H}/I_x \) and longitudinal resistivity \( \rho_{xx} = g_2 V_{H}/I_x \) where \( g_1 \) and \( g_2 \) are geometric constants related to the sample geometry. These resistivities are related to corresponding conductivities through the conductivity and resistivity tensors

\[
\begin{align*}
\rho_{xy} &= \frac{\sigma_{xy}}{\sigma_{xx} + \sigma_{yy}}, \\
\rho_{xx} &= \frac{\sigma_{xx}}{\sigma_{xx} + \sigma_{yy}}, \\
\rho_{yy} &= -\rho_{xy}.
\end{align*}
\]

Starting from a classical equation of motion for electrons in an electric field \( E_{\text{c}} \), magnetic field \( B_{\text{c}} \), and defining the cyclotron frequency \( \omega_{\text{c}} = eB_{\text{c}}/m^* \), the velocities perpendicular to the applied magnetic field can be deduced as

\[
\begin{align*}
\langle v_x \rangle &= \frac{E_x}{B_z} \sin \omega_{\text{c}} t \quad \text{and} \quad \langle v_y \rangle = \frac{E_x}{B_z} (\cos \omega_{\text{c}} t - 1)
\end{align*}
\]

with the time averaged velocities

\[
\langle v_x \rangle = \frac{E_x}{B_z} \quad \text{and} \quad \langle v_y \rangle = \frac{E_x}{B_z}
\]

leading to Hall and longitudinal resistivities of

\[
\sigma_{xy} = \sigma_{yx} = 0 \quad \text{and} \quad \sigma_{xx} = -\sigma_{yy} = \frac{n_e e}{B_z}.
\]

where \( n_e \) is the areal electron concentration. Below, a quantum approach is used to establish a relationship for the electron concentration \( n_e \). Note that the motion of the electrons in the crossed fields are quantized with allowed levels termed Landau levels \( E_n \)

\[
E_n = \left(n + \frac{1}{2}\right) \hbar \omega_{\text{c}} + g^* \mu_B B_{\text{c}} + \epsilon_z.
\]

where \( n \) is the quantum number describing the particular Landau Level, \( g^* \) is the Landé factor, \( \mu_B \) is the Bohr magneton and \( g^* \mu_B B_{\text{c}} \) is the spin magnetic energy, and \( \epsilon_z \) is the energy associated with the \( z \)-motion.
of the carriers, \(xy\)-plane carrier motion is characterized by the cyclotron energy term \(E_{cy}\),

\[
E_{cy} = \frac{\hbar^2 k_x^2}{2m^*} = \left( n + \frac{1}{2} \right) \hbar \omega_c .
\]  

(2.98)

Following this description and noting that the motion of electrons in the \(xy\)-plane may be expressed in terms of wavefunctions of the harmonic oscillator using the Landau gauge of the vector potential \((0, xB_z, 0)\), we may write the density of states per unit area, \(DOS_\Lambda\) as

\[
DOS_\Lambda = \frac{m^* \omega_L L_y}{2\pi \hbar} .
\]  

(2.99)

Since the degeneracy of each Landau level is one (i.e., since they are single spin states), this enables one to find \(n_s\) assuming Landau state filling up to the \(p^\text{th}\) level (for integer \(p\))

\[
n_s = \frac{m^* \omega_L p}{2\pi \hbar} .
\]  

(2.100)

Using the definition of the cyclotron frequency gives the final form

\[
n_s = \frac{p e B_z}{\hbar} .
\]  

(2.101)

which may be used to rewrite the previous expressions for Hall and longitudinal resistivity

\[
\rho_{xx} = \rho_{yy} = 0 \quad \text{and} \quad \rho_{xy} = -\rho_{yx} = \frac{\hbar}{e^2} .
\]  

(2.102)

Equation (2.102) shows that Hall resistivity is quantized in units of \(\hbar/e^2\) whenever the Fermi energy lies between filled Landau levels. Consistent with observation, the result is independent of the semiconductor being studied. Although this model provides an excellent basis for understanding experiments, understanding the details of the results (i.e., in particular the existence of a finite width for the Hall effect plateaus and zero longitudinal resistance dips) requires a more complete treatment involving so-called localized states.

**Acknowledgments.** The authors thank Dr. Robert Johnson for helpful discussions on the subject.

**References**

7. Kaye & Laby Tables of Physical and Chemical Constants (Section 2.6) (National Physical Laboratory, Teddington 2016) provided by the website of the National Physical Laboratory
2.37 R. Clausius: "Die Mechanische Wärmetheorie", Vol. 2 (Vieweg, Braunschweig 1879)
2.38 L. Lorenz: Ann. Phys. Lpz. 11, 70 (1880)
2.40 Lord Rayleigh: Philos. Mag. 34, 481 (1892)
2.41 K.W. Wagner: Arch. Electrochem. 2, 371 (1914)
2.44 H. Fricke: Phys. Rev. 24, 575 (1924)
2.45 J.A. Stratton: "Electromagnetic Theory" (McGraw-Hill, New York 1941)
2.46 J.M. Maxell-Garnett: Phil. Trans. R. Soc. Lond. 203, 385 (1904)
2.47 H. Looyenga: Physica 31, 401 (1965)
2.51 F. Ollendorf: Arch. Electrochem. 25, 416 (1931)
2.52 J.C.M. Maxwell-Garnett: Phil. Trans. R. Soc. Lond. 203, 385 (1904)
2.53 H. Fricke: Ann. Phys. Lpz. 11, 70 (1880)
Springer Handbook of Electronic and Photonic Materials
Kasap, S.; Capper, P. (Eds.)
2017, XXXVI, 1536 p. 1088 illus. in color., Hardcover
ISBN: 978-3-319-48931-5