

Chapter 8

DC conductivities

8.1. ELECTRON SCATTERING BY IMPURITIES

This chapter is mostly concerned with the methods for calculating the electrical conductivity. Four different methods are discussed: (1) solving the Boltzmann equation, (2) evaluating the Kubo formula for the current–current correlation function, (3) evaluating the force–force correlation function, and (4) solving the quantum Boltzmann equation. For scattering from fixed impurities they all give the same answer. For scattering by phonons two different answers are obtained. One is called the Ziman (1960) formula, and the other the Holstein (1964) formula. Two criteria are important in comparing these methods: which is the easiest to use, and which gives the most accurate answer?

The electrical resistivity, or conductivity, from impurity scattering is an important topic. From the experimental viewpoint, all solids have impurities which make a contribution to the total resistivity. In many metals and semiconductors, the low-temperature resistivity is dominated by impurities, since all other contributions are temperature dependent and vanish at low temperature. In metals, the resistivity from impurity scattering is largely temperature independent, except for temperature variations on the scale of the Fermi temperature $T_F = E_F/k_B$. The subject is important, on the theoretical side, because it was one of the earliest evaluations of the Kubo formula. The importance of vertex corrections became apparent. Indeed, the derivation showed that vertex corrections are usually very important and should be assumed important until shown otherwise. This conclusion, and message, continues to be relevant even for calculations of other quantities.

The present chapter is really about vertex corrections.

8.1.1. Boltzmann equation

The electrical resistivity from impurity scattering is easily derived by using the Boltzmann equation. This derivation is presented for several reasons. First, it is the easiest way to get the answer. Second, the resistivity was first found this way; the Green's function evaluation of the Kubo formula only confirmed the result known earlier from transport theory. The Green's function derivation is complicated and subtle, and it is useful to know and believe the right answer in order to recognize it when it is finally derived.

Our objective is to derive a formula for the electrical resistivity with the least possible fuss. The simplest possible model is adopted for the solid. It is a homogeneous system except for randomly located impurities. The electron states are plane waves except for occasional scattering from isolated impurities. The impurities are very dilute, so that interference between successive scatterings can be neglected. In the Boltzmann theory, the electrons are described by a classical distribution function $f(\mathbf{r}, \mathbf{k}, t)$. The time rate of change of this distribution function is governed by the Boltzmann equation

$$0 = \frac{df}{dt} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f + \frac{\partial \mathbf{k}}{\partial t} \cdot \nabla_{\mathbf{k}} f + \left(\frac{df}{dt} \right)_{\text{collisions}} \quad (8.1)$$

The last term is the time rate of change due to collisions with the impurities. There is no \mathbf{r} dependence in $f(\mathbf{r}, \mathbf{k}, t)$ since the material is assumed to be homogeneous. Also, for the dc conductivity, there is no time dependence. The system has a weak external electric field and the current flows in a steady-state fashion. The distribution function is only a function of wave vector $f(\mathbf{k})$ and obeys the equation

$$0 = \frac{\partial \mathbf{k}}{\partial t} \cdot \nabla_{\mathbf{k}} f + \left(\frac{df}{dt} \right)_{\text{collisions}} \quad (8.2)$$

In a solid, the factor $\partial \mathbf{k} / \partial t$ is equivalent to an acceleration which is equal to the forces on the electron (see Kittel, 1963):

$$\frac{\partial \mathbf{k}}{\partial t} = -e\mathbf{E} - \frac{e}{c} \mathbf{v} \times \mathbf{H}_0 \quad (8.3)$$

In the present problem, there is an electric field \mathbf{E} and no magnetic field $\mathbf{H}_0 = 0$, so that

$$e\mathbf{E} \cdot \nabla_{\mathbf{k}} f(\mathbf{k}) = \left(\frac{df}{dt} \right)_{\text{collisions}} \quad (8.4)$$

The collision term is the most interesting. It is evaluated in the *relaxation time approximation*. This approximation assumes that collisions seek to return the system to the equilibrium configuration $f_0(\mathbf{k})$, which is the configuration the system would have in the absence of an electric field. The rate of change of $f(\mathbf{k})$ due to collisions is assumed to be proportional to the degree that $f(\mathbf{k})$ is different from $f_0(\mathbf{k})$:

$$\left(\frac{df}{dt} \right)_{\text{collisions}} = - \frac{[f(\mathbf{k}) - f_0(\mathbf{k})]}{\tau_t(\mathbf{k})} \quad (8.5)$$

$$f_0(\mathbf{k}) = \frac{2}{e^{\beta \epsilon_{\mathbf{k}}} + 1} \quad (8.6)$$

The factor of two in f_0 is due to the spin degeneracy. The above equation defines the transport relaxation time $\tau_t(\mathbf{k})$. A more detailed derivation can be found in Ziman (1960). Here an equation for $\tau_t(\mathbf{k})$ is derived and solved. Then the distribution function is

$$f(\mathbf{k}) = f_0(\mathbf{k}) - e\tau_t(\mathbf{k})\mathbf{E} \cdot \nabla_{\mathbf{k}} f(\mathbf{k}) \quad (8.7)$$

When the electric field is small, only a small amount of current flows. The system is only slightly out of equilibrium. The distribution function $f(\mathbf{k}) = f_0(\mathbf{k}) + f_1(\mathbf{k})$, where the change $f_1(\mathbf{k})$ is small. It is only necessary to retain terms of $O(E)$. Iterate the above equation, which

effectively replaces f by f_0 on the right-hand side. Then $f(\mathbf{k})$ is evaluated to $O(E)$. Furthermore, it is assumed that the system is isotropic and $f_0(\mathbf{k}) = f_0(k)$

$$f(\mathbf{k}) \approx f_0(k) - e\tau_t(\mathbf{k})\mathbf{E} \cdot \nabla_{\mathbf{k}} f_0(k) \quad (8.8)$$

$$\approx f_0(k) - e\tau_t(\mathbf{k}) \frac{\mathbf{E} \cdot \mathbf{k}}{m} \frac{df_0(k)}{d\varepsilon_k} \quad (8.9)$$

The electrical current density \mathbf{J} is the product of the electron charge e , the electron's density n_0 , and the average velocity $\langle \mathbf{v} \rangle$, which is obtained by averaging over the electron distribution:

$$\mathbf{J} = en_0 \langle \mathbf{v} \rangle = e \int \frac{d^3k}{(2\pi)^3} f(\mathbf{k}) \frac{\hbar \mathbf{k}}{m} \quad (8.10)$$

$$n_0 = \int \frac{d^3k}{(2\pi)^3} f(\mathbf{k}) \quad (8.11)$$

The distribution function $f(\mathbf{k})$ is normalized to give the electron density n_0 . By using the result (8.9) for $f(\mathbf{k})$, the term f_0 gives an average $\langle \mathbf{v} \rangle$ of zero, since no current flows when there is no electric field; as many electrons are going one way as another. The current is proportional to the second term, and it is proportional to the electric field:

$$\mathbf{J} = e^2 \int \frac{d^3k}{(2\pi)^3} \tau_t(k) \mathbf{v}_k (\mathbf{v}_k \cdot \mathbf{E}) \left(-\frac{df_0(k)}{d\varepsilon_k} \right) \quad (8.12)$$

In a homogeneous, isotropic system, the current \mathbf{J} flows in the direction of \mathbf{E} . The quantity $f_0(k)$ is independent of \mathbf{k} direction. The only angular factors are $\mathbf{v}_k (\mathbf{v}_k \cdot \mathbf{E})$. The angular integrals will average this to $v_k^2 \mathbf{E}/3$ in three dimensions. The conductivity σ is the ratio of \mathbf{J} to \mathbf{E} :

$$\sigma = \frac{e^2}{3} \int \frac{d^3k}{(2\pi)^3} \tau_t(k) v_k^2 \left(-\frac{df_0(k)}{d\varepsilon_k} \right) \quad (8.13)$$

It is a positive quantity since $df_0/d\varepsilon_k$ is always negative. Equation (8.13) is the basis of all the calculations. There remains the important task of deriving a formula for the relaxation time $\tau_t(k)$. It is *not* just the time between scattering events, which is derived from the imaginary part of the retarded self-energy. This distinction is important, since it makes life difficult. The relaxation time in the Boltzmann equation is a special quantity.

The impurities are assumed to be static, fixed, objects with a spherically symmetric potential. They have no internal excitations, so the electron scatters from them elastically. The impurity causes the electron in state \mathbf{k} to scatter to \mathbf{k}' , which has the same energy, so that $|\mathbf{k}| = |\mathbf{k}'|$ and $\varepsilon_k = \varepsilon_{k'}$. The net rate of scattering out of the state \mathbf{k} is the rate of going from \mathbf{k} to \mathbf{k}' , which is proportional to $f(\mathbf{k})[1 - f(\mathbf{k}')] - f(\mathbf{k}')f(\mathbf{k})$ minus the rate from \mathbf{k}' to \mathbf{k} , which is proportional to $f(\mathbf{k}')[1 - f(\mathbf{k})]$:

$$\begin{aligned} -\left(\frac{df}{dt}\right)_{\text{collisions}} &= \frac{[f(\mathbf{k}) - f_0(\mathbf{k})]}{\tau_t(\mathbf{k})} = 2\pi n_i \int \frac{d^3k'}{(2\pi)^3} \delta(\varepsilon_k - \varepsilon_{k'}) \\ &\times \{ |T_{\mathbf{k}\mathbf{k}'}|^2 f(\mathbf{k})[1 - f(\mathbf{k}')] - |T_{\mathbf{k}'\mathbf{k}}|^2 f(\mathbf{k}') [1 - f(\mathbf{k})] \} \end{aligned} \quad (8.14)$$

where n_i is the concentration of impurities. The quantity $T_{\mathbf{k}\mathbf{k}'}$, is the T -matrix element for scattering from \mathbf{k} to \mathbf{k}' which was defined in Sec. 4.1. It is symmetric in its indices $T_{\mathbf{k}\mathbf{k}'} = T_{\mathbf{k}'\mathbf{k}}$, which simplifies the above equation

$$\frac{[f(\mathbf{k}) - f_0(\mathbf{k})]}{\tau_t(\mathbf{k})} = 2\pi n_i \int \frac{d^3 k'}{(2\pi)^3} \delta(\varepsilon_k - \varepsilon_{k'}) |T_{\mathbf{k}\mathbf{k}'}|^2 [f(\mathbf{k}) - f(\mathbf{k}')] \quad (8.15)$$

The integrand contains the factor $f(\mathbf{k}) - f(\mathbf{k}')$. The integral is evaluated by assuming the form in (8.9), which is written as

$$f(\mathbf{k}) = f_0(k) + \mathbf{k} \cdot \mathbf{E} C(k) \quad (8.16)$$

$$f(\mathbf{k}') = f_0(k') + \mathbf{k}' \cdot \mathbf{E} C(k') \quad (8.17)$$

$$f(\mathbf{k}) - f(\mathbf{k}') = (\mathbf{k} - \mathbf{k}') \cdot \mathbf{E} C(k) \quad (8.18)$$

Since $|\mathbf{k}| = |\mathbf{k}'|$ then $f_0(k') = f_0(k)$. The quantities $f(\mathbf{k})$ and $f(\mathbf{k}')$ differ only in the angular part of the second term. The angular part is treated as follows: define a coordinate system in which the \hat{z} direction is \hat{k} , so that

$$\hat{k} \cdot \hat{E} = \cos \theta \quad (8.19)$$

$$\hat{k} \cdot \hat{k}' = \cos \theta' \quad (8.20)$$

$$\hat{k}' \cdot \hat{E} = \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos \phi \quad (8.21)$$

where the law of cosines is used in the last identity. The difference of the two distribution functions is now

$$f(\mathbf{k}) - f(\mathbf{k}') = k E C(k) [\cos \theta (1 - \cos \theta') - \sin \theta \sin \theta' \cos \phi] \quad (8.22)$$

The last term on the right, which contains the factor $\cos \phi$, will vanish when doing the integral $\int d\phi$. There is no other ϕ dependence in the integrand of (8.15), and the average of $\cos \phi$ is zero. The remaining term may be written as

$$\int d\omega_{k'} [f(\mathbf{k}) - f(\mathbf{k}')] = k E C(k) \cos \theta \int d\omega_{k'} (1 - \cos \theta') \quad (8.23)$$

$$= [f(\mathbf{k}) - f_0(k)] \int d\omega_{k'} (1 - \cos \theta') \quad (8.24)$$

The term $f(\mathbf{k}) - f_0(k)$ is factored from both sides of (8.15), which leaves the definition for the reciprocal of the relaxation time:

$$\frac{1}{\tau_t(k)} = 2\pi n_i \int \frac{d^3 k'}{(2\pi)^3} \delta(\varepsilon_k - \varepsilon_{k'}) |T_{\mathbf{k}\mathbf{k}'}|^2 (1 - \cos \theta') \quad (8.25)$$

The important factor in the integrand is $(1 - \cos \theta') = 1 - \mathbf{k} \cdot \mathbf{k}'/k^2$. It makes the relaxation time in the Boltzmann equation different from the usual relaxation time $\tau(k)$, which is the time between scattering events. The latter quantity is simply

$$\frac{1}{\tau(k)} = 2\pi n_i \int \frac{d^3 k'}{(2\pi)^3} \delta(\varepsilon_k - \varepsilon_{k'}) |T_{\mathbf{k}\mathbf{k}'}|^2 \quad (8.26)$$

$$= \frac{4\pi n_i}{mk} \sum_l (2l+1) \sin^2 \delta_l(k) = v_k n_i \sigma(k) \quad (8.27)$$

$$\sigma(k) = \frac{4\pi}{k^2} \sum_l (2l+1) \sin^2 \delta_l(k) \quad (8.28)$$

The usual inverse relaxation time is just the scattering cross section $\sigma(k)$ times the particle velocity v_k times the impurity density n_i . An equivalent result may be obtained for the relaxation time $\tau_l(k)$ of the Boltzmann equation. The T matrix is expanded as in (4.47) and (4.52) for the case where $|\mathbf{k}| = |\mathbf{k}'|$,

$$T_{\mathbf{k}\mathbf{k}'} = -\frac{2\pi}{mk} \sum_l (2l+1) P_l(\cos \theta) e^{i\delta_l(k)} \sin(\delta_l(k)) \quad (8.29)$$

$$\frac{1}{\tau_l(k)} = \frac{4\pi n_i}{mk} \sum_{l=1} l \sin^2[\delta_l(k) - \delta_{l-1}(k)] \quad (8.30)$$

where θ is the angle between \hat{k} and \hat{k}' . The angular integrals in (8.25) are straightforward but cumbersome, so only the result for τ_l is just listed. The formula for $1/\tau_l$ has a different combination of phase shifts than the formula for the cross section.

The factor $(1 - \cos \theta')$ weights the amount of scattering of the electron by the impurity. Small-angle scattering, where $\cos \theta' \approx 1$, is relatively unimportant in contributing to $1/\tau_l$. These events do little to impede the flow of electrons and so contribute little to the resistivity. The factor $(1 - \cos \theta')$ obviously favors large-angle scattering events, which are more important for the electrical resistivity. The relaxation time in the transport equation is not identical to the average scattering rate because there is an additional factor to weight the amount of scattering.

An example of evaluating the conductivity in (8.13) is given for a free-electron metal. At low temperature the distribution function (8.6) becomes a delta function in energy which sets $k = k_F$

$$\left(-\frac{df_0}{d\xi_k} \right) = 2\delta(\xi_k) \quad (8.31)$$

The angular integrals have already been done, so that

$$\sigma = \frac{2e^2}{3} \frac{4\pi}{(2\pi)^3} v_F^2 m k_F \tau_l(k_F) = \frac{e^2 n_0 \tau_l(k_F)}{m} \quad (8.32)$$

where $n_0 = k_F^3/3\pi^2$. The relevant relaxation time is for electrons at the Fermi surface. Yet the conductivity is proportional to the density n_0 of all conduction electrons and not just to those at the Fermi surface. This surprising result is quite reasonable once the physics is understood. When the electric field is first imposed, the equation $\mathbf{k} = -e\mathbf{E}$ shows that all electrons in the Fermi sea start accelerating equally. The Fermi sea is translationally shifted in wave vector space. The scattering tends to relax the Fermi distribution back to its undisturbed configuration. As shown in Fig. 8.1, electrons in the leading edge of the displaced Fermi distribution

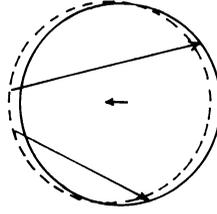


FIGURE 8.1 The circle represents the Fermi sea, which begins to move in response to an applied electric field. Steady state is maintained by relaxation back to other points on the Fermi surface.

are scattered back to the rear regions. Only those electrons at the Fermi surface can scatter. The electrons well below the Fermi surface cannot elastically scatter, since all states with the same energy are already occupied with other electrons. Above the Fermi surface there are no thermally excited electrons. Only electrons at the Fermi surface are available to elastically scatter to other points on the Fermi surface. The lifetime only involves $\tau_i(k_F)$.

The conductivity is relatively insensitive to temperature as long as the density of states of the metal is a smooth function of energy near the Fermi surface. The resistivity $\rho = 1/\sigma$ is the inverse of the conductivity, and is proportional to the concentration of impurities. This proportionality is experimentally verified.

Since impurity scattering is elastic, it does not change the energy of the electron. As the current of electrons moves through the solid, each electron gains energy from the electric field. How does the electron lose this energy, if it only scatters elastically? The next few paragraphs will answer this question.

Let $v = \cos \theta$, where θ is the angle between \mathbf{k} and \mathbf{E} . In order to keep the discussion simple, it is assumed that the distribution $f(\mathbf{k})$ is isotropic in the absence of a field. When the field is present, the distribution function can be expanded in a Legendre series in v . The first few terms are

$$f(\mathbf{k}) = f_s + v f_p(k) + P_2(v) f_d(k) + \dots \quad (8.33)$$

where $f_s(k)$ is the isotropic part of the distribution while $f_p(k)$ is the $l = 1$ part of the distribution. Note that $f_s(k)$ is not the equilibrium part of the distribution, which is $f_0(k)$. The electrical current is determined by the distribution f_p .

The conduction process can be viewed as having the following steps:

1. For $t < 0$, $\mathbf{E} = 0$ and the initial distribution is f_0 .
2. At $t = 0$ the field \mathbf{E} is switched on. It accelerates the particles and creates the distribution f_p .
3. The elastic scattering takes the particles from the p distribution f_p to the s distribution f_s . This step has a time constant τ_i .
4. The isotropic distribution f_s relaxes back to the equilibrium distribution f_0 . This step has a different time constant.

The energy relaxation occurs in the last step where f_s is brought to equilibrium. The electrons can lose energy to their heat bath, usually phonons. This process has a very different time constant than τ_i and is usually much slower. The relaxation time τ_i from elastic scattering determines the rate at which particles scatter out of the p distribution f_p into other distributions such as f_s and f_d . The current is determined by τ_i since it gives the steady state amplitude of f_p .

The energy relaxation occurs elsewhere in the chain of events. The rate for energy relaxation τ_e is calculated below.

8.1.2. Kubo formula: approximate solution

The electrical conductivity can be calculated from the Kubo formula by using the technique described in (3.390)–(3.393). The correlation function is evaluated for nonzero temperatures and frequencies:

$$\pi(i\omega_n) = -\frac{1}{3v} \int_0^\beta d\tau e^{i\omega_n\tau} \langle T_\tau \mathbf{j}(\tau) \cdot \mathbf{j}(0) \rangle \quad (8.34)$$

The retarded function $\tau_{\text{ret}}(\omega)$ is obtained by letting $i\omega_n \rightarrow \omega + i\delta$, and the dc conductivity is given by the limit of $\omega \rightarrow 0$:

$$\sigma = -\lim_{\omega \rightarrow 0} \left\{ \frac{\text{Im}[\pi_{\text{ret}}(\omega)]}{\omega} \right\} \quad (8.35)$$

This correlation function for σ is evaluated for the same model system described in Sec. 8.1.1. There is a free-particle system with a dilute concentration of simple scattering centers. Equation (8.34) is evaluated for the following Hamiltonian and current operator:

$$H = \sum_{\mathbf{p}\sigma} \xi_p C_{\mathbf{p}\sigma}^\dagger C_{\mathbf{p}\sigma} + V \quad (8.36)$$

$$V = \frac{1}{v} \sum_{\mathbf{q}} V(q) \rho_i(\mathbf{q}) \rho(\mathbf{q}) \quad (8.37)$$

$$\rho_i(\mathbf{q}) = \sum_j e^{i\mathbf{q} \cdot \mathbf{R}_j} \quad (8.38)$$

$$\rho(\mathbf{q}) = \sum_{\mathbf{p}\sigma} C_{\mathbf{p}+\mathbf{q},\sigma}^\dagger C_{\mathbf{p}\sigma} \quad (8.39)$$

$$\mathbf{j} = \frac{e}{m} \sum_{\mathbf{p}\sigma} \mathbf{p} C_{\mathbf{p}\sigma}^\dagger C_{\mathbf{p}\sigma} \quad (8.40)$$

The impurities are at positions \mathbf{R}_j , and an average will be taken over the possible distributions of impurity positions. This averaging technique was described earlier in Sec. 4.1.4.

The theoretical calculation is divided into two parts. The first part, in this section, is simply to reproduce the Boltzmann result which was derived above. How is a conductivity derived which is proportional to the relaxation time τ_i ? The derivation entails a summation over a set of vertex diagrams. The treatment is kept as introductory as possible, since it is one of the first summations over vertex diagrams. A formally exact solution to the correlation function due to Langer (1960) is presented in Sec. 8.1.3.

The logical way to evaluate the correlation function (8.34) is as a power series in the concentration of impurities. Averages over the impurity density operators in (4.91) are expressed as a power series in the number of impurities N_i :

$$\begin{aligned} f_n(\mathbf{q}_1, \dots, \mathbf{q}_n) &= \langle \rho_i(\mathbf{q}_1) \rho_i(\mathbf{q}_2) \cdots \rho_i(\mathbf{q}_n) \rangle \\ &= N_i \delta_{\sum \mathbf{q}_i} + N_i^2 \delta_{\sum \mathbf{q}_i} \delta_{\sum_j \mathbf{q}_j} + \cdots \end{aligned} \quad (8.41)$$

At first sight it appears possible to evaluate (8.34) by just expanding the S matrix and collecting all terms proportional to N_i , then N_i^2 , etc. A simple expansion in powers of N_i does not work, as is apparent from the desired answer. It has $\sigma \propto \tau_i \propto 1/n_i$. The first term has the conductivity inversely proportional to n_i , which is impossible to obtain as a simple series in N_i except by summing a set of diagrams. The expected answer is only going to be obtained by summing a series of diagrams.

The correlation function (8.34) is written in the interaction representation as

$$\pi(i\omega_n) = -\frac{1}{3v} \int_0^\beta d\tau e^{i\omega_n \tau} \langle T_\tau S(\beta) e^{\tau H_0} \mathbf{j} e^{-\tau H_0} \cdot \mathbf{j} \rangle \quad (8.42)$$

The logical way to evaluate this expression is to expand the S matrix and consider each term. The first term has $S = 1$, and this correlation function is called π_0 :

$$\begin{aligned} \pi_0(i\omega_n) &= -\frac{e^2}{3m^2v} \int_0^\beta d\tau e^{i\omega_n \tau} \sum_{\mathbf{p}\mathbf{p}'\sigma\sigma'} \mathbf{p} \cdot \mathbf{p}' \langle T_\tau C_{\mathbf{p}\sigma}^\dagger(\tau) C_{\mathbf{p}\sigma}(\tau) C_{\mathbf{p}'\sigma'}^\dagger(0) C_{\mathbf{p}'\sigma'}(0) \rangle \\ &= \frac{2e^2}{3m^2v} \int_0^\beta d\tau e^{i\omega_n \tau} \sum_{\mathbf{p}} p^2 \int_0^\beta d\tau e^{i\omega_n \tau} \mathcal{G}^{(0)}(\mathbf{p}, \tau) \mathcal{G}^{(0)}(\mathbf{p}, -\tau) \end{aligned} \quad (8.43)$$

This expression is zero unless $i\omega_n = 0$, since the number operators $C_{\mathbf{p}\sigma}^\dagger C_{\mathbf{p}\sigma}$ are τ -independent. The τ -integral gives zero unless $i\omega_n = 0$ and then it gives β . The term π_0 gives a conductivity of zero. The result is not surprising, since it is the correlation function of noninteracting particles, and they have zero resistivity. Perhaps a better answer is the correlation function for the conductivity is infinity. The zero is sufficient to alert us that a nonsensical question was asked, and a nonsensical answer was obtained. A resistive system requires putting damping into the particle motion.

The next logical step is to replace all $\mathcal{G}^{(0)}$ by \mathcal{G} . The self-energy of the particles, from impurity scattering, is included in all the particle Green's functions. Of course, \mathcal{G} is obtained by summing a series of diagrams, which is Dyson's equation. The self-energy $\Sigma(\mathbf{p}, ip_n)$ from impurity scattering was evaluated in Sec. 4.1 in the limit of low n_i . It is a retarded function with real and imaginary parts, where the imaginary parts are due to the damping of the particle motion. The step of replacing $\mathcal{G}^{(0)}$ by \mathcal{G} does put in damping of the particle motion. The first correlation function which will be evaluated is shown in Fig. 8.2. It is a simple bubble diagram, with the smooth lines denoting \mathcal{G} and the two vertices having the vector vertex \mathbf{p} . This correlation function is called $\pi^{(0)}(i\omega)$:

$$\pi^{(0)}(i\omega_n) = \frac{2e^2}{3m^2v} \sum_{\mathbf{p}} p^2 \frac{1}{\beta} \sum_{ip} \mathcal{G}(\mathbf{p}, ip + i\omega) H(\mathbf{p}, ip) \quad (8.44)$$

The wiggly lines at the two ends of the bubble, which are connected to the vertices, represent the incoming frequency $i\omega_n$. The first step is to evaluate the summation over Matsubara frequencies.

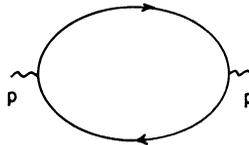


FIGURE 8.2

The procedure for doing this summation was described in Sec. 3.5. The easiest way is to represent the interacting Green's functions by the Lehmann representation

$$\begin{aligned} S(i\omega_n) &= \frac{1}{\beta} \sum_{ip} \mathcal{G}(\mathbf{p}, ip + i\omega) \mathcal{G}(\mathbf{p}, ip) \\ &= \int \frac{d\varepsilon_1}{2\pi} A(\mathbf{p}, \varepsilon_1) \int \frac{d\varepsilon_2}{2\pi} A(\mathbf{p}, \varepsilon_2) \frac{1}{\beta} \sum_{ip_m} \frac{1}{ip_m + i\omega_n - \varepsilon_1} \frac{1}{ip_m - \varepsilon_2} \end{aligned} \quad (8.45)$$

$$= \int \frac{d\varepsilon_1}{2\pi} A(\mathbf{p}, \varepsilon_1) \int \frac{d\varepsilon_2}{2\pi} A(\mathbf{p}, \varepsilon_2) \frac{n_F(\varepsilon_2) - n_F(\varepsilon_1)}{i\omega_n + \varepsilon_2 - \varepsilon_1} \quad (8.46)$$

The next step is to convert to a retarded function ($i\omega_n \rightarrow \omega + i\delta$) and then to take the imaginary part:

$$\begin{aligned} \text{Im}[S(\omega)] &= -\pi \int \frac{d\varepsilon_1}{2\pi} A(\mathbf{p}, \varepsilon_1) \int \frac{d\varepsilon_2}{2\pi} A(\mathbf{p}, \varepsilon_2) \{n_F(\varepsilon_2) - n_F(\varepsilon_1)\} \delta(\omega + \varepsilon_2 - \varepsilon_1) \\ &= - \int \frac{d\varepsilon_2}{4\pi} A(\mathbf{p}, \varepsilon_2) A(\mathbf{p}, \varepsilon_2 + \omega) \{n_F(\varepsilon_2) - n_F(\varepsilon_2 + \omega)\} \end{aligned} \quad (8.47)$$

$$\text{Im}[\pi_{\text{ret}}^{(0)}(\omega)] = -\frac{e^2}{3m^2v} \sum_{\mathbf{p}} p^2 \int \frac{d\varepsilon_2}{2\pi} A(\mathbf{p}, \varepsilon_2) A(\mathbf{p}, \varepsilon_2 + \omega) \{n_F(\varepsilon_2) - n_F(\varepsilon_2 + \omega)\}$$

The next step in the derivation (8.35) is to divide by ω and then to take the limit $\omega \rightarrow 0$. The important frequency dependence is in the last factor:

$$\lim_{\omega \rightarrow 0} \frac{1}{\omega} \{n_F(\varepsilon_2) - n_F(\varepsilon_2 + \omega)\} = -\frac{dn_F(\varepsilon_2)}{d\varepsilon_2} \quad (8.48)$$

and the conductivity from this contribution is called $\sigma^{(0)}$

$$\sigma^{(0)} = \frac{e^2}{3m^2} \int \frac{d^3p}{(2\pi)^3} p^2 \int \frac{d\varepsilon}{2\pi} A^2(\mathbf{p}, \varepsilon) \left(-\frac{dn_F(\varepsilon)}{d\varepsilon} \right) \quad (8.49)$$

In the last step, the limit $v \rightarrow \infty$ changed the summation over \mathbf{p} into a continuous integral. The right-hand side is positive since $dn_F/d\varepsilon$ is negative. Before discussing this result, it is useful to review the order of the steps in the derivation. They will be used in all evaluations of the Kubo formula:

1. Do all summations over Matsubara frequencies ip_n .
2. Analytically continue $i\omega \rightarrow \omega + i\delta$ to get the retarded function $\pi_{\text{ret}}(\omega)$.
3. Take the imaginary part $\text{Im}[\pi_{\text{ret}}(\omega)]$.
4. Divide by ω , and then take the limit $\omega \rightarrow 0$.

These steps cannot be taken out of order. Equation (8.49) for $\sigma^{(0)}$ has several interesting features. The factor $dn_F/d\varepsilon = -\delta(\varepsilon)$ at zero temperature, which is rather convenient, since it serves to eliminate the integral over $d\varepsilon$. The sharp step in $n_F(\varepsilon)$ is in contrast to the momentum distribution (3.135),

$$n_{\mathbf{p}} = \int \frac{d\varepsilon}{2\pi} n_F(\varepsilon) A(\mathbf{p}, \varepsilon) \quad (8.50)$$

which no longer has a discontinuous step at $p = k_F$ since the impurity scattering causes a general smearing of this distribution. However, the energy distribution $n_F(\varepsilon)$ is always a sharp function of ε at zero temperature regardless of the interactions.

Another crucial feature of (8.49) is that the spectral function is squared, $A(\mathbf{p}, \varepsilon)^2$. That this is important may be shown by examining the limit as the impurity concentration n_i vanishes. Since the self-energy is proportional to n_i , it will vanish in this limit. Define $\Delta = -\text{Im}(\Sigma)$, and then

$$\lim_{n_i \rightarrow 0} A(\mathbf{p}, \varepsilon) = \lim_{\Delta \rightarrow 0} \frac{2\Delta}{[\varepsilon - \xi_{\mathbf{p}} - \text{Re}(\Sigma)]^2 + \Delta^2} = 2\pi\delta(\varepsilon - \xi_{\mathbf{p}}) \quad (8.51)$$

The spectral function becomes a delta function when $n_i \rightarrow 0$. This limiting behavior is reasonable, since in the absence of self-energy effects, the particles are all free, and the spectral function is indeed a delta function. The question at hand is what happens to A^2 as $n_i \rightarrow 0$, since it appears to go as the square of a delta function. In fact, A^2 does diverge as $n_i \rightarrow 0$, which makes the conductivity diverge to infinity when $n_i \rightarrow 0$. A method is needed for handling this divergence. The answer is provided by considering the integrals

$$\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{2\Delta}{\omega^2 + \Delta^2} = 1 \quad (8.52)$$

$$\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \left(\frac{2\Delta}{\omega^2 + \Delta^2} \right)^2 = \frac{1}{\Delta} \quad (8.53)$$

where $\omega = \varepsilon - \xi_{\mathbf{p}} - \text{Re}(\Sigma)$. The first integral has the right behavior as $\Delta \rightarrow 0$, since it gives the same result as does $A = 2\Delta/(\omega^2 + \Delta^2) = 2\pi\delta(\omega)$. The second integral suggests the replacement

$$\lim_{\Delta \rightarrow 0} a^2 = \lim_{\Delta \rightarrow 0} \left(\frac{2\Delta}{\omega^2 + \Delta^2} \right)^2 = \lim_{\Delta \rightarrow 0} \frac{2\pi\delta(\omega)}{\Delta} \quad (8.54)$$

which will give the right behavior as $\Delta \rightarrow 0$. The replacement $A^2 \rightarrow 2\pi\delta(\omega)/\Delta$ will be made in the limit as $n_i \rightarrow 0$. Furthermore, the quantity 2Δ is recognized as the inverse scattering time of electrons on the Fermi surface,

$$2\Delta(k_F, \varepsilon = 0) = \frac{1}{\tau(k_F)} = -2 \text{Im}(\Sigma) \quad (8.55)$$

The conductivity formula may now be written as

$$\sigma^{(0)} = \frac{2e^2}{3m^2} \int \frac{d^3p}{(2\pi)^3} \delta(\xi_p) p^2 \tau(p) \quad (8.56)$$

This equation looks like the right answer for σ since it seems to have exactly the same combination of factors as (8.13). But there is a very important difference between (8.56) and the Boltzmann result (8.13)—in the relaxation time. The formula (8.56) has a relaxation time without the $(1 - \cos \theta')$ factor, since the relaxation time in (8.56) is from $\text{Im}(\Sigma)$, which is just the average time between scattering events. The $(1 - \cos \theta')$ factor was important for weighting the large-angle scattering processes, which were important for the resistivity. The preliminary answer (8.56) is not the Boltzmann result and has serious deficiencies.

The above derivation has one achievement. It succeeds in deriving a term in σ which is inversely proportional to n_i . The relaxation time $\tau(p)$, although the wrong one for the

conductivity, at least has the virtue that it is inversely proportional to n_i , which makes $\sigma^{(0)}$ also inversely proportional to n_i .

The calculation has not yet yielded the right result. More diagrams need to be evaluated. Among those remaining, there must be a subset which, when evaluated, will give the right answer. The S matrix is expanded for impurity scattering, and the remaining terms contain higher powers in the impurity interaction V . At first it appears that higher powers in the impurity interaction must imply that the additional terms are higher powers in the impurity concentration n_i . This conclusion is false, as is obvious from the answer. The final σ is proportional to $1/n_i$, while our preliminary term $\sigma^{(0)}$ is also proportional to $1/n_i$. The important correction terms in the S -matrix expansion must also yield terms in the conductivity which are proportional to $1/n_i$. The S -matrix expansion is examined to find the terms which cause σ to diverge as $1/n_i$ when $n_i \rightarrow 0$, although these terms must come from higher terms in the S -matrix expansion. Higher-order terms in S can be proportional to $1/n_i$ if they also contain higher powers of the spectral function A^n .

The correlation function $\sigma^{(0)}$, contained Green's functions \mathcal{G} which include all self-energy effects. The remaining diagrams are called *vertex corrections*. They are defined as diagrams in which the impurity scattering links the Green's functions on both sides of the bubble. Some vertex diagrams are shown in Fig. 8.3(a). There is a single impurity with a varying number of scattering events from the electron line on either side of the bubble. If there were no scattering line connecting one side of the bubble, the diagram would be a self-energy term on the other side. A diagram in which the two electron lines, on both sides of the bubble, scatter from the same impurity cannot be a self-energy diagram of either one and so is called a vertex diagram. Figure 8.3(a) only shows vertex diagrams with a single impurity participating in the scattering. Vertex diagrams can occur with scattering from several impurities. These are equally important and are considered later.

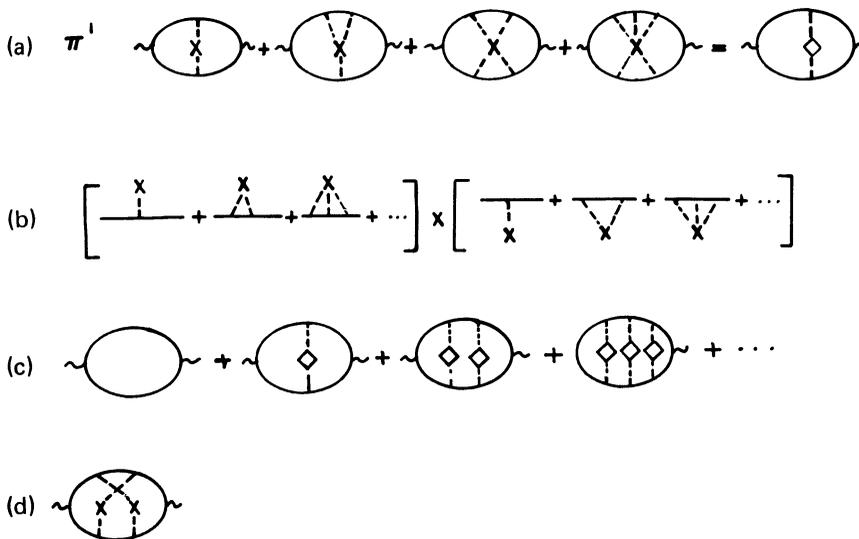


FIGURE 8.3

The sum of all the diagrams in Fig. 8.3(a) can be evaluated in a simple way and gives the correlation function:

$$W_{\mathbf{pp}'}^{(1)}(ip, ip + i\omega) = n_i T_{\mathbf{pp}'}(ip) T_{\mathbf{p}'\mathbf{p}}(ip + i\omega) \quad (8.57)$$

$$\begin{aligned} \pi^{(1)}(i\omega) &= \frac{2e^2}{3m^2v^2} \sum_{\mathbf{pp}'} \frac{1}{\beta} \sum_{ip} \mathbf{p} \cdot \mathbf{p}' \mathcal{G}(\mathbf{p}, ip) \mathcal{G}(\mathbf{p}, ip + i\omega) \mathcal{G}(\mathbf{p}', ip) \\ &\quad \times \mathcal{G}(\mathbf{p}', ip + i\omega) W_{\mathbf{pp}'}^{(1)}(ip, ip + i\omega) \end{aligned} \quad (8.58)$$

Equation (8.58) is easy to prove once the rules are recalled for scattering from a single impurity: (1) total momentum is conserved, and (2) energy is not changed from ip or $ip + i\omega$. The second rule says that all electron lines in one string, say the top, are at the same energy ip , while the ones on the bottom are at $ip + i\omega$. The momentum conservation requires that the momentum transfer on the top electron line be $\mathbf{p} - \mathbf{p}'$, which is exactly the opposite of the other: $\mathbf{p}' - \mathbf{p}$. The evaluation is easy, since the two sides decouple, as is illustrated schematically in Fig. 8.3(b). The effective interaction $W^{(1)}$ is represented by a diamond in the figures, which is illustrated to the right of Fig. 8.3(a). The diamond is the total vertex scattering from a single impurity.

Figure 8.3(c) shows the sum of correlation functions which have increasing numbers of diamonds in them. These contributions are called *ladder diagrams*. Summing the ladder diagrams achieves the objective of a relaxation time with the factor $(1 - \cos \theta')$. It is important to realize that the sum of diagrams in Fig. 8.3(c) is not the only contribution with scattering from several impurity sites. An example of a nonladder diagram is shown in Fig. 8.3(d). This contribution is not included in the series shown in Fig. 8.3(c). It is of order $O(n_i^2)$ and is neglected. The sum of ladder diagrams omits many terms. However, the omitted terms are not as important in the limit where $n_i \rightarrow 0$.

The first two terms in the sum of ladder diagrams have already been derived; they are $\pi^{(0)}$, and $\pi^{(1)}$, in (8.44) and (8.58). The superscript denotes the number of ladders. The superscript (L) denotes sum of ladder diagrams. The series of terms in the ladder sum can be generated by representing them as a vector vertex function $\Gamma^{(L)}$:

$$\pi^{(L)}(i\omega) = \frac{2e^2}{3m^2v} \sum_{\mathbf{p}} \frac{1}{\beta} \sum_{ip} \mathbf{p} \cdot \Gamma^{(L)}(\mathbf{p}, ip, ip + i\omega) \mathcal{G}(\mathbf{p}, ip) \mathcal{G}(\mathbf{p}, ip + i\omega) \quad (8.59)$$

$$\begin{aligned} \Gamma^{(L)}(\mathbf{p}; ip, ip + i\omega) &= \mathbf{p} + \frac{1}{v} \sum_{\mathbf{p}'} \Gamma^{(L)}(\mathbf{p}'; ip, ip + i\omega) W_{\mathbf{pp}'}^{(1)}(ip, ip + i\omega) \\ &\quad \times \mathcal{G}(\mathbf{p}', ip) \mathcal{G}(\mathbf{p}', ip + i\omega) \end{aligned} \quad (8.60)$$

Repeated iteration of (8.60) will generate the series of terms in the ladder summation shown in Fig. 8.3(c). The ladder summation will be evaluated, approximately, in order to obtain the factor of $(1 - \cos \theta')$ in the relaxation time.

The correlation function is a function of ip and $i\omega$ in the combination ip and $ip + i\omega$. Define the quantity

$$\begin{aligned} P(ip, ip + i\omega) &= \frac{2e^2}{3m^2v} \sum_{\mathbf{p}} \mathbf{p} \cdot \Gamma^{(L)}(\mathbf{p}; ip, ip + i\omega) \mathcal{G}(\mathbf{p}, ip) \mathcal{G}(\mathbf{p}, ip + i\omega) \\ \pi^{(L)}(i\omega) &= \frac{1}{\beta} \sum_{ip} P(ip, ip + i\omega) \end{aligned} \quad (8.61)$$

The summation over Matsubara frequencies ip is evaluated, as usual, by examining the contour integral $\int dz n_F(z)P(z, z + i\omega)$. The integrand has the poles of $n_F(z)$, which give the summation over ip , and also branch cuts along the two axes $z = \epsilon$ and $z = \epsilon + i\omega$, where ϵ is real. These cuts are shown in Fig. 8.4. The integral equals the contribution from the two branch cuts, where one has to subtract the parts above and below each cut:

$$\begin{aligned} \pi^{(L)}(i\omega_n) = & - \int_{-\infty}^{\infty} \frac{d\epsilon}{2\pi i} n_F(\epsilon) [P(\epsilon + i\delta, \epsilon + i\omega_n) - P(\epsilon - i\delta, \epsilon + i\omega_n) \\ & + P(\epsilon - i\omega_n, \epsilon + i\delta) - P(\epsilon - i\omega_n, \epsilon - i\delta)] \end{aligned} \quad (8.62)$$

The infinitesimal part $\pm i\delta$ is unnecessary when there is a term with $\pm i\omega_n$. Next find the retarded function from the analytical continuation $i\omega_n \rightarrow \omega + i\delta$. A variable change $\epsilon \rightarrow \epsilon + \omega$ in the last two terms brings us to the point

$$\begin{aligned} \pi_{\text{ret}}^{(L)}(\omega) = & \int_{-\infty}^{\infty} \frac{d\epsilon}{2\pi i} \{ [n_F(\epsilon) - n_F(\epsilon + \omega)] P(\epsilon - i\delta, \epsilon + \omega + i\delta) \\ & - n_F(\epsilon) P(\epsilon + i\delta, \epsilon + \omega + i\delta) + n_F(\epsilon + \omega) P(\epsilon - i\delta, \epsilon + \omega - i\delta) \} \end{aligned} \quad (8.63)$$

The next step is to take the imaginary part of this expression. Because of the factor of i in the denominator $2\pi i$, this step gives the real part of the integrand. The subsequent step is to take the limit that $\omega \rightarrow 0$. The function $P(\epsilon - i\delta, \epsilon + \omega + i\delta)$ becomes $P(\epsilon - i\delta, \epsilon + i\delta)$ at zero frequency, which is real. It is real because the function is symmetric in its arguments, $P(z_1, z_2) = P(z_2, z_1)$, so that the complex conjugate of $P(\epsilon - i\delta, \epsilon + i\delta)$ equals itself. The past two factors of P in the above equation are complex conjugates of each other. Taking the real part of the integrand removes their imaginary parts, and the real parts are equal.

The limit $\omega \rightarrow 0$ again only involves $[n_F(\epsilon) - n_F(\epsilon + \omega)]/\omega$. These steps brings us to a formal expression for the dc conductivity:

$$\begin{aligned} \sigma = & - \lim_{\omega \rightarrow 0} \left(\frac{\text{Im}[\pi_{\text{ret}}^{(L)}]}{\omega} \right) \\ = & \int_{-\infty}^{\infty} \frac{d\epsilon}{2\pi i} \left(- \frac{dn_F(\epsilon)}{d\epsilon} \right) \{ P(\epsilon - i\delta, \epsilon + i\delta) - \text{Re}[P(\epsilon + i\delta, \epsilon + i\delta)] \} \end{aligned}$$

There are only two functions which need to be found, $P(\epsilon - i\delta, \epsilon + i\delta)$ and $P(\epsilon + i\delta, \epsilon + i\delta)$. At zero temperature, where $-dn_F/d\epsilon = \delta(\epsilon)$, they need to be found only at $\epsilon = 0$. These two functions have quite different behavior and are obtained by different methods. Both are usually important, but the most singular is $P(\epsilon - i\delta, \epsilon + i\delta)$, and this term leads to the factor of $(1 - \cos \theta')$ in the lifetime.

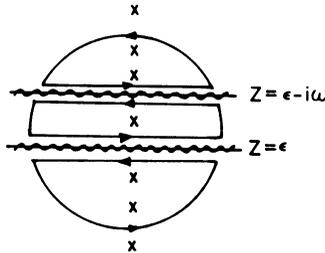


FIGURE 8.4 Branch cuts in the contour integral.

Consider the definition of P in (8.61):

$$P(\varepsilon - i\delta, \varepsilon + i\delta) = \frac{2e^2}{3m^2v} \sum_{\mathbf{p}} \mathbf{p} \cdot \mathbf{\Gamma}^{(L)}(\mathbf{p}, \varepsilon - i\delta, \varepsilon + i\delta) \mathcal{G}(\mathbf{p}, \varepsilon - i\delta) \mathcal{G}(\mathbf{p}, \varepsilon + i\delta)$$

The product of Green's functions is entirely real:

$$\begin{aligned} \mathcal{G}(\mathbf{p}, \varepsilon - i\delta) \mathcal{G}(\mathbf{p}, \varepsilon + i\delta) &= G_{\text{adv}}(\mathbf{p}\varepsilon) G_{\text{ret}}(\mathbf{p}, \varepsilon) = \frac{1}{\omega^2 + \Delta^2} \\ &= \frac{A(\mathbf{p}, \varepsilon)}{2\Delta(\mathbf{p}, \varepsilon)} \end{aligned} \quad (8.65)$$

$$\begin{aligned} \omega &= \varepsilon - \xi_{\mathbf{p}} - \text{Re}[\Sigma(\mathbf{p}, \varepsilon)], \quad \Delta = -\text{Im}[\Sigma] \\ P(\varepsilon - i\delta, \varepsilon + i\delta) &= \frac{e^2}{3m^2} \int \frac{d^3p}{(2\pi)^3} \frac{A(\mathbf{p}, \varepsilon)}{\Delta(\mathbf{p}, \varepsilon)} \mathbf{p} \cdot \mathbf{\Gamma}^{(L)}(\mathbf{p}, \varepsilon - i\delta, \varepsilon + i\delta) \end{aligned} \quad (8.66)$$

The combination $G_{\text{adv}}G_{\text{ret}}$ is rigorously defined as the spectral function $A(\mathbf{p}, \varepsilon)$ divided by $2\Delta(\mathbf{p}, \varepsilon)$. There is no assumption that $A^2 = A/2\Delta$. The same combination is found in the equation for the vertex function:

$$\begin{aligned} \mathbf{\Gamma}^{(L)}(\mathbf{p}, \varepsilon - i\delta, \varepsilon + i\delta) &= \mathbf{p} + \int \frac{d^3p'}{(2\pi)^3} \mathbf{\Gamma}^{(L)}(\mathbf{p}', \varepsilon - i\delta, \varepsilon + i\delta) \\ &\quad \times W_{\mathbf{p}\mathbf{p}'}^{(L)}(\varepsilon - i\delta, \varepsilon + i\delta) G_{\text{adv}}(\mathbf{p}', \varepsilon) G_{\text{ret}}(\mathbf{p}', \varepsilon) \end{aligned} \quad (8.67)$$

The vector function $\mathbf{\Gamma}$ must point in the direction of \mathbf{p} since that is the only vector in its function arguments. It is convenient to define an integral equation for the scalar function:

$$\mathbf{p}\gamma(\mathbf{p}, \varepsilon) = \mathbf{\Gamma}^{(L)}(\mathbf{p}, \varepsilon - i\delta, \varepsilon + i\delta) \quad (8.68)$$

$$\gamma(\mathbf{p}, \varepsilon) = 1 + \int \frac{d^3p'}{(2\pi)^3} \frac{A(\mathbf{p}', \varepsilon)}{2\Delta(\mathbf{p}', \varepsilon)} \frac{\mathbf{p} \cdot \mathbf{p}'}{p^2} W_{\mathbf{p}\mathbf{p}'}^{(L)}(\varepsilon - i\delta, \varepsilon + i\delta) \gamma(\mathbf{p}', \varepsilon) \quad (8.69)$$

$$P(\varepsilon - i\delta, \varepsilon + i\delta) = \frac{e^2}{3m^2} \int \frac{d^3p}{(2\pi)^3} \frac{A(\mathbf{p}, \varepsilon)}{\Delta(\mathbf{p}, \varepsilon)} p^2 \gamma(\mathbf{p}, \varepsilon) \quad (8.70)$$

Equation (8.69) is a one-dimensional integral equation for the scalar function $\gamma(\mathbf{p}, \varepsilon)$, where p' is the integration variable. The angular integrals $d\omega_{p'}$ just average the quantity $\mathbf{p} \cdot \mathbf{p}' W_{\mathbf{p}\mathbf{p}'}^{(L)}$ over angles to provide the kernel for the integral equation. The integral equation should not be difficult to solve with modern computers for realistic self-energies and T -matrices.

The Boltzmann result is obtained in the twin limits $T \rightarrow 0$ ($\varepsilon \rightarrow 0$) and $n_i \rightarrow 0$, where $A(\mathbf{p}', 0) \rightarrow 2\pi\delta(\xi_{\mathbf{p}'})$. Equation (8.69) then reduces to the integral equation

$$\gamma(\mathbf{p}, \varepsilon) = 1 + \int \frac{d^3p'}{(2\pi)^3} \frac{2\pi\delta(\xi_{\mathbf{p}'})}{2\Delta(\mathbf{p}', \varepsilon)} \frac{\mathbf{p} \cdot \mathbf{p}'}{p^2} n_i |T_{\mathbf{p}\mathbf{p}'}|^2 \gamma(\mathbf{p}', \varepsilon) \quad (8.71)$$

$$\gamma(k_F) = 1 + \gamma(k_F) \frac{\Delta - \Delta_1}{\Delta} \quad (8.72)$$

$$2\Delta_1 = 2\pi n_i \int \frac{d^3p'}{(2\pi)^3} \delta(\xi_{\mathbf{p}'}) \left[1 - \frac{\mathbf{p} \cdot \mathbf{p}'}{p^2} \right] |T_{\mathbf{p}\mathbf{p}'}|^2 = \frac{1}{\tau_t} \quad (8.73)$$

which is easily solved to give

$$\gamma(k_F) = \frac{\Delta}{\Delta_1} \quad (8.74)$$

The factor $(1 - \mathbf{p} \cdot \mathbf{p}'/p^2) = (1 - \cos \theta')$ since $|\mathbf{p}'| = |\mathbf{p}|$. The solution $\gamma = \Delta/\Delta_1$ is put into (8.65) to give

$$P(\varepsilon - i\delta, \varepsilon + i\delta) = \frac{e^2}{3m^2} \int \frac{d^3p}{(2\pi)^3} p^2 \frac{A(\mathbf{p}, \varepsilon)}{\Delta_1(\mathbf{p}, \varepsilon)} \quad (8.75)$$

which gives the same conductivity as (8.13) when put into (8.64). The term $P(\varepsilon - i\delta, \varepsilon + i\delta)$ leads to the important contribution as $T \rightarrow 0$ and $n_i \rightarrow 0$.

The other term $P(\varepsilon + i\delta, \varepsilon + i\delta)$ should not be neglected. It is complex but not very singular in the limit where $n_i \rightarrow 0$. As is evident from the definition (8.61) the singular parts should arise from the Green's function product:

$$G(\mathbf{p}, \varepsilon + i\delta)^2 = G_{\text{ret}}(\mathbf{p}, \varepsilon)^2 = \frac{1}{(\omega - i\Delta)^2} = \frac{\omega^2 - \Delta^2 - 2i\omega\Delta}{(\omega^2 + \Delta^2)^2} \quad (8.76)$$

$$= \frac{1}{\omega^2 + \Delta^2} - \frac{2\Delta^2}{(\omega^2 + \Delta^2)^2} - \frac{2i\omega\Delta}{(\omega^2 + \Delta^2)^2} \quad (8.77)$$

$$= \frac{A}{2\Delta} - \frac{1}{2}A^2 - i \text{Re}[G]A \quad (8.78)$$

As $n_i \rightarrow 0$ then $A^2 \rightarrow A/\Delta$ and the real part of this expression vanishes. The imaginary part becomes $\omega\delta(\omega)$, which is also zero. The singular parts of this expression vanish as $n_i \rightarrow 0$. The vertex corrections to $P(\varepsilon + i\delta, \varepsilon + i\delta)$ are not of order unity but of order $O(n_i)$. In this case, the vertex corrections may actually be a series of terms which are successively smaller, so that the vertex corrections may be obtained by just evaluating the first few. The situation is quite different than for $P(\varepsilon - i\delta, \varepsilon + i\delta)$, where one has to solve the vertex equation and sum all the ladder diagrams.

One can evaluate $P(\varepsilon + i\delta, \varepsilon + i\delta)$ by solving a vertex equation similar to (8.69). Define the scalar vertex γ' by $\mathbf{\Gamma}^{(L)}(\mathbf{p}, \varepsilon + i\delta, \varepsilon + i\delta) = \mathbf{p}\gamma'(p, \varepsilon)$, and then

$$\gamma'(p, \varepsilon) = 1 + n_i \int \frac{d^3p'}{(2\pi)^3} \frac{\mathbf{p} \cdot \mathbf{p}'}{p^2} T_{\mathbf{p}\mathbf{p}'}(\varepsilon)^2 \gamma'(p', \varepsilon) G_{\text{ret}}(\mathbf{p}', \varepsilon)^2 \quad (8.79)$$

The vertex function $\gamma'(p, \varepsilon)$ is complex, as is the vertex $T_{\mathbf{p}\mathbf{p}'}(\varepsilon)^2$, and the product $G_{\text{ret}}(\mathbf{p}', \varepsilon)^2$. There are two coupled equations for the real and imaginary parts. It may be obtained from the *Ward identity* (1950),

$$\mathbf{\Gamma}(\mathbf{p}, \varepsilon + i\delta, \varepsilon + i\delta) = \mathbf{p} + m\mathbf{\nabla}_{\mathbf{p}}\Sigma(\mathbf{p}, \varepsilon + i\delta) \quad (8.80)$$

which is an exact identity between the exact vertex function $\mathbf{\Gamma}(\mathbf{p}, \varepsilon + i\delta, \varepsilon + i\delta)$ and the exact retarded self-energy $\Sigma(\mathbf{p}, \varepsilon + i\delta)$ from impurity scattering. This self-energy is not the one derived in Sec. 4.1, since that was only for scattering from a single impurity and only had unperturbed Green's functions $\mathcal{G}^{(0)}$ as the internal lines in the scattering equation. The exact self-energy is found from scattering from all numbers of impurities and using exact Green's functions as internal lines of diagram—although this procedure must be done carefully in order not to count the same contribution twice. The Ward identity is very convenient, since it

permits the vertex function to be obtained from the self-energy function by a simple operation. It will be proved in the next section.

The present derivation has an approximate vertex function which contains only the ladder diagrams. In that case the Ward identity is still valid as long as the self-energy is found using only diagrams where an electron scatters from one impurity at a time.

The Green's function derivation of this result is certainly harder than the transport equation of Sec. 8.1.1. The one advantage of the Green's function method is that there is no need to make the approximation that $n_i \rightarrow 0$. The $n_i \rightarrow 0$ limit is made implicitly in solving the Boltzmann equation when it is assumed the particles are plane waves except for occasional scattering events from isolated impurities.

There are several lessons to be learned from this summation of ladder diagrams. The ladder diagrams, although they appear in a higher order of perturbation theory, do not lead to terms in the final answer which are smaller than lower-order terms. A term is not necessarily small if it occurs in a higher order of perturbation theory. It is this feature of vertex corrections which means they can never be dismissed without an investigation. They may be small, but one should always check. It would be nice to have some rules of thumb which would establish whether vertex corrections are important. The way to tell is to examine the scattering process which causes the vertex functions. If the two particle states, on each side of the bubble, can scatter quasielastically so that their relative energy changes little, then vertex corrections are large. The vertex correction is basically a potential divided by an energy denominator $V/\Delta E$, where ΔE is the change in energy of the two particles. If it is small, the small denominator will compensate for the small potential V , so that vertex corrections become sizable. Repeated scatterings, as in a series of ladder diagrams, just cause additional powers of the factor $(V/\Delta E)^n$. Vertex corrections are large when the scattering by the potential causes only a small change in the relative energy of the two particles.

8.1.3. Ward identities

The evaluation of a two-particle correlation function, such as the Kubo formula for the conductivity, often requires an evaluation of a vertex function. The Ward (1950) identity is an exact relationship between the vertex and self-energy functions in the problem. As an example, two types of Ward identities permit the evaluation of the scalar vertex function $\Gamma(\mathbf{p}, ip)$ or the vector vertex function $\mathbf{\Gamma}(\mathbf{p}, ip)$ which satisfy the equations

$$\Gamma(\mathbf{p}, ip) = 1 + \int \frac{d^3 p'}{(2\pi)^3} \Gamma(\mathbf{p}', ip) \mathcal{G}(\mathbf{p}', ip)^2 W_{\mathbf{p}\mathbf{p}'}(ip, ip) \quad (8.81)$$

$$\mathbf{\Gamma}(\mathbf{p}, ip) = \mathbf{p} + \int \frac{d^3 p'}{(2\pi)^3} \mathbf{\Gamma}(\mathbf{p}', ip) \mathcal{G}(\mathbf{p}', ip)^2 W_{\mathbf{p}\mathbf{p}'}(ip, ip) \quad (8.82)$$

The Ward identity states that these two functions are given by

$$\Gamma(\mathbf{p}, ip) = 1 - \left[\frac{\partial \Sigma(\mathbf{p}, z)}{\partial z} \right]_{z=ip} \quad (8.83)$$

$$\mathbf{\Gamma}(\mathbf{p}, ip) = \mathbf{p} + m \mathbf{\nabla}_p \Sigma(\mathbf{p}, ip) \quad (8.84)$$

An evaluation of the self-energy function $\Sigma(\mathbf{p}, ip)$ permits an easy evaluation of these two vertex functions. These relationships are proved below.

An important point regarding the Ward identities is that they are not useful for evaluating all vertex functions. An example is provided in the last section, where the Ward identities were useful for finding $P(\varepsilon + i\delta, \varepsilon + i\delta)$ but not $P(\varepsilon - i\delta, \varepsilon + i\delta)$. The Ward identities cannot be applied blindly; they must be used only when appropriate. These circumstances are delineated after the identities are proved.

The Ward identities for impurity scattering were derived by Langer (1961). The similar theorems for the electron–phonon interaction were derived by Engelsberg and Schrieffer (1963). We shall prove the result for the ladder diagrams obtained by scattering from a single impurity. In this case, the self-energy diagram is that for scattering from a single impurity:

$$\begin{aligned} \Sigma(\mathbf{p}, ip) = n_i \left\{ V_{\mathbf{pp}} + \int \frac{d^3 p'}{(2\pi)^3} V_{\mathbf{pp}'} V_{\mathbf{p}'\mathbf{p}} \mathcal{G}(\mathbf{p}', ip) \right. \\ \left. + \int \frac{d^3 p' d^3 p''}{(2\pi)^6} V_{\mathbf{pp}'} V_{\mathbf{p}'\mathbf{p}''} V_{\mathbf{p}''\mathbf{p}} \mathcal{G}(\mathbf{p}', ip) \mathcal{G}(\mathbf{p}'', ip) + \dots \right\} \end{aligned} \quad (8.85)$$

An important condition is that the Green's functions in this self-energy diagram are those calculated with the self-energy $\mathcal{G} = 1/[ip - \xi_p - \Sigma]$. Equation (8.85) is a self-consistent equation for the self-energy Σ , since it depends functionally on itself. Unfortunately, the Ward identities do not let one avoid solving an integral equation. Instead, one integral equation is exchanged for another. In this sense the Ward identities are not very useful in practice.

Rather than prove the two separate identities (8.83) and (8.84), a general theorem is proved for which these are two limiting cases. The general theorem is obtained by subtracting the expressions (8.85) for $\Sigma(\mathbf{p}, ip) \equiv \Sigma(p)$ by the same result for $\Sigma(\mathbf{p} + \mathbf{q}, ip + i\omega) \equiv \Sigma(p + q)$:

$$\begin{aligned} \Sigma(p + q) - \Sigma(p) = n_i \int \frac{d^3 p_1}{(2\pi)^3} V_{\mathbf{pp}_1} V_{\mathbf{p}_1\mathbf{p}} [\mathcal{G}(p_1 + q) - \mathcal{G}(p_1)] \\ + n_i \int \frac{d^3 p_1 d^3 p_2}{(2\pi)^6} V_{\mathbf{pp}_1} V_{\mathbf{p}_1\mathbf{p}_2} V_{\mathbf{p}_2\mathbf{p}} [\mathcal{G}(p_1 + q) \mathcal{G}(p_2 + q) - \mathcal{G}(p_1) \mathcal{G}(p_2)] \\ + n_i \int \frac{d^3 p_1 d^3 p_2 d^3 p_3}{(2\pi)^9} V_{\mathbf{pp}_1} V_{\mathbf{p}_1\mathbf{p}_2} V_{\mathbf{p}_2\mathbf{p}_3} V_{\mathbf{p}_3\mathbf{p}} \\ \times [\mathcal{G}(p_1 + q) \mathcal{G}(p_2 + q) \mathcal{G}(p_3 + q) - \mathcal{G}(p_1) \mathcal{G}(p_2) \mathcal{G}(p_3)] + \dots \end{aligned} \quad (8.86)$$

By purely algebraic manipulations, this series can be shown to be identical to

$$\begin{aligned} \Sigma(p + q) - \Sigma(p) = n_i \int \frac{d^3 p_1}{(2\pi)^3} T_{\mathbf{pp}_1}(ip) T_{\mathbf{p}_1+\mathbf{q},\mathbf{p}+\mathbf{q}}(ip + i\omega) [\mathcal{G}(p_1 + q) \mathcal{G}(p_1)] \\ \times [\Sigma(p_1 + q) - \Sigma(p_1) + \xi_{\mathbf{p}_1+\mathbf{q}} - \xi_{\mathbf{p}_1} - i\omega] \end{aligned} \quad (8.87)$$

$$= n_i \int \frac{d^3 p_1}{(2\pi)^3} T_{\mathbf{pp}_1}(ip) T_{\mathbf{p}_1+\mathbf{q},\mathbf{p}+\mathbf{q}}(ip + i\omega) [\mathcal{G}(p_1 + q) - \mathcal{G}(p_1)] \quad (8.88)$$

This rather startling result may be demonstrated term by term. The first nonvanishing term has the integrand

$$\begin{aligned} V_{\mathbf{pp}_1} V_{\mathbf{p}_1\mathbf{p}} [\mathcal{G}(p_1 + q) - \mathcal{G}(p_1)] = V_{\mathbf{pp}_1} V_{\mathbf{p}_1\mathbf{p}} \mathcal{G}(p_1 + q) \mathcal{G}(p_1) \\ \times [\Sigma(p_1 + q) - \Sigma(p_1) + \xi_{\mathbf{p}_1+\mathbf{q}} - \xi_{\mathbf{p}_1} - i\omega] \end{aligned} \quad (8.89)$$

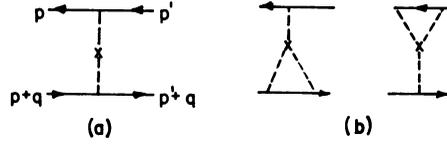


FIGURE 8.5

which is just the vertex diagram shown in Fig. 8.5(a). The next term in the series (8.86) has the integrand

$$V_{pp_1} V_{p_1 p_2} V_{p_2 p} [\mathcal{G}(p_1 + q) \mathcal{G}(p_2 + q) - \mathcal{G}(p_1) \mathcal{G}(p_2)] \tag{8.90}$$

The Green's function factors in brackets may be rearranged into

$$\begin{aligned} \mathcal{G}(p_1 + q) \mathcal{G}(p_2 + q) - \mathcal{G}(p_1) \mathcal{G}(p_2) &= \mathcal{G}(p_1 + q) [\mathcal{G}(p_2 + q) - \mathcal{G}(p_2)] \\ &\quad - \mathcal{G}(p_2) [\mathcal{G}(p_1 + q) - \mathcal{G}(p_1)] \end{aligned} \tag{8.91}$$

The first bracket on the right is the first diagram in Fig. 8.5(b). The factor $V_{pp_1} V_{p_1 p_2} \mathcal{G}(p_1 + q)$ is just the multiple scattering from the impurity by the electron on the lower line. Similarly, the second term in (8.91) corresponds to the second diagram in Fig. 8.5(b), where the factor $V_{p_1 p_2} V_{p_2 p} \mathcal{G}(p_2)$ is the multiple scattering of the top electron line. These terms are just the first in the series which generates the T -matrices. The further terms in (8.86) provide the remaining terms. In this manner, one can establish the validity of (8.88).

Equation (8.87) shows the quantity

$$\Lambda(p, p + q) = \Sigma(p + q) - \Sigma(p) + \xi_{p+q} - \xi_p - i\omega \tag{8.92}$$

obeys the vertex equation

$$\begin{aligned} \Lambda(p, p + q) &= \xi_{p+q} - \xi_p - i\omega \\ &\quad + n_i \int \frac{d^3 p_1}{(2\pi)^3} \mathcal{G}(p_1) \mathcal{G}(p_1 + q) T_{pp_1} T_{p_1+q, p+q} \Lambda(p_1, p_1 + q) \end{aligned} \tag{8.93}$$

The two equations (8.92) and (8.93) provide the most general type of Ward identity. They are useful, since any equation which can be cast into the form of (8.93) has the solution (8.92). Langer (1961) and Engelsberg and Schrieffer (1963) show that this equation is related to the equation of continuity $\nabla \cdot \mathbf{j} + \dot{\rho} = 0$.

The first Ward identity (8.83) is obtained by taking the limit $\mathbf{q} = 0$ and then dividing Eqn. (8.93) by $-i\omega$ with the result

$$\begin{aligned} \frac{\Lambda(\mathbf{p}, ip, ip + i\omega)}{-i\omega} &= 1 + n_i \int \frac{d^3 p_1}{(2\pi)^3} \mathcal{G}(\mathbf{p}_1, ip) \mathcal{G}(\mathbf{p}_1, ip + i\omega) \\ &\quad \times T_{pp_1}(ip) T_{p_1, p}(ip + i\omega) \frac{\Lambda(\mathbf{p}_1, ip, ip + i\omega)}{-i\omega} \end{aligned} \tag{8.94}$$

The quantity $\Gamma(\mathbf{p}, ip)$ in (8.81) obeys the same equation as $\Lambda/(-i\omega)$ in the limit $i\omega \rightarrow 0$ so they are equal. From (8.92) one has the solution

$$\Gamma(\mathbf{p}, ip) = \lim_{i\omega \rightarrow 0} \left[\frac{\Lambda(\mathbf{p}, ip, ip + i\omega)}{-i\omega} \right] = 1 - \left[\frac{\partial \Sigma(\mathbf{p}, Z)}{\partial Z} \right]_{Z=ip} \tag{8.95}$$

which proves the Ward identity (8.83).

The other Ward identity is found as the limit $i\omega = 0$, followed by letting $\mathbf{q} \rightarrow 0$. The latter limit is taken slowly, so that one can retain terms proportional to \mathbf{q} . In this limit Eqns. (8.92) and (8.93) become

$$\lim_{\mathbf{q} \rightarrow 0} \Lambda = \frac{\mathbf{q} \cdot \mathbf{p}}{m} + \mathbf{q} \cdot \nabla_{\mathbf{p}} \Sigma(\mathbf{p}, ip) \quad (8.96)$$

$$= \frac{\mathbf{q} \cdot \mathbf{p}}{m} + n_i \int \frac{d^3 p_1}{(2\pi)^3} \mathcal{G}(\mathbf{p}_1, ip) \mathcal{G}(\mathbf{p}_1 + \mathbf{q}, ip) T_{\mathbf{p}\mathbf{p}_1} T_{\mathbf{p}_1 + \mathbf{q}, \mathbf{p} + \mathbf{q}} \Lambda(p_1, p_1 + q) \quad (8.97)$$

The vertex function Λ is proportional to \mathbf{q} , so define the vector vertex function by the limit ($i\omega = 0$):

$$\lim_{\mathbf{q} \rightarrow 0} \Lambda = \frac{1}{m} \mathbf{q} \cdot \Gamma(\mathbf{p}, ip) \quad (8.98)$$

Then the preceding two equations can be expressed in terms of this vector vertex function:

$$\Gamma(\mathbf{p}, ip) = \mathbf{p} + m \nabla_{\mathbf{p}} \Sigma(\mathbf{p}, ip) \quad (8.99)$$

$$\Gamma(\mathbf{p}, ip) = \mathbf{p} + n_i \int \frac{d^3 p_1}{(2\pi)^3} \mathcal{G}(\mathbf{p}_1, ip)^2 |T_{\mathbf{p}\mathbf{p}_1}(ip)|^2 \Gamma(\mathbf{p}_1, ip) \quad (8.100)$$

This equation is the same as (8.82), which proves the other Ward identity (8.84). Both are now understood to be limiting cases of the general result (8.93). The Ward identities are useful anytime one can cast the vertex equation into the form (8.93).

The factor $[1 - \partial \Sigma / \partial \Sigma]$ is recognized as the inverse of the renormalization Z defined earlier and discussed, for example, in Sec. 5.8.1. This quantity is sometimes called the effective charge. Similarly, the vector vertex is one of the factors which give the effective mass of the particle. The Ward identities relate the vertex corrections to a change in the effective charge and mass of the particle, which is why they are related to the equation of continuity.

8.2. MOBILITY OF FRÖHLICH POLARONS

The Fröhlich Hamiltonian between electrons and Einstein phonons ($\omega_0 = \omega_{LO}$) is

$$H = \sum_{\mathbf{p}\sigma} \varepsilon_{\mathbf{p}} C_{\mathbf{p}\sigma}^\dagger C_{\mathbf{p}\sigma} + \omega_0 \sum_{\mathbf{q}} a_{\mathbf{q}}^\dagger a_{\mathbf{q}} + \frac{M_0}{\sqrt{V}} \sum_{\mathbf{p}\mathbf{q}} \frac{1}{|\mathbf{q}|} C_{\mathbf{p}+\mathbf{q},\sigma}^\dagger C_{\mathbf{p}\sigma} (a_{\mathbf{q}} + a_{-\mathbf{q}}^\dagger) \quad (8.101)$$

$$M_0^2 = \frac{4\pi\alpha\hbar(\hbar\omega_0)^{3/2}}{\sqrt{2m_B}}, \quad \varepsilon_{\mathbf{p}} = \frac{p^2}{2m_B}$$

$$\alpha = \frac{e^2}{\hbar} \left(\frac{m_B}{2\hbar\omega_0} \right)^{1/2} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0} \right) \quad (8.102)$$

It was discussed in Sec. 7.1, where several important quantities were derived, such as the effective mass m^* and the ground state energy E_0 .

The effective mass m^* of a particle can be measured by cyclotron resonance. Such experiments have been done for polarons (Hodby, 1972). The effective mass m^* is a function of the band mass m_B and the polar coupling constant α . A separate measurement of the two dielectric functions ε_0 and ε_∞ , as well as ω_{LO} , permits a determination of the band mass m_B

and α from m^* . This analysis takes a theory of the polaron mass $m^*(m_B, \alpha)$, which was provided in Sec. 7.1. Another way to check the theory is to measure the mobility of electrons in insulators. The mobility also depends upon m_B and α .

A typical experimental result for the mobility is shown in Fig. 8.6. It shows the Hall mobility of CdTe measured by Segall *et al.* (1963). The steep rise around 200 K is due to optical mode scattering. At lower temperatures the mobility saturates because of the scattering from impurities in the crystals. Impurity scattering varies from sample to sample, depending on the concentration and type of impurity.

The average value of the current operator is the particle density n_0 times the charge e times the average velocity $\langle \mathbf{v} \rangle$. The average velocity $\langle \mathbf{v} \rangle$ is proportional to the applied electric field \mathbf{E} , and the constant of proportionality is the mobility, where $E > 0$ so the electron charge is $-e$:

$$\langle \mathbf{v} \rangle = \mu \mathbf{E} \tag{8.103}$$

$$\mathbf{J} = -en_0\mu\mathbf{E} \tag{8.104}$$

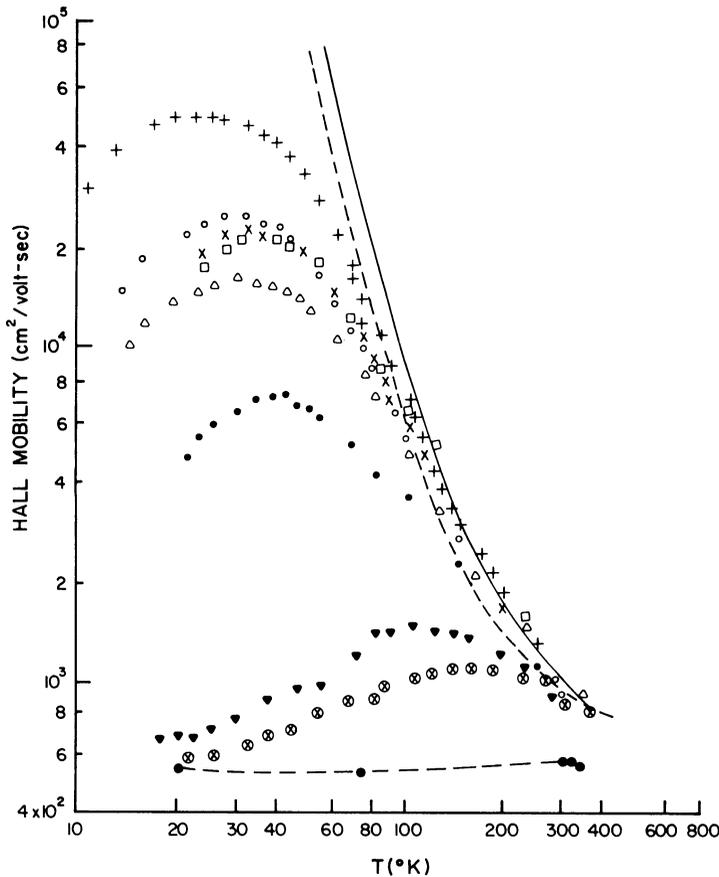


FIGURE 8.6 The temperature dependence of the electron mobility of several samples of *n*-type CdTe. Different samples have different kinds and concentrations of defects or impurities. The solid line is the theoretical mobility from optical mode scattering, including the temperature dependence of the static dielectric constant. The dashed line neglects this change. *Source:* Segall *et al.* (1963) (used with permission).

The mobility μ is the average velocity of each electron per unit applied electric field. Of course, it is strictly defined in the limit of vanishing electric field. Since the electrical conductivity $\sigma = n_0 e^2 \tau_i / m = -e \mu n_0$ is the ratio of the current \mathbf{J} to the field \mathbf{E} , then $\mu = -e \tau_i / m$.

The theories of the electron mobility in insulating materials, such as alkali halides and II–VI semiconductors, treat it as a property of a single electron. The electron lifetime is calculated for the scattering from impurities and by acoustical and optical phonons. The electron–electron interactions can be ignored in the limit where the concentration of electrons is very low. Spin is also unimportant, and the spin index is omitted in this section.

There are just as many different ways to calculate the polaron mobility as there are to calculate the effective mass m^* or ground state energy E_0 . Each theoretical technique was applied to the mobility as well as to the other quantities. As summarized by Langreth (1967), these various methods usually agree in the limit of weak coupling ($\alpha \ll 1$) and low temperature $k_B T \ll \hbar \omega_0$. In this case

$$\lim_{T \rightarrow 0} \mu = -\frac{e \tau_0}{m_B} \equiv \mu_0 \quad (8.105)$$

$$\frac{1}{\tau_0} = 2\alpha N_0 \omega_0, \quad N_0 = \frac{1}{e^{\beta \hbar \omega_0} - 1} \quad (8.106)$$

The lifetime τ_0 is the result obtained in (7.46) as $1/\tau_0 = -2 \operatorname{Im}[\Sigma^{(1)}(p, \varepsilon)]$ in the limit where $p \rightarrow 0$ and $\varepsilon \rightarrow 0$, where $\Sigma^{(1)}$ is the one-phonon self-energy. This limit is appropriate, since at very low temperatures the electrons are in states within $k_B T \ll \hbar \omega_0$ of the bottom of the band. These low-energy particles cannot emit phonons, since this event is prevented by energy conservation. They can only absorb them, and the rate of absorption is proportional to the thermal average density of phonons N_0 . The factor N_0 makes the mobility increase exponentially with decreasing temperature, since the electron scattering becomes less likely as the number density of phonons declines. The exponential increase in the mobility is evident in the experimental data of Fig. 8.6. The behavior of large polarons is opposite to that of small polarons, whose mobility increases with increasing temperature.

One feature of the mobility formula (8.105) is that it is proportional to the inverse of α . Our starting point for the theoretical calculation is again the Kubo formula, which will be evaluated for using electron–phonon interactions. The expansion of the S matrix for this potential will generate a series in the parameter α . To obtain a leading term in the inverse power of α requires the summation of a subset of diagrams. The situation is similar to the mobility from impurity scattering, where diagrams were summed to get the conductivity inversely proportional to the impurity concentration n_i . There are important differences between the lifetime from optical phonon scattering and that of impurity scattering. This conclusion is evident from the result presented in (8.105). Here the relaxation time is not calculated with the factor of $(1 - \cos \theta')$ in the angular average. The polaron mobility is calculated in a different way than the scattering from impurities. Actually it is calculated in the same way, but a different result is obtained. This difference arises from the inelastic nature of the polaron scattering as first shown by Howarth and Sondheimer (1953).

Langreth and Kadanoff (1964) showed that the polaron mobility μ is a power series in α , with the leading term in (8.105) of order $1/\alpha$:

$$\mu = \frac{a_{-1}}{\alpha} + a_0 \alpha^0 + a_1 \alpha^1 + \cdots \quad (8.107)$$

They tried to calculate the coefficient a_{-1} , a_0 in this series in the limit of low temperature. This objective is simple but has slippery aspects (Mahan, 1966). It is a subtle procedure to examine each term in the S -matrix expansion and to determine its leading term in α . The situation is similar to that for impurity scattering. Each term in the series (8.107) is obtained by summing subsets of diagrams.

The first term of Langreth and Kadanoff (a_{-1}) is quite simple, and it is probably worthwhile to state it in advance. Define μ_0 as the result (8.105) for the limits $\alpha \rightarrow 0$, $T \rightarrow 0$. They found

$$\frac{\mu}{\mu_0} = 1 - \frac{\alpha}{6} + O(\alpha^2) \quad (8.108)$$

They observed that μ_0/μ is precisely the expansion in α given by the equations

$$\mu = -\frac{e\tau}{m^*} \quad (8.109)$$

$$\frac{\tau}{\tau_0} = 1 + O(\alpha^2) \quad (8.110)$$

$$\frac{m^*}{m_B} = 1 + \frac{\alpha}{6} + O(\alpha^2) \quad (8.111)$$

The ballistic formula $\mu = -e\tau/m^*$ supports the quasiparticle picture that the particle acts as if it has an effective mass m^* and lifetime τ . They speculate that the quasiparticle picture would be valid for all values of α and that the inclusion of all terms in α would just reproduce the product series of m^* and τ .

Before the derivation of the electron mobility, it is necessary to derive some single-particle properties. These will be needed in the limit of zero temperature. All terms are dropped of order N_0 compared to unity. Many of these single-particle properties were derived in Sec. 7.1. The first self-energy term, proportional to α , is the one-phonon result in (7.15):

$$\text{Re}[\Sigma^{(1)}(p, \omega)] = -\alpha \frac{\omega^{3/2}}{\sqrt{\varepsilon_p}} \sin^{-1} \left(\frac{\varepsilon_p}{\omega_0 - \omega + \varepsilon_p} \right)^{1/2} \quad (8.112)$$

At zero temperature, this self-energy is evaluated at small ε_p and small ω ,

$$\text{Re}[\Sigma^{(1)}(p, \omega)] = -\alpha[\omega_0 + \frac{1}{2}\omega - \frac{1}{3}\varepsilon_p + O(\omega^2, \varepsilon_p^3, \omega\varepsilon_p)] \quad (8.113)$$

This expansion permits a quick derivation of the effective mass m^* and the renormalization coefficient Z . The zero subscript means $p = 0$, $\omega = 0$:

$$Z_0 = \left(1 - \frac{\partial \Sigma}{\partial \omega} \right)_0^{-1} = \frac{1}{1 + \alpha/2} \approx 1 - \frac{\alpha}{2} + O(\alpha^2) \quad (8.114)$$

$$\left(\frac{m_B}{m^*} \right)_0 = Z_0 \left(1 + \frac{\partial \Sigma}{\partial \varepsilon_p} \right) = \frac{1 + \alpha/3}{1 + \alpha/2} \approx 1 - \frac{\alpha}{6} + O(\alpha^2) \quad (8.115)$$

Another important quantity is the lifetime τ which is defined as

$$\frac{1}{\tau(p)} = Z(p) \{-2 \text{Im}[\Sigma(p, E_p)]\} \quad (8.116)$$

where $Z(p)$ is the renormalization coefficient evaluated at $\omega = E_p$, which is the ground state energy. The ground state energy is only needed to order α , which is the simple result $E_p = -\alpha\omega_0 + p^2/2m^* + O(p^4)$.

Equation (8.116) shows that the renormalization coefficient $Z(p)$ enters into the definition of the lifetime. The argument for this is as follows. The spectral function is defined as

$$A(p, \omega) = -2 \operatorname{Im}[G_{\text{ret}}(p, \omega)] \quad (8.117)$$

$$= \frac{-2 \operatorname{Im}[\Sigma(p, \omega)]}{\{\omega - \varepsilon_p - \operatorname{Re}[\Sigma(p, \omega)]\}^2 + \{\operatorname{Im}[\Sigma(p, \omega)]\}^2} \quad (8.118)$$

A suitable definition of $\tau(p)$ is obtained by examining this limit more carefully when $\operatorname{Im}(\Sigma)$ is small but not infinitesimal. In the vicinity of the peak $\omega \approx E_p$ of the spectral function the term in the denominator is

$$\begin{aligned} \omega - \varepsilon_p - \operatorname{Re}[\Sigma(p, \omega)] &\approx \omega - \varepsilon_p - \operatorname{Re}[\Sigma(p, E_p)] - (\omega - E_p) \frac{\partial \operatorname{Re} \Sigma}{\partial \omega} \\ &\approx (\omega - E_p) \left[1 - \frac{\partial \operatorname{Re} \Sigma}{\partial \omega} \right] = \frac{(\omega - E_p)}{Z(p)} \end{aligned} \quad (8.119)$$

so that the spectral function is approximately

$$A(p, \omega) = \frac{-2 \operatorname{Im}[\Sigma(p, \omega)]}{(\omega - E_p)^2 / Z(p)^2 + [\operatorname{Im}(\Sigma)]^2} \quad (8.120)$$

$$= Z(p) \left(\frac{1/\tau(p)}{((\omega - E_p)^2 + (1/2\tau)^2)} \right) \quad (8.121)$$

The last step used the definition (8.116) of the relaxation time. The relaxation time $\tau(p)$ is treated as a function of p but not ω . This form for the spectral function is used in the Green's function (Problem 6 in Chapter 3),

$$G_{\text{ret}}(p, t) = \Theta(t) \int_{-\infty}^{\infty} \frac{d\omega}{2\pi i} e^{-\omega t} A(p, \omega) \quad (8.122)$$

$$= -iZ(p)\Theta(t) \exp[-itE_p - t/(2\tau)] \quad (8.123)$$

The relaxation time is defined from the decay of the Green's function. It has the desired form, with the relaxation time $\tau(p)$ determining the decay of the excitation.

There is another way to understand the factor Z in the definition (8.116) of $\tau(p)$. The quantity $-2 \operatorname{Im}(\Sigma)$ is the rate of decay of a state (p, ω) . The factor Z is the fraction of the quasiparticle strength at the value (p, ω) . The rest of the quasiparticle strength is usually dispersed throughout the spectrum.

The one-phonon self-energy is evaluated for the quasiparticle lifetime. The imaginary self-energy is evaluated at the quasiparticle energy $\omega = E_0 = -\alpha\omega_0$ and is multiplied by the factor of Z given in (8.115). The imaginary self-energy is calculated from the expression

$$-2 \operatorname{Im}[\Sigma(p, \omega)] = 2\pi \int \frac{d^3q}{(2\pi)^3} \frac{M_0^2}{q^2} \quad (8.124)$$

$$\begin{aligned} & \times [N_0 \delta(\omega + \omega_0 - \varepsilon_{\mathbf{q}+\mathbf{p}}) + (N_0 + 1) \delta(\omega - \omega_0 - \varepsilon_{\mathbf{q}+\mathbf{p}})] \\ & = \frac{\alpha(\omega_0)^{3/2}}{\sqrt{\varepsilon_p}} \left[N_0 \Theta(\omega + \omega_0) \ln \left| \frac{\sqrt{\omega + \omega_0} + \sqrt{\varepsilon_p}}{\sqrt{\omega + \omega_0} - \sqrt{\varepsilon_p}} \right| \right. \\ & \quad \left. + (N_0 + 1) \Theta(\omega - \omega_0) \ln \left| \frac{\sqrt{\omega - \omega_0} + \sqrt{\varepsilon_p}}{\sqrt{\omega - \omega_0} - \sqrt{\varepsilon_p}} \right| \right] \end{aligned} \quad (8.125)$$

Expanding for small $\omega < \omega_0$ and small ε_p gives

$$-2 \operatorname{Im}[\Sigma(0, \omega)] = \frac{2\alpha N_0 \omega_0^{3/2}}{\sqrt{\omega_0 + \omega}} \approx 2\alpha N_0 [\omega_0 - \frac{1}{2}\omega + \dots] \quad (8.126)$$

$$-2 \operatorname{Im}[\Sigma(0, \omega = -\alpha\omega_0)] = 2\alpha N_0 \omega_0 \left[1 + \frac{\alpha}{2} + O(\alpha^2) \right] \quad (8.127)$$

$$\begin{aligned} \frac{1}{\tau} & = Z[-2 \operatorname{Im}(\Sigma)] = \frac{1}{\tau_0} \left(1 - \frac{\alpha}{2} \right) \left(1 + \frac{\alpha}{2} \right) \\ & = \frac{1}{\tau_0} [1 + O(\alpha^2)] \end{aligned} \quad (8.128)$$

The first correction terms in α from $Z = 1 - \alpha/2$ and $\operatorname{Im}[\Sigma] = 1 + \alpha/2$ cancel to order $O(\alpha)$. The one-phonon term provides no correction term to τ of $O(\alpha)$.

The electron mobility is evaluated by the same method which was used for impurity scattering. Starting from the Kubo formula for the electrical conductivity, the first important diagram is the bubble with interacting Green's functions shown in Fig. 8.7(a). The solid lines are total Green's functions $\mathcal{G}(p, ip_m)$ which include the self-energies

$$\pi(i\omega) = - \int_0^\beta d\tau e^{i\omega_n \tau} \langle T_\tau \mathbf{j}(\tau) \cdot \mathbf{j}(0) \rangle \quad (8.129)$$

$$\pi^{(0)}(i\omega) = \frac{2e^2}{3m_B^2} \int \frac{d^3p}{(2\pi)^3} p^2 \frac{1}{\beta} \sum_{ip_n} \mathcal{G}(p, ip_m) \mathcal{G}(\mathbf{p}, ip_m + i\omega_n) \quad (8.130)$$

The solid lines are total Green's functions $\mathcal{G}(p, ip_m)$ which include the self-energies. The evaluation of this term is identical to that used for the same bubble diagram for impurity

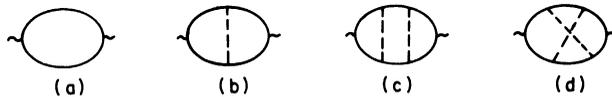


FIGURE 8.7

scattering. The contribution to the Kubo formula contains the two Green's functions as in Eqn. (8.44), which are evaluated by the standard series of steps to yield Eqn. (8.49):

$$\sigma^{(0)} = \frac{e^2}{3m_B^2} \int \frac{d^3p}{(2\pi)^3} p^2 \int \frac{d\varepsilon}{2\pi} A(p, \varepsilon)^2 \left[-\frac{dn_F(\varepsilon)}{d\varepsilon} \right] \quad (8.131)$$

In the present application, the self-energies in the Green's functions are evaluated from the Fröhlich Hamiltonian rather than impurity scattering. In a more realistic model of a solid, both self-energy expressions should be included: from phonons and impurity scattering.

An approximate evaluation of this contribution to the conductivity is obtained in the limit where $\alpha \ll 1$. The electron distribution is assumed to be Maxwell-Boltzmann:

$$n_F(\varepsilon) \approx e^{\beta(\mu-\varepsilon)} = \frac{n_0}{2} \left(\frac{2\pi\beta}{m^*} \right)^{3/2} e^{-\beta\varepsilon} \quad (8.132)$$

where $n_F(\varepsilon)$ is the energy distribution for each spin state, while n_0 is the total concentration for both spin states. This term in the conductivity is divided by $-en_0$ to get the corresponding term in the mobility of each electron:

$$\mu^{(0)} = -\frac{e\beta}{6m_B^2} \left(\frac{2\pi\beta}{m^*} \right)^{3/2} \int \frac{d^3p}{(2\pi)^3} p^2 \int \frac{d\varepsilon}{2\pi} A(p, \varepsilon)^2 e^{-\beta\varepsilon} \quad (8.133)$$

Using the expression (8.121) for the spectral function, the square of the spectral function is

$$A(p, \varepsilon)^2 = \frac{(Z/\tau)^2}{\{(\varepsilon - E_p)^2 + (2\tau)^{-2}\}^2} \approx A\pi\tau Z^2 \delta(\varepsilon - E_p) \quad (8.134)$$

The electron lifetime τ is inversely proportional to α , so that the mobility $\mu^{(0)}$ is inversely proportional to α . The preceding integral is evaluated in the limit where the temperature $T \rightarrow 0$. The integral is easy if it is assumed at low T that $Z(p)$ and $\tau(p)$ are evaluated at $p = 0$

$$\mu^{(0)} = -e\tau Z^2 \frac{m^*}{m_B^2} = -\frac{e\tau_0}{m_B} \left(1 - \frac{5}{6}\alpha + O(\alpha^2) \right) \quad (8.135)$$

where $Z(0) = 1 - \alpha/2$ and $m^*/m_B = 1 + \alpha/6$. The result (8.135) is the contribution from the simple bubble diagram of Fig. 8.8(a). This result does not resemble (8.108). The differences disappear when higher-order diagrams are considered, such as those in Fig. 8.7.

Other contributions to μ can be derived from the other diagrams, which are the vertex corrections. In the limits $T \rightarrow 0$ and $\alpha \ll 1$ the vertex corrections do not contribute to the mobility a term which goes as $O(1/\alpha)$. The simple bubble result is the final answer at low temperature and weak coupling. This conclusion is quite different from the situation for impurity scattering. There a series of ladder diagrams was summed in order to derive the final answer, and each ladder diagram gave a term which was the same inverse power of the coupling constant the impurity concentration n_i . The vertex diagrams for optical phonon scattering are less important because of the inelastic nature of the phonon scattering. Mahan (1966) showed that the two-phonon ladder diagrams in Fig. 8.7(c) provide the largest vertex corrections, which are of $O(T/\omega_0)$.

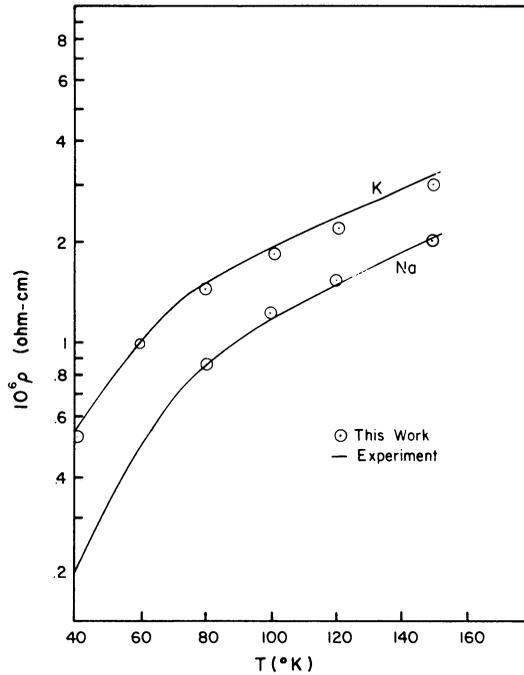


FIGURE 8.8 Resistivity as a function of temperature for Na and K. The solid line is experiment, and points are theory. *Source:* Dynes and Carbotte (1968) (used with permission).

8.3. ELECTRON-PHONON RELAXATION TIMES

This section is concerned with the calculation of electron relaxation times in semiconductors from the scattering by acoustical phonons. The prior sections have introduced two different relaxation times for the electron in a solid. One of these is the average time between scattering events and is denoted as $\tau(k)$. The mean-free-path (mfp) for this relaxation time is denoted as $l(k) = v_k \tau(k)$. The energy bands in the solid are assumed to be isotropic, so that the relaxation time and mfp depend only upon the magnitude of the wave vector. The other relaxation time τ_t enters into the electrical conductivity ($\sigma = n_0 e^2 \tau_t / m^*$). The equivalent quantity, before averaging over wave vector, is called the momentum relaxation time and is denoted with the subscript “t” since it is used in transport of current. The equivalent mfp is $l_t(k) = v_k \tau_t(k)$.

To further confuse the topic, the transport lifetimes τ_t for scattering from impurities had a factor of $(1 - \cos \theta')$ in the integrand of the scattering integral. The scattering from optical phonons did not have such a factor. The difference is that at low temperature, impurity scattering is elastic, while optical phonon is highly inelastic. Vertex corrections are relatively unimportant for inelastic scattering.

In writing the Boltzmann equation, the last term in the time development is the rate of change of the distribution function from collisions. The discussion of the lifetime from impurity scattering in Sec. 8.1.1 started with this collision term. Here the similar collision

term is presented from the scattering due to the electron–phonon interaction:

$$\begin{aligned} \left(\frac{\partial f(\mathbf{p})}{\partial t}\right)_{ep} &= \frac{2\pi}{\hbar} \int \frac{d^3q}{(2\pi)^3} |M_{\mathbf{q}}|^2 \{f(\mathbf{p})[1 - f(\mathbf{p} + \mathbf{q})] \\ &\quad \times [(N_{\mathbf{q}} + 1)\delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}+\mathbf{q}} - \hbar\omega_{\mathbf{q}}) + N_{\mathbf{q}}\delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}+\mathbf{q}} + \hbar\omega_{\mathbf{q}})] \\ &\quad - f(\mathbf{p} + \mathbf{q})[1 - f(\mathbf{p})][(N_{\mathbf{q}} + 1)\delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}+\mathbf{q}} + \hbar\omega_{\mathbf{q}}) + N_{\mathbf{q}}\delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}+\mathbf{q}} - \hbar\omega_{\mathbf{q}})]\} \end{aligned} \quad (8.136)$$

The integrand has four terms. The first two correspond to an electron initially in state \mathbf{p} scattering to $\mathbf{p} + \mathbf{q}$ by either the emission ($N_{\mathbf{q}} + 1$) or absorption ($N_{\mathbf{q}}$) of a phonon. This term is multiplied by the occupation factors $f(\mathbf{p})[1 - f(\mathbf{p} + \mathbf{q})]$ which ensures that the initial state \mathbf{p} is occupied and the final state is empty. The other two terms correspond to processes whereby electrons initially in $\mathbf{p} + \mathbf{q}$ scatter back into the state \mathbf{p} by either phonon emission or absorption. The back scattering has the occupation factors of $f(\mathbf{p} + \mathbf{q})[1 - f(\mathbf{p})]$. The above expression vanishes when the system is in thermal equilibrium and $f(\mathbf{p}) = n_F(\xi_{\mathbf{p}})$, $N_{\mathbf{q}} = n_B(\omega_{\mathbf{q}})$.

An immediate simplification of the above formula is attained by grouping together the terms with the same delta function for energy conservation

$$\begin{aligned} \left(\frac{\partial f(\mathbf{p})}{\partial t}\right)_{ep} &= \frac{2\pi}{\hbar} \int \frac{d^3q}{(2\pi)^3} |M_{\mathbf{q}}|^2 \\ &\quad \times [\delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}+\mathbf{q}} - \hbar\omega_{\mathbf{q}})\{f(\mathbf{p})[N_{\mathbf{q}} + 1 - f(\mathbf{p} + \mathbf{q})] - N_{\mathbf{q}}f(\mathbf{p} + \mathbf{q})\} \\ &\quad + \delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}+\mathbf{q}} + \hbar\omega_{\mathbf{q}})\{f(\mathbf{p})[N_{\mathbf{q}} + f(\mathbf{p} + \mathbf{q})] - f(\mathbf{p} + \mathbf{q})[N_{\mathbf{q}} + 1]\}] \end{aligned} \quad (8.137)$$

The above expression is quite general.

The terms which multiply $f(\mathbf{p})$ are those which define the relaxation time $1/\tau_p(\mathbf{p})$.

$$\begin{aligned} \frac{1}{\tau_p(k)} &= \frac{2\pi}{\hbar} \int \frac{d^3q}{(2\pi)^3} |M_{\mathbf{q}}|^2 \{\delta(\varepsilon_k - \varepsilon_{\mathbf{k}+\mathbf{q}} = \hbar\omega_{\mathbf{q}})[N_{\mathbf{q}} + 1 - n_F(\varepsilon_{\mathbf{k}+\mathbf{q}})] \\ &\quad + \delta(\varepsilon_k - \varepsilon_{\mathbf{k}+\mathbf{q}} + \hbar\omega_{\mathbf{q}})[N_{\mathbf{q}} + n_F(\varepsilon_{\mathbf{k}+\mathbf{q}})]\} \end{aligned} \quad (8.138)$$

$$\begin{aligned} \frac{1}{\tau_{pt}(k)} &= \frac{2\pi}{\hbar} \int \frac{d^3q}{(2\pi)^3} |M_{\mathbf{q}}|^2 \left(-\frac{\mathbf{q} \cdot \mathbf{k}}{k^2}\right) \\ &\quad \times \{\delta(\varepsilon_k - \varepsilon_{\mathbf{k}+\mathbf{q}} - \hbar\omega_{\mathbf{q}})[N_{\mathbf{q}} + 1 - n_F(\varepsilon_{\mathbf{k}+\mathbf{q}})] \\ &\quad + \delta(\varepsilon_k - \varepsilon_{\mathbf{k}+\mathbf{q}} + \hbar\omega_{\mathbf{q}})[N_{\mathbf{q}} + n_F(\varepsilon_{\mathbf{k}+\mathbf{q}})]\} \end{aligned} \quad (8.139)$$

$$|M_{\mathbf{q}}|^2 = D^2 \frac{\hbar q^2}{2\rho\omega_{\mathbf{q}}} \quad (8.140)$$

$$N_{\mathbf{q}} = n_B(\omega_{\mathbf{q}}) \quad (8.141)$$

The subscript “p” denotes scattering by phonons. In Eqn. (8.138) the relaxation time from phonon scattering is all of the terms which multiply $f(\mathbf{p})$ in (8.137). It is identical to the expression obtained from the imaginary part of the electron self-energy, as calculated in the one-phonon approximation. The occupation factors $f(\mathbf{p} + \mathbf{q})$, $N_{\mathbf{q}}$ are represented by their values in thermal equilibrium. The matrix element is from deformation potential interactions.

The transport lifetime τ_{pt} from the electron scattering by phonons is defined in (8.140). It has an additional factor of $(-\mathbf{q} \cdot \mathbf{k}/k^2)$ in the integrand. If the scattering were elastic, then this

factor equals $(1 - \cos \theta)$. The difference between the impurity scattering times τ_i , and τ_{it} , is the factor of $[1 - \cos(\theta)]$ in the integrand of $1/\tau_{it}$, where θ is the scattering angle $\hat{k} \cdot \hat{p} = \cos(\theta)$.

$$\frac{1}{\tau_i(k)} = \frac{2\pi n_i}{\hbar} \int \frac{d^3 p}{(2\pi)^3} |T_{\mathbf{k}\mathbf{p}}|^2 \delta(\varepsilon_k - \varepsilon_p) \quad (8.142)$$

$$\frac{1}{\tau_{it}(k)} = \frac{2\pi n_i}{\hbar} \int \frac{d^3 p}{(2\pi)^3} |T_{\mathbf{k}\mathbf{p}}|^2 [1 - \hat{k} \cdot \hat{p}] \delta(\varepsilon_k - \varepsilon_p) \quad (8.143)$$

The identical factor is used to calculate the momentum relaxation (τ_{pt}) from scattering by phonons. If $\mathbf{p} = \mathbf{k} + \mathbf{q}$ then the angular factor $[1 - \cos(\theta)] \approx -\mathbf{q} \cdot \mathbf{k}/k^2$. The identity is exact if $p = k$. Using this factor of $[1 - \cos \theta]$ in the impurity scattering is rigorously correct. Using this factor in the scattering by phonons is not rigorously correct. The difference is that the scattering by phonons is inelastic. Instead, the correct result is found by solving the Boltzmann equation, which is done in Sec. 8.4. However, the above approximation is actually quite good, and therefore is useful. The lifetimes for scattering by phonons are evaluated below.

Besides these two relaxation times, there are several others which are occasionally useful. The most important is the scattering by electron–electron interactions in a metal. The formulas for this case are derived in Chapter 11 in the discussion of Fermi liquid theory. Two others are the relaxation times for temperature, and the relaxation times for energy. These two are discussed below.

The relaxation time for energy determines the rate at which the electron loses or gains energy from the scattering. Since impurity scattering is elastic, the electron does not change its energy. There is no contribution to energy relaxation from impurity scattering. However, in the scattering by phonons the electron changes its energy by $\varepsilon_{\mathbf{k}+\mathbf{q}} - \varepsilon_{\mathbf{k}} = \pm \hbar \omega_{\mathbf{q}}$. One process gains energy ($+\hbar \omega_{\mathbf{q}}$) while the other loses energy ($-\hbar \omega_{\mathbf{q}}$). There is a net gain or loss of energy if one process dominates over the other. For example, if an energetic electron is injected into the solid, it will gradually lose energy until it equilibrates thermally. The process of coming to thermal equilibrium requires that it emit more phonons than it absorbs.

The rate of energy relaxation is calculated by starting from (8.138) and inserting the energy change $\varepsilon_{\mathbf{k}+\mathbf{q}} - \varepsilon_{\mathbf{k}}$ into the integrand

$$\begin{aligned} \left(\frac{dE}{dt} \right)_{pe} = & -\frac{2\pi}{\hbar} \int \frac{d^3 q}{(2\pi)^3} |M_q|^2 \hbar \omega_q \{ \delta(\varepsilon_k - \varepsilon_{\mathbf{k}+\mathbf{q}} - \hbar \omega_q) [N_q + 1 - n_F(\varepsilon_{\mathbf{k}+\mathbf{q}})] \\ & - \delta(\varepsilon_k - \varepsilon_{\mathbf{k}+\mathbf{q}} + \hbar \omega_q) [N_q + n_F(\varepsilon_{\mathbf{k}+\mathbf{q}})] \} \end{aligned} \quad (8.144)$$

Converting this expression to a relaxation time (τ_{pe}) and mfp (l_{pe}) is done below.

The evaluation of the relaxation times for the electron–phonon systems is done for metals and semiconductors. In metals, the evaluation assumes the existence of a Fermi surface. The electron–phonon scattering affects those few electrons within about $k_B T$ of the Fermi surface. In semiconductors, the interesting effects are for a single electron in the band, and the reference energy is the band edge. These two cases are different and are treated separately.

8.3.1. Metals

In metals the derivation is simplified by introducing the McMillan function denoted as $\alpha^2 F(\omega)$. In this function are collected all of the complicated parts of the phonon dispersion, and the matrix element (see Sec. 7.4)

$$\alpha^2 F(E, \omega) = \int \frac{d^3 q}{(2\pi)^3} |M_q|^2 \delta(\omega - \omega_q) \delta(E - \varepsilon_{\mathbf{k}+\mathbf{q}}) \quad (8.145)$$

$$\alpha_t^2(E, \omega) = \int \frac{d^3 q}{(2\pi)^3} |M_q|^2 \left(-\frac{\mathbf{q} \cdot \mathbf{k}}{k^2} \right) \delta(\omega - \omega_q) \delta(E - \varepsilon_{\mathbf{k}+\mathbf{q}}) \quad (8.146)$$

$$\begin{aligned} \frac{1}{\tau_{p(t)}(k)} &= \frac{2\pi}{\hbar} \int d\omega \{ \alpha_{(t)}^2 F(\varepsilon_k - \omega, \omega) [n_B(\omega) + n_F(\omega - \varepsilon_k)] \\ &\quad + \alpha_{(t)}^2 F(\varepsilon_k + \omega, \omega) [n_B(\omega) + n_F(\omega + \varepsilon_k)] \} \end{aligned} \quad (8.147)$$

The dimensionless functions $\alpha_{(t)}^2 F(E, \omega)$ depends upon the variable E . However, this dependence is similar to the variation in the density of states at the Fermi surface. The variation with E is usually smooth on the energy scale of the Debye energy, which is all that is relevant for interactions with phonons. The important variation is

$$\alpha_{(t)}^2 F(\varepsilon_k \pm \omega, \omega) = \alpha_{(t)}^2 F(\varepsilon_k, \omega) \pm \omega \frac{\partial}{\partial \varepsilon_k} \alpha_{(t)}^2 F(\varepsilon_k, \omega) + \dots \quad (8.148)$$

The second term is usually smaller than the first by a factor of $\hbar\omega/W$, where $\hbar\omega$ is a phonon energy while W is an electronic bandwidth. The usual approximation is to neglect the dependence upon the factor of E . The above expression is rewritten as

$$\frac{1}{\tau_{p(t)}(k)} = \frac{2\pi}{\hbar} \int_0^{\omega_D} d\omega \alpha_{(t)}^2 F(\omega) [2n_B(\omega) + n_F(\omega - \varepsilon_k) + n_F(\omega + \varepsilon_k)] \quad (8.149)$$

High temperature is defined as T greater than the Debye temperature. Most solids have Debye temperatures less than room temperature, so that $T = 300$ K is a high temperature. In this limit, the largest term in the bracket is $n_B(\omega) \approx k_B T / \hbar\omega$ and the lifetimes have the simple expression

$$\frac{\hbar}{\tau_p} = 2\pi\lambda k_B T \quad (8.150)$$

$$\frac{\hbar}{\tau_{pt}} = 2\pi\lambda_r k_B T \quad (8.151)$$

$$\lambda_{(t)} = 2 \int_0^{\omega_D} \frac{d\omega}{\omega} \alpha_{(t)}^2 F(\omega) \quad (8.152)$$

The inverse lifetime is proportional to temperature, and the constant of proportionality is the dimensionless electron–phonon coupling constant $\lambda_{(t)}$. The resistivity of metals [$\rho = m/(e^2 n_0 \tau_t)$] at high temperature is linear in T , and the slope is given by λ_r . A measurement of $\rho(T)$, gives λ_r , which can then be used to predict the transition temperature for superconductivity in that metal. This process works quite well.

The equation for energy relaxation is

$$\begin{aligned} \left(\frac{dE}{dt}\right)_{pe} &= 2\pi\hbar \int_0^{\omega_D} d\omega \omega \alpha^2 F(\omega) [1 - n_F(\varepsilon_k - \omega) - n_F(\varepsilon_k + \omega)] \\ &\approx [1 - 2n_F(\varepsilon_k)] \hbar \pi \lambda \langle \omega^2 \rangle \end{aligned} \quad (8.153)$$

$$\lambda \langle \omega^2 \rangle = 2 \int_0^{\omega_D} d\omega \omega \alpha^2 F(\omega) \quad (8.154)$$

The frequency integrals go from zero to the Debye frequency $\omega_D = k_B \Theta_D / \hbar$. The occupation numbers were expanded $n_F(\varepsilon_k \pm \omega) \approx n_F(\varepsilon_k) \pm \omega n'_F$. The first derivative terms cancel and the second derivative terms are small and are neglected. The factor $[1 - 2n_F(\varepsilon_k)]$ is one for electrons and minus one for holes. The energy relaxation always takes the particle to the chemical potential.

The energy relaxation (dE/dt) is governed by the quantity $\lambda \langle \omega^2 \rangle$. It is a single function, although the notation gives the impression that it is a product of two functions. The quantity $\hbar \lambda \langle \omega^2 \rangle$ has the units of Watts. In this case there is no obvious lifetime τ_{pe} nor mfp (l_{pe}). The phonons give a constant value to the energy relaxation as long as the electron's energy $\xi = \varepsilon_k - \mu > \hbar \omega_D$.

8.3.2. Semiconductors

The relaxation time for a semiconductor is calculated assuming that there is only one electron in the band. Most semiconductors have electrons or hole in one or several band minimum. In order to keep the discussion simple, the present calculation will be done assuming there is a single conduction band at the center of the Brillouin zone. This situation applies to GaAs and other III–V and II–VI semiconductors. The phonon wave vectors are rather small. An electron with energy $\varepsilon_k \approx k_B T$ has a small wave vector k . If it emits or absorbs a phonon and goes to energy $\varepsilon_{\mathbf{k}+\mathbf{q}} \pm \hbar \omega_q$, then $|\mathbf{k} + \mathbf{q}|$ is a small wave vector. The consequence is that q is also a small wave vector. At small wave vectors, it is a good approximation to represent acoustical phonons by the Debye model ($\omega_q = c_s q$) and optical phonons by an Einstein model ($\omega_q = \omega_0$). The results for acoustical phonons are derived here. The case of optical phonons is assigned as problems, although the derivation is similar to that for Fröhlich polarons in the last section.

The wave vector integrals can be done analytically since the Debye approximation is accurate. Define

$$\alpha^2 F^{(\pm)}(k, \omega) = \int \frac{d^3 q}{(2\pi)^3} |M_q|^2 \delta(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}+\mathbf{q}} \pm \hbar \omega_q) \delta(\omega - \omega_q) \quad (8.155)$$

$$= \frac{D^2 m^*}{8\pi^2 \hbar^2 \rho k c_s^4} \omega^2 |_{\omega_l < \omega < \omega_u} \quad (8.156)$$

$$\omega_u = 2c_s(k \pm k_s), \quad \omega_l = 2c_s(\pm k_s - k) \quad (8.157)$$

$$k_s = m^* c_s / \hbar \quad (8.158)$$

The lower limit ω_l can be set equal to zero since the factor of ω^2 makes this contribution negligible. The lifetime is

$$\frac{1}{\tau_p(k)} = 2\pi \int d\omega \{ \alpha^2 F^{(-)}(k, \omega) [n_B(\omega) + 1] + \alpha^2 F^{(+)}(k, \omega) n_B(\omega) \}$$

There are two interesting limits to this expression. The first is at zero temperature where the phonon occupation numbers $n_B(\omega) = 0$. The particle is assumed to be energetic so that $k \gg k_s$. In this case the answer is

$$\frac{1}{\tau_p(k)} = \gamma_p \varepsilon_k \quad (8.159)$$

$$\gamma_p = \frac{4 D^2 m^{*2}}{3\pi \hbar^4 \rho c_s^2} \quad (8.160)$$

The lifetime depends upon the value of kinetic energy. The second case is at high temperature, where $n_B \approx k_B T / \hbar \omega$ so the result is proportional to temperature. In this case the natural quantity is the mfp [$l_p(k) = v_k \tau_p(k)$]

$$\frac{1}{\tau_p(k)} = \frac{v_k}{l_p} \quad (8.161)$$

$$\frac{1}{l_p} = \frac{D^2 k_B T m^{*2}}{\pi \hbar^4 \rho c_s^2} \quad (8.162)$$

Another mfp for the electron is called l_ϕ . The symbol ϕ denotes the phase of the electron. The value of l_ϕ is the distance over which the electron travels before it breaks its phase coherence. For electrons in a pure semiconductor, where there are no electron–electron interactions, then $l_\phi = l_p$. The distance for phase coherence is given by the mfp for scattering by phonons. The phonon can carry away an arbitrary amount of phase, so such scattering does change the phase of the electron in a random fashion. In contrast, the scattering by impurities does not change the phase coherence. Impurity scattering changes the phase of the electron, but it changes the phase of each electron by the same amount. So coherence is maintained in scattering by impurities, while it is not maintained in the scattering by phonons. The impurities are just part of the one-electron potential which guides the electron as it wanders through the crystal. Electron–electron interactions also break phase coherence.

A calculation similar to finding l_p can be done for the momentum relaxation

$$\begin{aligned} \alpha_l^2 F^{(\pm)}(\omega) &= \int \frac{d^3 q}{(2\pi)^3} |M_q|^2 (-\mathbf{k} \cdot \mathbf{q} / k^2) \delta(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}+\mathbf{q}} \pm \hbar \omega_{\mathbf{q}}) \delta(\omega - \omega_{\mathbf{q}}) \\ &= \frac{D^2 m^*}{16\pi^2 \hbar^2 \rho k^3 c_s^6} \omega^4 |_{\omega_l < \omega < \omega_u} \\ \frac{1}{\tau_{pl}(k)} &= 2\pi \int_0^{\omega_D} d\omega \{ \alpha_l^2 F^{(-)}(\omega) [n_B(\omega) + 1] + \alpha_l^2 F^{(+)}(\omega) n_B(\omega) \} \end{aligned}$$

The results are at low temperature

$$\frac{1}{\tau_{pt}(k)} = \gamma_{pt} \varepsilon_k \quad (8.163)$$

$$\gamma_{pt} = \frac{8}{5\pi} \frac{D^2 m^{*2}}{\hbar^4 \rho c_s} = \frac{6}{5} \gamma_p \quad (8.164)$$

and at high temperature

$$\frac{1}{l_{pt}} = \frac{D^2 k_B T m^{*2}}{\pi \hbar^4 \rho c_s^2} = \frac{1}{l_p} \quad (8.165)$$

The results are the same at high temperatures for the transport and regular scattering mfp: $l_{pt} = l_p$. This equality is expected for the following reason. At high temperatures, the wave vector dependence of the interaction is effectively

$$\frac{|M_{\mathbf{q}}|^2}{\hbar \omega_{\mathbf{q}}} = \frac{D^2 q^2}{2\rho \omega_q^2} = \frac{D^2}{2\rho c_s^2} \quad (8.166)$$

The interaction is independent of wave vector, so the scattering is isotropic. The angular factor of $\cos(\theta)$ averages to zero for isotropic scattering, and the two mfp's are identical.

Some numerical results are presented in Table 8.1 for the high-temperature case using $T = 300$ K. The factor of $\rho c_s^2 = C_{11}$, where C_{11} is the elastic constant for LA phonons. The values for l_{pt} are several microns.

Also included are data for the polar scattering by optical phonons, which produces a lifetime τ_0 , which can be converted to a mfp using the thermal velocity $v_T = \sqrt{2k_B T/m^*}$.

Energy relaxation in semiconductors is discussed using the formula (8.144)

$$\left(\frac{dE}{dt}\right) = -2\pi \int_0^{\omega_D} d\omega \omega \{ \alpha^2 F^{(-)}(k, \omega) [n_B(\omega) + 1] - \alpha^2 F^{(+)}(k, \omega) n_B(\omega) \}$$

At low temperature the phonon occupation number $n_B = 0$ and the only process is phonon emission. Assuming that the particle has enough energy to emit phonons, the above formula gives

$$\left(\frac{dE}{dt}\right) = \frac{D^2 m^* k^3}{\pi \hbar^2 \rho} = -\frac{v_k \varepsilon_k}{l_{pe}} \quad (8.167)$$

$$\left(\frac{dE}{dx}\right) = \frac{1}{v_k} \left(\frac{dE}{dt}\right) = -\frac{\varepsilon_k}{l_{pe}} \quad (8.168)$$

$$\frac{1}{l_{pe}} = \frac{2D^2 m^{*3}}{\pi \hbar^4 \rho} \quad (8.169)$$

The mfp path for energy relaxation l_{pe} from phonons is quite long. Some values are shown in Table 8.1. Typical distances are millimeters. Most semiconductor devices are much smaller than these length scales, so that electrons do not achieve energy relaxation in traversing most devices unless there is sufficient impurities to cause them to random walk this distance.

The other interesting case is for high temperature. Expand $n_B = k_B T / \omega_q - \frac{1}{2}$ and find

$$\left(\frac{dE}{dt}\right) = -2\pi \int_0^{\omega_D} d\omega \omega \left\{ \frac{k_B T}{\hbar \omega} [\alpha^2 F^{(-)} - \alpha^2 F^{(+)}] + \frac{1}{2} [\alpha^2 F^{(+)} + \alpha^2 F^{(+)}] \right\}$$

TABLE 8.1 Electron mfp data at $T = 300$ K. The top lines have material constants. The second group of lines has estimated mpf from optical phonon scattering. The last two lines are from deformation potential scattering by LA phonons

| | Units | GaAs | InP | InAs | InSb |
|--------------------|-----------------|-------|-------|-------|-------|
| m^* | m_e | 0.064 | 0.078 | 0.027 | 0.013 |
| ϵ_0 | | 12.8 | 12.5 | 15.15 | 17.7 |
| ϵ_∞ | | 10.9 | 9.5 | 12.25 | 15.7 |
| C_{11} | GPa | 118 | 101 | 83 | 66.7 |
| ρ | g/cm^3 | 5.32 | 4.81 | 5.67 | 5.78 |
| $\hbar\omega_{LO}$ | meV | 35.4 | 42.8 | 29.6 | 23.6 |
| D | eV | 8.0 | 6.4 | 6.0 | 14.0 |
| α | | 0.067 | 0.125 | 0.055 | 0.020 |
| τ_0 | ps | 0.80 | 0.51 | 0.85 | 2.09 |
| v_T | km/s | 378 | 341 | 581 | 837 |
| $v_T\tau_0$ | μm | 0.30 | 0.18 | 0.50 | 1.75 |
| l_{pt} | μm | 2.0 | 1.8 | 13.9 | 8.8 |
| l_{pe} | mm | 3.2 | 2.5 | 80 | 134 |

The last term gives the same mfp l_{pe} as was found for zero temperature. The first term, which is proportional to $k_B T$, is nonzero when the limits to the frequency integrals are taken to be $2c_s(k \pm k_s)$, so the integral is

$$\int^{2c_s(k-k_s)} d\omega\omega^2 - \int^{2c_s(k+k_s)} d\omega\omega^2 = \quad (8.170)$$

$$-\frac{1}{3}(2c_s)^3[(k+k_s)^3 - (k-k_s)^3] = -16k^2 m^* c_s^4 \quad (8.171)$$

$$\left(\frac{dE}{dx}\right) = -\frac{1}{l_{pe}}[\epsilon_k - 2k_B T] \quad (8.172)$$

The energy relaxation has the same formula l_{pe} for the mfp at room temperature as at low temperature. The only change at higher temperature is that the energy ϵ_k relaxes to the value $2k_B T$, which was zero at zero temperature.

8.3.3. Temperature relaxation

The relaxation time for temperature is required when the electron temperature T_e and the phonon temperature T_p are different. The electron–phonon interaction will allow energy exchange between these two systems, and gradually bring them to the same temperature. The rate at which they equilibrate defines the temperature relaxation time. The result for metals was introduced by Allen (1987). The same quantity is important in semiconductors. There are several situations where an experimentalist might encounter $T_p \neq T_e$. Since the electrons absorb electromagnetic radiation, an intense laser pulse could raise T_e above T_p . Similarly, a strong dc electric field could accelerate electrons to have an average kinetic energy well above the phonon temperature, and then the two systems would mutually relax to the same temperature.

The starting expression for this calculation is the collision term (8.137) in the Boltzmann equation. The expression is simplified with the following assumptions:

- The electrons are in thermal equilibrium among themselves with an effective temperature T_e , $\beta_e = 1/k_B T_e$ and distribution function $f(\mathbf{p}) = n_F(\xi_{\mathbf{p}})$. This relative equilibrium is maintained by rapid electron–electron scattering.
- The phonons are in thermal equilibrium among themselves with an effective temperature T_p , $\beta_p = 1/k_B T_p$ which is maintained by rapid anharmonic phonon–phonon scattering. They are described by a distribution function $n_B(\omega_{\mathbf{q}})$.
- The electrons and phonons exchange energy according to (8.137).

Always keep in mind that the boson distribution function $n_B(\omega)$ is at a different temperature than the electron distribution functions $n_F(\varepsilon)$, $n_F(\varepsilon')$. The expression can be further simplified by using the function $\alpha^2 F(\omega)$ to express the integrals over the phonon states

$$\begin{aligned} \left(\frac{\partial f(\varepsilon)}{\partial t}\right)_{ep} &= 2\pi \int d\omega \alpha^2 F(\omega) \int d\varepsilon' \\ &\times \{ \delta(\varepsilon - \varepsilon' - \omega) [n_F(\varepsilon) [1 - n_F(\varepsilon')] [n_B(\omega) + 1] - n_F(\varepsilon') n_B(\omega) [1 - n_F(\varepsilon)]] \\ &+ \delta(\varepsilon - \varepsilon' + \omega) [n_F(\varepsilon) n_B(\omega) [1 - n_F(\varepsilon')] - n_F(\varepsilon') (n_B(\omega) + 1) [1 - n_F(\varepsilon)]] \} \end{aligned} \quad (8.173)$$

where $\varepsilon = \varepsilon_{\mathbf{p}}$, $\varepsilon' = \varepsilon_{\mathbf{p}+\mathbf{q}}$. Next, the change in the internal energy is calculated. The change in energy is accomplished by inserting under in the integral the factor of $\varepsilon' - \varepsilon = \pm \hbar\omega$.

$$\begin{aligned} \left(\frac{\partial E(\varepsilon)}{\partial t}\right)_{ep} &= -2\pi\hbar \int d\omega \alpha^2 F(\omega) \omega \int d\varepsilon' (8.174) \\ &\times \{ \delta(\varepsilon - \varepsilon' - \omega) [n_F(\varepsilon) [1 - n_F(\varepsilon')] [n_B(\omega) + 1] - n_F(\varepsilon') n_B(\omega) [1 - n_F(\varepsilon)]] \\ &- \delta(\varepsilon - \varepsilon' + \omega) [n_F(\varepsilon) n_B(\omega) [1 - n_F(\varepsilon')] - n_F(\varepsilon') (n_B(\omega) + 1) [1 - n_F(\varepsilon)]] \} \\ &= -2\pi\hbar \int d\omega \alpha^2 F(\omega) \omega \\ &\times \{ n_F(\varepsilon) [1 - n_F(\varepsilon - \omega)] [n_B(\omega) + 1] - n_F(\varepsilon - \omega) n_B(\omega) [1 - n_F(\varepsilon)] \\ &- n_F(\varepsilon) n_B(\omega) [1 - n_F(\varepsilon + \omega)] + n_F(\varepsilon + \omega) (n_B(\omega) + 1) [1 - n_F(\varepsilon)] \} \end{aligned} \quad (8.175)$$

In order to average over the entire system of electrons, it is necessary to also integrate over $d\varepsilon$. If $N(\varepsilon)$ is the density of states for the electron system, the energy change in the electron system is

$$\frac{dU}{dt} = C \frac{dT_e}{dt} = \int d\varepsilon N(\varepsilon) \left(\frac{\partial E(\varepsilon)}{\partial t}\right)_{ep} \quad (8.176)$$

where C is the heat capacity. The integrals over $d\varepsilon$ all converge within a thermal energy of the chemical potential. On this small energy scale, the density of states $N(\varepsilon)$ can be taken to be a constant $N(0)$, where the zero of energy is the chemical potential. Then all of the integrals over $d\varepsilon$ have the typical form

$$I(\omega) = \int d\varepsilon n_F(\varepsilon) [1 - n_F(\varepsilon - \omega)] = \frac{\omega}{e^{\beta_e \omega} - 1} = \omega n_{Be}(\omega) \quad (8.177)$$

The subscript “e” is added to the boson occupation function $n_{Be}(\omega)$ to emphasize that it is evaluated at the electron temperature T_e . The above result can be derived by changing variables of integration to $s = \exp(\beta_e \varepsilon)$, $ds = s\beta_e d\varepsilon$, $t = \exp(\beta_e \omega)$ which changes the integral to

$$I(\omega) = k_B T_e \int_0^\infty \frac{ds}{(s+1)(s+t)} = \frac{k_B T_e}{e^{\beta_e \omega} - 1} \ln \left[\frac{s+1}{s+t} \right]_0^\infty = \omega n_{Be}(\omega)$$

The other integrals over $d\varepsilon$ are evaluated in a similar fashion. The result for the energy relaxation is

$$C \frac{dT_e}{dt} = -4\pi\hbar^2 N(0) \int_0^{\omega_D} d\omega \omega^2 \alpha^2 F(\omega) [n_{Be}(\omega) - n_{Bp}(\omega)] \quad (8.178)$$

The final formula is quite simple. The right-hand side of this equation obviously vanishes in equilibrium when $T_e = T_p$. Another simplification occurs because the heat capacity is proportional to the density of states $C = \pi^2 k_B^2 T N(0)/3$ which simplifies the expression to

$$\frac{dT_e}{dt} = -\frac{12\hbar^2}{\pi k_B^2 T_e} \int_0^{\omega_D} d\omega \omega^2 \alpha^2 F(\omega) [n_{Be}(\omega) - n_{Bp}(\omega)] \quad (8.179)$$

The only factor which relates to the particular solid is $\alpha^2 F(\omega)$.

There are two interesting limits when evaluating this expression. The first is at very low temperature, $T_{e,p} \ll \Theta$, where the Debye temperature is Θ . Then the integral over ω has its main contribution at small values of ω . The small values of frequency come from sound waves, where it is suitable to use the Debye approximation. The evaluation of $\alpha^2 F(\omega)$ in this limit is identical to the semiconductor case, which gives that $\alpha^2 F = \gamma \omega^2$ where γ is a constant. The frequency limits can be extended to infinity, which gives the expression

$$\frac{dT_e}{dt} = -\Gamma [T_e^5 - T_p^5] \quad (8.180)$$

$$\Gamma = \frac{12k_B^3 \gamma}{\pi T_e \hbar^3} I_4 \quad (8.181)$$

$$I_4 = \int_0^\infty \frac{dx x^4}{e^x - 1} = 4! \zeta(5) \quad (8.182)$$

The relaxation obeys a T^5 law, which means very little heat is exchanged between electrons and phonons at low temperatures.

The other interesting limit is at room temperature, or at least above the Debye temperature. Then the boson occupation factors can be expanded $n_B \approx k_B T / \hbar \omega$ which gives the simple result

$$\frac{dT_e}{dt} = -\frac{(T_e - T_p)}{\tau_{pT}} \quad (8.183)$$

$$\frac{1}{\tau_{pT}} = \frac{6\hbar}{\pi k_B T_e} \lambda \langle \omega^2 \rangle \quad (8.184)$$

At high temperature the temperature relaxation is determined by the temperature difference $\delta T = T_e - T_p$, as well as by the function $\lambda \langle \omega^2 \rangle$. The latter function also determines the energy relaxation.

The temperature relaxation is measured by lasers using pulse-probe techniques. The first laser pulse is absorbed by the electrons, which thermally excites them, which causes their

temperature to differ from the phonons. The time-delayed probe measurement determines how some property, such as the refractive index, varies with the time interval after the initial pulse. An observation of an exponential relaxation curve is a measurement of τ_{pT} .

The above result is valid for a metal. A similar derivation can be done for a semiconductor, which is assigned in the problems. The same calculation for scattering of electrons in a semiconductor, by optical phonons, can be done in an interesting and rigorous way. For optical phonon scattering in a semiconductor, all of the phonons have the same energy $\hbar\omega_0$ so there is no need for an integral over frequency. In fact, just set

$$\alpha^2 F(\omega) = \frac{\lambda\omega_0}{2} \delta(\omega - \omega_0) \quad (8.185)$$

where λ is the dimensionless coupling constant. Starting from (8.173) gives

$$\begin{aligned} \left(\frac{\partial f(\varepsilon)}{\partial t} \right) &= -\pi\lambda\omega_0 \{ [f(\varepsilon)(N_0 + 1) - N_0 f(\varepsilon - \omega_0)] \Theta(\varepsilon - \omega_0) \\ &\quad + f(\varepsilon)N_0 - f(\varepsilon + \omega_0)(N_0 + 1) \} \end{aligned} \quad (8.186)$$

The states $f(\varepsilon + l\omega_0)$ can be regarded as a ladder of energy states, where each level is separated by ω_0 . The above equation can be solved exactly by matrix methods. Define $f_l = f(\varepsilon + l\omega_0)$ where the energy ε is now limited to the range $0 < \varepsilon < \omega_0$. The vector \tilde{f} has f_l as its elements. The above equation can now be cast into a matrix equation, where N_0 is the occupation number of the optical phonons.

$$\frac{d}{dt} \tilde{f} = -\frac{1}{\tau_{oT}} \mathcal{M} \cdot \tilde{f} \quad (8.187)$$

$$\mathcal{M} = \begin{pmatrix} N_0 & -(N_0 + 1) & 0 & \cdots \\ -N_0 & 2N_0 + 1 & -(N_0 + 1) & \cdots \\ 0 & -N_0 & 2N_0 + 1 & \cdots \end{pmatrix} \quad (8.188)$$

$$\frac{1}{\tau_{oT}} = \pi\lambda\omega_0 \quad (8.189)$$

The matrix \mathcal{M} has tridiagonal form. Only three rows are shown, but its dimensionality is infinite. All row except the first have elements: $-N_0$, $2N_0 + 1$, $-(N_0 + 1)$ to the left of the diagonal, on the diagonal, and to the right of the diagonal. If the matrix \mathcal{M} has eigenvalues ε_j and eigenfunctions ψ_j then the solution is

$$\tilde{f}(t) = \sum_j a_j \psi_j e^{-\varepsilon_j t / \tau_{oT}} \quad (8.190)$$

where the coefficients a_j are determined by the initial conditions. The equilibrium distribution is given by $\varepsilon_0 = 0$, $\psi_0 = \exp(-l\beta\omega_0)$, where β is determined by the phonon temperature. The system relaxes towards the equilibrium distribution, which is why its eigenvalue has to be zero. The other eigenvalues and eigenvectors are given by Mahan (1985).

8.4. ELECTRON–PHONON INTERACTIONS IN METALS

8.4.1. Force–force correlation function

In pure metals the electrical resistivity has two components. There is usually a constant resistivity from electron scattering by impurities, which is the largest part of the resistance at small temperatures. There is also a resistivity from electron scattering by phonons, which is temperature dependent and becomes large at high temperature. “Matthiessen’s rule” (1862) is that these two contributions to the resistance are additive. It should be regarded as a rule of thumb, rather than an ironclad rule. There are enough “deviations from Matthiessen’s rule” to make the abbreviation DMR a familiar acronym (Bass, 1982).

Electron scattering by acoustical phonons presents a hard problem in transport theory. The scattering is slightly inelastic. This problem is solved using neither the elastic scattering theory of Sec. 8.1 nor the inelastic scattering theory of Sec. 8.2. Instead an integral equation must be derived and solved for the energy dependence of the scattering process. The slightly inelastic nature of the scattering process makes this calculation much harder than the previous cases.

Two methods for obtaining the electrical conductivity are emphasized in this book. One uses equilibrium methods and evaluates the Kubo formula for the current–current correlation function. The resistivity from phonons will be found using this method, which follows the original derivation by Holstein (1964). The second method utilizes the quantum Boltzmann equation (QBE), which is a nonequilibrium theory. Mahan and Hansch (1983) used the QBE to derive the Holstein formula. Both of these derivations are complicated. They end by deriving the same integral equation for the scattering function, which must be solved by further work. Their virtue is that they are formally exact starting points, although approximations are made in obtaining the solution.

Other methods for obtaining the resistivity have been proposed, partly to avoid all of the work associated with the exact methods. These other methods are approximate. However, often the theories are both simple and accurate, which make them useful approximations. One of them is the force–force correlation function. If $\mathbf{F}(t)$ is the fluctuating force that acts on the electron, then define $R(i\omega)$ as the force–force correlation function:

$$R(i\omega) = -\frac{1}{3} \int_0^\beta d\tau e^{i\omega_n \tau} \langle T_\tau \mathbf{F}(\tau) \cdot \mathbf{F}(0) \rangle \quad (8.191)$$

$$\rho = \frac{1}{e^2 n_0^2} \lim_{\omega \rightarrow 0} \left[\frac{\text{Im } R_{\text{ret}}(\omega)}{\omega} \right] \quad (8.192)$$

This formula is just the quantum analogy of the Nyqvist theorem (Kubo *et al.*, 1985). After calculating this correlation function, the retarded function is obtained by letting $i\omega \rightarrow \omega + i\delta$. The resistance ρ is found by dividing by ω and taking the limit of $\omega \rightarrow 0$. For example, assume that the force on the electron has two terms: \mathbf{F}_i from impurities and \mathbf{F}_{ph} from phonons. If they are uncorrelated, then the correlation function has no cross terms. Symbolically write

$$R = \langle (\mathbf{F}_i + \mathbf{F}_{ph}) \cdot (\mathbf{F}_i + \mathbf{F}_{ph}) \rangle = \langle \mathbf{F}_i \cdot \mathbf{F}_i \rangle + \langle \mathbf{F}_{ph} \cdot \mathbf{F}_{ph} \rangle \quad (8.193)$$

In this case the resistivities from impurities and phonons are additive, in agreement with Matthiessen’s rule.

As an example, the resistivity is calculated from impurity scattering. The potential energy of the electron scattering from the impurities at \mathbf{R}_i is discussed in Sec. 4.1.5

$$V(\mathbf{r}) = \sum_i V_{ei}(\mathbf{r} - \mathbf{R}_i) = \frac{1}{v} \sum_{i\mathbf{q}} V(\mathbf{q}) \exp[i\mathbf{q} \cdot (\mathbf{r} - \mathbf{R}_i)] \quad (8.194)$$

The force \mathbf{F} is the gradient of the potential. The factor of $\exp(i\mathbf{q} \cdot \mathbf{r})$ can also be written as the electron density operator $\rho(\mathbf{q})$. The factor of $\exp(-i\mathbf{q} \cdot \mathbf{R}_i)$ can be written as the impurity density operator $\rho_i(-\mathbf{q})$

$$\mathbf{F}(\mathbf{r}) = -\frac{i}{v} \sum_{i\mathbf{q}} \mathbf{q} V(\mathbf{q}) \exp[i\mathbf{q} \cdot (\mathbf{r} - \mathbf{R}_i)] = -\frac{i}{v} \sum_{i\mathbf{q}} \mathbf{q} V(\mathbf{q}) \rho(\mathbf{q}) \rho_i(-\mathbf{q})$$

The next step is to evaluate the force–force correlation function $R(i\omega)$. In correlating \mathbf{F} with itself, there are two separate factors. One is $\langle \rho_i(-\mathbf{q}) \rho_i(-\mathbf{q}') \rangle$, which equals the number of impurities N_i if $\mathbf{q} = -\mathbf{q}'$. The other is the electron density–density correlation function, which is given exactly in terms of the inverse dielectric function

$$R_i(i\omega) = -\frac{N_i}{3v^2} \sum_{\mathbf{q}} \frac{q^2}{v_q} V(\mathbf{q})^2 \left[\frac{1}{\epsilon(\mathbf{q}, i\omega)} - 1 \right] \quad (8.195)$$

The next step is to take the imaginary part of the retarded function. The only retarded function on the right-hand side of the above equation is the inverse dielectric function $\epsilon = \epsilon_1 + i\epsilon_2$. Its imaginary part is $-\epsilon_2/(\epsilon_1^2 + \epsilon_2^2)$. At low frequency $\epsilon_2 = 2\omega e^2 m^2 / q^3$. Dividing by ω eliminates the factor of frequency. The formula for the resistivity from impurity scattering is

$$\rho = \frac{n_i m^2}{6\pi n_0^2 e^2} \int \frac{d^3 q}{(2\pi)^3} q \left| \frac{V(q)}{\epsilon(q)} \right|^2 \quad (8.196)$$

This formula is the exact result for the zero-temperature resistivity from impurity scattering, when the scattering is calculated in the second Born approximation. If $V(q)/\epsilon(q)$ is replaced by the T matrix for scattering, then it is the exact result, period. It is the formula $\rho = m/(n_0 e^2 \tau_i)$, where τ_i is defined in (8.25). It even includes the factor of $(1 - \cos \theta')$, although this assertion is not immediately obvious. One has to perform the angular integral in (8.25), which eliminates the delta function, in order to show its equivalence with the above formula for the resistivity.

The force–force correlation function gives the right resistivity for impurity scattering. No vertex correction or integral equation was needed in the derivation. The ease of derivation has made this approach popular.

Several caveats are needed. One is that impurity scattering is the only known example where the force–force correlation function gives the correct answer. In other cases it give an approximate answer. The second caveat is that the right answer is obtained by a wrong derivation. The derivation contains two important limits. One is setting the volume $v \rightarrow \infty$, while the second is $\omega \rightarrow 0$. The above answer is obtained by taking these limits in the wrong order. If they had been done correctly, in the right order, a different answer is obtained. Using the right order in evaluating the force–force correlation function gives a zero result as $\omega \rightarrow 0$. These points are discussed by Argyres and Sigel (1974), Huberman and Chester (1975), Kubo *et al.* (1985), and Fishman (1989).

The force–force correlation function may also be evaluated for the electron scattering by phonons. The result is

$$\rho(T) = C' \sum_{\lambda} \int q d^3 q |W(\mathbf{q})|^2 (\hat{\xi}_{\lambda} \cdot \mathbf{q})^2 \left[-\frac{dn_B(\omega)}{d\omega} \right] \quad (8.197)$$

$$C' = \frac{3\hbar v_0}{16Me^2 v_F^2 k_F^4} \quad (8.198)$$

Here $W(\mathbf{q})$ is the screened electron–ion interaction and v_0 is the unit cell volume. This formula was first derived by Ziman (1960) as a variational solution to the Boltzmann equation. It is the formula that is most often evaluated when calculating the temperature dependence of the resistivity of metals. Figure 8.8 shows a theoretical calculation of Dynes and Carbotte (1968) compared with experiments for Na and K. An important feature of these calculations is numerically integrating over the Brillouin zone for all the phonon states, while employing accurate values for the phonon frequencies $\omega_{\lambda}(\mathbf{q})$ and polarization vectors $\hat{\xi}_{\lambda}$.

8.4.2. Kubo formula

A rigorous theory of the dc electrical conductivity is calculated for the scattering of electrons by all phonons: acoustical and optical. The starting point is the Kubo formula, and the derivation follows the original one by Holstein (1964). His theory sums the ladder diagrams for phonons and reduces the vertex function to an integral equation which is solved numerically. So far a solution is available only for a spherical Fermi surface, so that the result is the Kubo formula analogy of (8.198). The results are expressed in terms of the McMillan function $\alpha^2 F(\omega)$ and a similar function $\alpha'^2 F(\omega)$, which is used in transport theory.

The goal is to evaluate the current–current correlation function in the presence of the electron–phonon interaction. This correlation function can always be expressed as a product of two Green’s functions and the vertex function. The Green’s function $\mathcal{G}(\mathbf{p}, ip)$ used here represents fully interacting particles, with a self-energy found from the electron–phonon interaction plus any additional interactions of interest. The Feynman diagram for the correlation function is shown in Fig. 8.9 where the vertex function is put only at one end of the bubble in order not to overcount the vertex terms:

$$\begin{aligned} \pi(i\omega) &= -\frac{1}{3v} \int_0^{\beta} d\tau e^{i\omega_n \tau} \langle T_{\tau} \mathbf{j}(\tau) \cdot \mathbf{j}(0) \rangle \\ &= -\frac{e^2}{3m^2 v} \sum_{\mathbf{p}\mathbf{p}'\sigma\sigma'} \mathbf{p} \cdot \mathbf{p}' \int_0^{\beta} d\tau e^{i\omega_n \tau} \langle T_{\tau} C_{\mathbf{p}\sigma}^{\dagger}(\tau) C_{\mathbf{p}\sigma}(\tau) C_{\mathbf{p}'\sigma'}^{\dagger}(0) C_{\mathbf{p}'\sigma'}(0) \rangle \end{aligned} \quad (8.199)$$

$$\pi(i\omega) = \frac{2e^2}{3m^2} \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\beta} \sum_{ip} \mathcal{G}(\mathbf{p}, ip) \mathcal{G}(\mathbf{p}, ip + i\omega) \mathbf{p} \cdot \mathbf{\Gamma}(\mathbf{p}; ip, ip + i\omega) \quad (8.200)$$

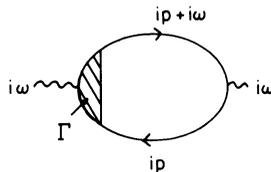


FIGURE 8.9

The dc conductivity is found by the same steps used in Secs. 8.1 and 8.2. One evaluates the correlation function for values of $i\omega$ and analytically continues $i\omega \rightarrow \omega + i\delta$ to find the retarded function. The dc conductivity is the imaginary part of the retarded function divided by ω , in the limit where $\omega \rightarrow 0$.

The vertex function $\Gamma(\mathbf{p}; ip, ip + i\omega)$ is evaluated below. It depends on both the frequency variables ip and $i\omega$, which is written in the combination $(ip, ip + i\omega)$. The two arguments ip and $ip + i\omega$ come from the electron Green's functions which have the same frequency arguments. In a homogeneous electron gas, the vector vertex function $\Gamma(\mathbf{p}; ip, ip + i\omega)$ must point in the vector direction \mathbf{p} , although in real metals the crystalline potential defines other possible directions. However, the approximation of treating the Fermi surface as strictly spherical is equivalent to neglecting crystal directions, so assume Γ points in the direction \mathbf{p} . It is convenient to introduce the scalar function $\gamma(\mathbf{p}; ip, ip + i\omega)$, which is the amplitude of the vector vertex function:

$$\Gamma(\mathbf{p}; ip, ip + i\omega) = \mathbf{p}\gamma(\mathbf{p}; ip, ip + i\omega) \quad (8.201)$$

$$\pi(i\omega) = \frac{2e^2}{3m^2} \int \frac{d^3p}{(2\pi)^3} p^2 S(\mathbf{p}, i\omega) \quad (8.202)$$

$$S(\mathbf{p}, i\omega) = \frac{1}{\beta} \sum_{ip} \mathcal{G}(\mathbf{p}, ip) \mathcal{G}(\mathbf{p}, ip + i\omega) \gamma(\mathbf{p}; ip, ip + i\omega) \quad (8.203)$$

The scalar function γ is not the same as the scalar vertex function in (8.81).

The next step is to evaluate the summation over Matsubara frequencies ip to obtain $S(\mathbf{p}, i\omega)$. To this end, construct the usual contour integral which has cuts along the axes where $ip \rightarrow \text{real}$ and also $ip + i\omega \rightarrow \text{real}$. These series of algebraic operations are the same as used to derive (8.64) for impurity scattering. That earlier result can be used here by identifying $P = GG\gamma$

$$\begin{aligned} \text{Im}[S_{\text{ret}}(\mathbf{p}, \omega)] = \text{Re} \left\{ \int_{-\infty}^{\infty} \frac{d\varepsilon}{2\pi} [n_F(\varepsilon + \omega) - n_F(\varepsilon)] \right. \\ \times [G_{\text{adv}}(\mathbf{p}, \varepsilon) G_{\text{ret}}(\mathbf{p}, \varepsilon + \omega) \gamma(\mathbf{p}; \varepsilon - i\delta, \varepsilon + \omega + i\delta) \\ \left. - G_{\text{ret}}(\mathbf{p}, \varepsilon) G_{\text{ret}}(\mathbf{p}, \varepsilon + \omega) \gamma(\mathbf{p}; \varepsilon + i\delta, \varepsilon + \omega + i\delta)] \right\} \quad (8.204) \end{aligned}$$

and the conductivity is

$$\begin{aligned} \sigma = \frac{2e^2}{3m^2} \int \frac{d^3p}{(2\pi)^3} p^2 \int_{-\infty}^{\infty} \frac{d\varepsilon}{2\pi} \left[-\frac{dn_F(\varepsilon)}{d\varepsilon} \right] \\ \times \{ |G_{\text{ret}}(\mathbf{p}, \varepsilon)|^2 \gamma(\mathbf{p}, \varepsilon - i\delta, \varepsilon + i\delta) - \text{Re}[G_{\text{ret}}(\mathbf{p}, \varepsilon)^2 \gamma(\mathbf{p}, \varepsilon + i\delta, \varepsilon + i\delta)] \} \quad (8.205) \end{aligned}$$

Equation (8.205) is exact for the exact vertex function. It is expressed in terms of the two functions $\gamma(\mathbf{p}; \varepsilon - i\delta, \varepsilon + i\delta)$ and $\gamma(\mathbf{p}; \varepsilon + i\delta, \varepsilon + i\delta)$. These two functions are expected to be quite different, as they were in Sec. 8.1 for impurity scattering. The function $\gamma(\mathbf{p}; \varepsilon + i\delta, \varepsilon + i\delta)$ could be obtained from a Ward identity. That is true here, as first shown

for the electron–phonon system in metals by Englesberg and Schrieffer (1963). This Ward identity can be expressed in terms of $\Gamma(\mathbf{p}, \varepsilon + i\delta, \varepsilon + i\delta)$ or the scalar $\gamma(\mathbf{p}, \varepsilon + i\delta, \varepsilon + i\delta)$:

$$\Gamma(\mathbf{p}; \varepsilon + i\delta, \varepsilon + \omega + i\delta) = \mathbf{p} + m\mathbf{V}_p \Sigma(\mathbf{p}, \varepsilon) \quad (8.206)$$

$$\gamma(\mathbf{p}; \varepsilon + i\delta, \varepsilon + \omega + i\delta) = 1 + \frac{\partial}{\partial \xi_p} \Sigma(\mathbf{p}, \varepsilon) \quad (8.207)$$

$$G_{\text{ret}}(\mathbf{p}, \varepsilon)^2 \gamma(\mathbf{p}; \varepsilon + i\delta, \varepsilon + \omega + i\delta) = \frac{\partial}{\partial \xi_p} G_{\text{ret}}(\mathbf{p}, \varepsilon) \quad (8.208)$$

For metals, the electron–phonon system has the feature that the self-energy function $\Sigma(\mathbf{p}, \varepsilon)$ is not very \mathbf{p} dependent, and the derivative of $\Sigma(\mathbf{p}, \varepsilon)$ with respect to ξ_p is small: $d\Sigma/d\xi_p \approx \Sigma/E_F \ll 1$. A good approximation is to set $\gamma(\mathbf{p}, \varepsilon + i\delta, \varepsilon + i\delta) = 1$.

There is another result which is even stronger, since it applies for all values of ε near the Fermi energy. Since the self-energy $\Sigma(\mathbf{p}, \varepsilon)$ does not have significant \mathbf{p} dependence near $p \sim k_F$, call it $\Sigma(k_F, \varepsilon) = \text{Re}[\Sigma(\varepsilon)] - i\Gamma(\varepsilon)$ where the p notation is suppressed. The function $\Gamma(\varepsilon) = -\text{Im}[\Sigma(k_F, \varepsilon)]$. The retarded and advanced Green's function have their important dependence on p through the kinetic energy term, $\xi \equiv \xi_p$. Define $\Omega(\varepsilon) = \varepsilon - \text{Re}[\Sigma(\varepsilon)]$:

$$G_{\text{ret}}(p, \varepsilon) = \frac{1}{\varepsilon - \xi - \text{Re}[\Sigma(\varepsilon)] + i\Gamma(\varepsilon)} = \frac{1}{\Omega(\varepsilon) - \xi + i\Gamma(\varepsilon)} \quad (8.209)$$

$$G_{\text{adv}}(p, \varepsilon) = \frac{1}{\varepsilon - \xi - \text{Re}[\Sigma(\varepsilon)] - i\Gamma(\varepsilon)} = \frac{1}{\Omega(\varepsilon) - \xi - i\Gamma(\varepsilon)} \quad (8.210)$$

Similarly, the vertex function $\gamma(\mathbf{p}; \varepsilon - i\delta, \varepsilon + \omega + i\delta)$ is assumed to not have a significant dependence upon p except on the order of k_F . This assumption will be justified later. The second term in brackets in (8.205) can be neglected; it vanishes by doing the kinetic energy integration. By neglecting terms of order $O(\xi/E_F)$ then

$$d^3 p p^2 = 4\pi p^4 dp = 4\pi m k_F^3 d\xi [1 + O(\xi/E_F)] = 12\pi m n_0 d\xi \quad (8.211)$$

where the electron density is $n_0 = k_F^3/3\pi^2$. When the kinetic energy integral is evaluated, the only ξ variation is in the Green's functions. Of the three combinations which occur, only one makes a nonzero contribution:

$$\int_{-\infty}^{\infty} d\xi G_{\text{ret}}(p, \varepsilon) G_{\text{ret}}(p, \varepsilon') = 0 \quad (8.212)$$

$$\int_{-\infty}^{\infty} d\xi G_{\text{ret}}(p, \varepsilon) G_{\text{adv}}(p, \varepsilon') = \frac{2\pi i}{\Omega(\varepsilon) - \Omega(\varepsilon') + i[\Gamma(\varepsilon) + \Gamma(\varepsilon')]} \quad (8.213)$$

$$\int_{-\infty}^{\infty} d\xi G_{\text{adv}}(p, \varepsilon) G_{\text{adv}}(p, \varepsilon') = 0 \quad (8.214)$$

Each integral is evaluated by closing the contour at infinity. The two integrals which vanish have both their poles in the same half plane (upper or lower), so the integration contour can be chosen to avoid them both, which encircles no poles and hence gives zero. The integral over the combination $G_{\text{adv}} G_{\text{ret}}$ has one pole in each plane. Closing the contour always picks up one

pole whose residue produces the result. By using these integration results, the dc conductivity (8.205) becomes

$$\sigma = \frac{e^2 n_0}{2m} \int_{-\infty}^{\infty} d\varepsilon \frac{\Lambda(\varepsilon)}{\Gamma(\varepsilon)} \left[-\frac{d}{d\varepsilon} n_F(\varepsilon) \right] \tag{8.215}$$

$$\Lambda(\varepsilon) = \gamma(k_F, \varepsilon - i\delta, \varepsilon + i\delta) \tag{8.216}$$

$$\Gamma(\varepsilon) = -2 \operatorname{Im}[\Sigma(k_F, \varepsilon)] \tag{8.217}$$

Equation (8.215) is the final result of the formal derivation. There only remains the evaluation of the vertex function $\Lambda(\varepsilon)$ and the imaginary self-energy $\Gamma(\varepsilon)$. If the vertex function were absent ($\Lambda = 1$), then the evaluation would be easy. The imaginary self-energy $\Gamma(\varepsilon)$ has been evaluated for many metals, and a result was given in Fig. 7.14 for Pb. The quantity $1/2\Gamma = \tau(\varepsilon)$, where $\tau(\varepsilon)$ is the relaxation time defined as the average time between scattering events. The result (8.215) when $\Lambda = 1$, is just the average of the relaxation time over the thermally smeared Fermi distribution. Of course, the earlier solution for impurity scattering showed that the neglect of $\Lambda(\varepsilon)$ is a serious error. The vertex function $\Lambda(\varepsilon)$ serves the important role of weighting the scattering events and favoring those at high momentum transfer. For impurity scattering, $\Lambda = \Gamma/\Gamma_i$ or $\Lambda/\Gamma = 1/\Gamma_i$, where Γ_i is the scattering rate which contains the equivalent of $(1 - \cos \theta')$.

The vertex equation for phonon scattering is now solved to derive Γ_i . Since the phonon scattering is inelastic, the results are not identical to those for impurity scattering. In summing the ladder diagrams for the vertex function, however, the vertex contributions are found to be important and significantly different from unity. Migdal's (1958) theorem, which asserts that vertex terms are unimportant, is contradicted.

The vertex function $\Gamma(\mathbf{p}, ip, ip + i\omega)$ is calculated by solving the integral equation

$$\Gamma(\mathbf{p}, ip, ip + i\omega) = \mathbf{p} + \frac{1}{\beta} \sum_{iq,\lambda} \int \frac{d^3q}{(2\pi)^3} M_\lambda(\mathbf{q})^2 \mathcal{D}(\mathbf{q}, iq) \mathcal{G}(\mathbf{p} + \mathbf{q}, ip + iq) + \mathcal{G}(\mathbf{p} + \mathbf{q}, ip + iq + i\omega) \Gamma(\mathbf{p} + \mathbf{q}, ip + iq, ip + iq + i\omega) \tag{8.128}$$

This vertex sums the ladder diagrams for phonons. It is illustrated in Fig. 8.10. Iteration of Eqn. (8.218) produces a series in which each additional term has one more ladder diagram. The solution to the integral equation produces an expression which contains all terms with any number of phonon ladder diagrams. This solution is not an exact evaluation of the vertex function $\Gamma(\mathbf{p}, ip, ip + i\omega)$, since other vertex contributions occur which are not ladders but have the phonon lines crossed. One expects these terms to be smaller, but detailed calculations are lacking, so this is only a supposition.

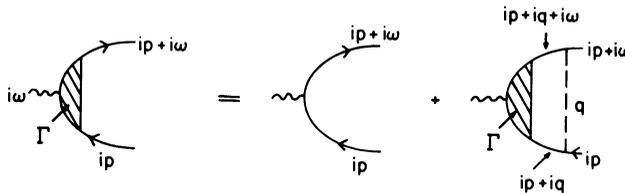


FIGURE 8.10

It is unfortunate that the vertex function is a vector. The scalar vertex function obtained by replacing \mathbf{p} by 1 in Eqn. (8.218) is easily obtained from a Ward identity:

$$\begin{aligned}\Gamma(\mathbf{p}, ip, ip + i\omega) &= 1 + \frac{1}{\beta} \sum_{iq,\lambda} \int \frac{d^3q}{(2\pi)^3} M_\lambda(\mathbf{q})^2 \mathcal{D}(\mathbf{q}, iq) \mathcal{G}(\mathbf{p} + \mathbf{q}, ip + iq) \\ &\quad \times \mathcal{G}(\mathbf{p} + \mathbf{q}, ip + iq + i\omega) \Gamma(\mathbf{p} + \mathbf{q}, ip + iq, ip + iq + i\omega) \\ &= 1 - \frac{1}{i\omega} [\Sigma(\mathbf{p}, ip + i\omega) - \Sigma(\mathbf{p}, ip)]\end{aligned}\quad (8.219)$$

The Ward identity is not helpful for our problem with a vector vertex function. The integral equation must be attacked and solved.

Define the scalar function $\gamma(\mathbf{p}, ip, ip + i\omega)$ as the scalar amplitude of the vector vertex function $\mathbf{\Gamma} = \mathbf{p}\gamma$. It obeys the integral equation

$$\begin{aligned}\gamma(\mathbf{p}, ip, ip + i\omega) &= 1 + \frac{1}{\beta} \sum_{iq,\lambda} \int \frac{d^3q}{(2\pi)^3} M_\lambda(\mathbf{q})^2 \mathcal{D}(q) \frac{\mathbf{p} \cdot (\mathbf{p} + \mathbf{q})}{p^2} \\ &\quad \times \mathcal{G}(p + q) \mathcal{G}(p + q + i\omega) \mathbf{\Gamma}(\mathbf{p} + \mathbf{q}, ip + iq, ip + iq + i\omega)\end{aligned}$$

This equation is not the same one which is obeyed by the scalar vertex function in (8.219), and these two functions are quite different.

First do the integrals over angles and wave vector. First write $d^3q = 2\pi q^2 dq dv$, where $v = \cos\theta$ is the angle between \mathbf{p} and \mathbf{q} . The angle variable is changed to p_1 , defined as

$$p_1^2 = (\mathbf{p} + \mathbf{q})^2 = p^2 + q^2 + 2pqv \quad (8.220)$$

$$dv = \frac{p_1 dp_1}{pq} \quad (8.211)$$

$$\int d^3q = \frac{2\pi}{p} \int_0^\infty q dq \int_{|p-q|}^{|p+q|} p_1 dp_1 = \frac{2\pi}{v_0} \int_0^\infty q dq \int_{\xi_{p-q}}^{\xi_{p+q}} d\xi_1 \quad (8.222)$$

Next uncouple the limits of integration. Only electrons at the Fermi surface contribute, so that values of $\xi_1 = \varepsilon_{p_1} - \mu \approx 0$ and $\xi_p \approx 0$ are important in the integration process. Approximate $v_p = v_F$ as the Fermi velocity. Similarly, the integral $q dq$ is understood to be over the spherical Fermi surface from one point \mathbf{p} to all other points \mathbf{p}_1 , where p and p_1 both have magnitude k_F :

$$\int d^3q \rightarrow \frac{1}{v_F} \int d^2q \int_{-\infty}^\infty d\xi_1 \quad (8.223)$$

The limits on the $d\xi_1$ integral are extended between $\pm\infty$, since most of the integrand has large q values, where the actual limits on ξ_1 are from a very negative number to a very positive one. Since the main contribution is in the region $\xi_1 \approx 0$, this error is small. Consider the other angular factors in the integrand:

$$\frac{\mathbf{p} \cdot (\mathbf{p} + \mathbf{q})}{p^2} = 1 + \frac{qv}{p} = 1 + \frac{1}{2p^2} [p_1^2 - p^2 - q^2] \quad (8.224)$$

$$\approx 1 - \frac{q^2}{2k_F^2} + O\left(\frac{\xi_1}{E_F}, \frac{\xi_p}{E_F}\right) \quad (8.225)$$

The terms ξ_1/E_F , ξ_p/E_F are neglected, since it is expected that ξ_1 and ξ_p will be small—on the order of a Debye energy. On the other hand, the factor $q^2/2k_F^2$ need not be small, since the

integration over phonon states has a significant contribution from high values of q near the edge of the Brillouin zone. This term must be retained. The difference in treating the factor $(p_1^2 - p^2)/k_F^2$, which is neglected, and q^2/k_F^2 , which is retained, is that the former enters the average over electron states and the latter over phonon states. The integrals over the phonon wave vector can be expressed in terms of the function $\alpha^2 F(\omega)$, which was introduced in Sec. 7.4:

$$\alpha^2 F(\omega) = \frac{1}{v_F} \sum_{\lambda} \int \frac{d^3 q}{(2\pi)^3} M_{\lambda}(\mathbf{q})^2 \delta[\omega - \omega_{\lambda}(\mathbf{q})] \quad (8.226)$$

$$\alpha_i^2 F(\omega) = \frac{1}{v_F} \sum_{\lambda} \frac{d^3 q}{(2\pi)^3} M_{\lambda}(\mathbf{q})^2 \frac{q^2}{2k_F^2} \delta[\omega - \omega_{\lambda}(\mathbf{q})] \quad (8.227)$$

The first of these is just the McMillan function, which was defined in Sec. 7.4. Since the Fermi surface has been assumed to be spherical, $\alpha^2 F$ has the same value at each point on the surface, and the \mathbf{p} subscript is omitted. The other form of coupling $\alpha_i^2 F$ is called the “transport form of alpha-squared- F ,” which was introduced by Allen (1971). It differs from the McMillan form by having the additional factor of $q^2/2k_F^2$ in the integrand, which gives more weight to the scattering processes at large wave vector. The factor of $q^2/2k_F^2$ is identical to $(1 - \cos \theta')$ when the scattering is elastic.

After completing all these angular and wave vector integrations, the vertex function $\gamma(\mathbf{p}, ip, ip + i\omega)$ is found to be not very dependent on p . The only variation is as a function of p , which can be set equal to k_F with an error of only ξ_p/E_F . Define $\gamma(ip, ip + i\omega) = \gamma(k_F, ip, ip + i\omega)$ and arrive at the equations:

$$\begin{aligned} \gamma(ip, ip + i\omega) &= 1 + \int_0^{\omega_D} du [\alpha^2 F(u) - \alpha_i^2 F(u)] \\ &\quad \times \int_{-\infty}^{\infty} d\xi_1 S(\xi_1, u; ip, ip + i\omega) \end{aligned} \quad (8.228)$$

$$\begin{aligned} S(\xi, u; ip, ip + i\omega) &= \frac{1}{\beta} \sum_{iq} \frac{2u}{(iq)^2 - u^2} \gamma(ip + iq, ip + iq + i\omega) \\ &\quad \times \mathcal{G}(\xi, ip + iq) \mathcal{G}(\xi, ip + iq + i\omega) \end{aligned} \quad (8.229)$$

This integral equation is not too difficult to solve. The basic approximation has been to decouple the integrations over dq and $d\xi_1$, which permits all the phonon information to be collected into the functions $\alpha^2 F$, $\alpha_i^2 F$. The primary assumption in this decoupling is that the Fermi degeneracy E_F is very much larger than other energies such as $k_B T$ or $\hbar\omega_D$. The present integral equation is actually much easier to solve than the one for polarons in Sec. 8.2, since there the integration variables cannot be accurately decoupled in the same way.

The preceding is the basic integral equation which needs to be solved for the vertex function. The ξ dependence of S is only in the Green's functions $\mathcal{G}(\xi, ip + iq)$ and $\mathcal{G}(\xi, ip + iq + i\omega)$. This integral is done later, using the result (8.213) that only the integral over the pair $G_{\text{ret}} G_{\text{adv}} \equiv |G_{\text{ret}}|^2$ is nonzero. The factor $2u/[(iq)^2 - u^2]$ is the phonon Green's function for a phonon of energy u . The next step in the derivation is to do the summation over Matsubara frequency iq , which is done in the usual way by constructing a contour integral.

$$\oint \frac{dz}{2\pi i} \frac{2u}{z^2 - u^2} n_B(z) \gamma(z + ip, z + ip + i\omega) G(z + ip) G(ip + i\omega)$$

The contour in integration is a circle at infinity. The integrand has poles from the phonon Green's functions at $z = \pm u$ and cuts along the axes where the electron Green's functions are real, $z = -ip + \varepsilon'$ and $z = -ip - i\omega + \varepsilon'$. The contour integral from these three contributions are

$$S = S_1 + S_2 + S_3 \quad (8.230)$$

$$S_1 = n_B(u)\gamma(ip + u, ip + i\omega + u)G(ip + u)G(ip + i\omega + u) \\ + [n_B(u) + 1]\gamma(ip - u, ip + i\omega - u)G(ip - u)G(ip + i\omega - u) \quad (8.231)$$

$$S_2 = \int \frac{d\varepsilon'}{2\pi i} \frac{2u}{(\varepsilon' - ip)^2 - u^2} G(\varepsilon' + i\omega)[\gamma(\varepsilon' + i\delta, \varepsilon' + i\omega)G_{\text{ret}}(\varepsilon') \\ - \gamma(\varepsilon' - i\delta, \varepsilon' + i\omega)G_{\text{adv}}(\varepsilon')] \quad (8.232)$$

$$S_3 = \int \frac{d\varepsilon'}{2\pi i} \frac{2u}{(\varepsilon' - ip - i\omega)^2 - u^2} G(\varepsilon' - i\omega)[\gamma(\varepsilon' - i\omega, \varepsilon' + i\delta)G_{\text{ret}}(\varepsilon') \\ - \gamma(\varepsilon' - i\omega, \varepsilon' - i\delta)G_{\text{adv}}(\varepsilon')] \quad (8.233)$$

The next step is to perform the analytic continuations. The final result should be $\Lambda(\varepsilon) = \gamma(\varepsilon - i\delta, \varepsilon + i\delta)$. First set $ip \rightarrow \varepsilon - i\delta$. The next step is to set $ip + i\omega \rightarrow \varepsilon + i\omega \rightarrow \varepsilon + \omega + i\delta \rightarrow \varepsilon + i\delta$, since $\omega \rightarrow 0$. The Green's functions become advanced or retarded according to the side of the cut, so this analytical continuation produces

$$S = n_B(u)\Lambda(\varepsilon + u)|G_{\text{ret}}(\xi, \varepsilon + u)|^2 + [n_B(u) + 1]\Lambda(\varepsilon - u)|G_{\text{ret}}(\xi, \varepsilon - u)|^2 \\ + \int \frac{d\varepsilon'}{2\pi i} n_F(\varepsilon') \left[|G_{\text{ret}}(\xi, \varepsilon')|^2 \Lambda(\varepsilon') \left(\frac{2u}{\varepsilon' - \varepsilon - i\delta)^2 - u^2} - \frac{2u}{(\varepsilon' - \varepsilon + i\delta)^2 - u^2} \right) \right. \\ \left. + O(G_{\text{ret}}^2, G_{\text{adv}}^2) \right] \quad (8.234)$$

The next step is to do the integration over $d\xi_1$. According to (8.213), this integral eliminates all combinations of the Green's functions except $G_{\text{ret}}G_{\text{adv}} \equiv |G_{\text{ret}}|^2$, which removes the terms $O(G_{\text{ret}}^2, G_{\text{adv}}^2)$. Also note that the factor below is the phonon spectral function

$$\frac{2u}{(\varepsilon' - \varepsilon - i\delta)^2 - u^2} - \frac{2u}{(\varepsilon' - \varepsilon + i\delta)^2 - u^2} = 2\pi[\delta(\varepsilon' - \varepsilon - u) - \delta(\varepsilon' - \varepsilon + u)]$$

which eliminates the integral over $d\varepsilon'$. These manipulations provide the final form of the integral equation for $\Lambda(\varepsilon)$. It was first derived by Holstein (1964), although his result is modified by expressing it in the α^2F formalism:

$$\Lambda(\varepsilon) = 1 + \pi \int_0^{\omega_D} du [\alpha^2 F(u) - \alpha_i^2 F(u)] \\ \times \left[[n_B(u) + n_F(\varepsilon + u)] \frac{\Lambda(\varepsilon + u)}{\Gamma(\varepsilon + u)} + [n_B(u) + 1 - n_F(\varepsilon - u)] \frac{\Lambda(\varepsilon - u)}{\Gamma(\varepsilon - u)} \right] \quad (8.235) \\ \sigma = \frac{n_0 e^2}{2m} \int_{-\infty}^{\infty} d\varepsilon \frac{\Lambda(\varepsilon)}{\Gamma(\varepsilon)} \left[\frac{dn_F(\varepsilon)}{d\varepsilon} \right]$$

The integral (8.215) for the dc conductivity was rewritten in order to present the two important results together. The integral equation for the vertex function $\Lambda(\varepsilon)$ must be solved, and the solution is used in the integral over ε for the conductivity. Recall that $\Lambda(\varepsilon)$ is a real

function, so the integral equation is not complicated. The form of the equations suggests the unknown quantity is actually $\bar{\tau}(\varepsilon) = 2\Lambda/\Gamma$, which might be called the effective relaxation time for transport.

The energy uncertainty $\Gamma(\varepsilon)$ of the electron from the electron–phonon interaction was given previously in (7.305). The equivalent result for the transport kernel is:

$$\begin{aligned}\Gamma(\varepsilon) &= -\text{Im}[\Sigma] = \pi \int_0^{\omega_D} du \alpha^2 F(u) [2n_B(u) + n_F(u + \varepsilon) + n_F(u - \varepsilon)] \\ \Gamma_t(\varepsilon) &= \pi \int_0^{\omega_D} du \alpha^2 F_t(u) [2n_B(u) + n_F(\varepsilon + u) + n_F(u - \varepsilon)]\end{aligned}\quad (8.236)$$

If the self-energy function Γ and the vertex function $\Lambda(\varepsilon)$ were both constants and independent of ε , the solution of the vertex equation (8.235) would be easy:

$$\Lambda = 1 + \frac{\Lambda}{\Gamma} [\Gamma - \Gamma_t] \quad (8.237)$$

$$= \frac{\Gamma}{\Gamma_t} \quad (8.238)$$

This model gives $\Lambda/\Gamma = 1/\Gamma_t$, so that the conductivity integral has only the transport form of the relaxation time. The earlier solution to impurity scattering also assumed that $\Lambda/\Gamma = 1/\Gamma_t$. For impurity scattering the functions $\Gamma(\varepsilon)$ and $\Lambda(\varepsilon)$ are usually insensitive to ε near $\varepsilon \approx 0$; the exception is when the impurity has a scattering resonance near the Fermi surface or if the density of states is not smooth. However, for the electron–phonon interaction in metals, it is not a good approximation to treat $\Gamma(\varepsilon)$, $\Gamma_t(\varepsilon)$ or $\Lambda(\varepsilon)$ as constants. The calculated results in Sec. 7.4, shown in Fig. 7.14 for Pb, illustrate that $\Gamma(\varepsilon)$ has substantial energy variations near the Fermi energy. The vertex function does also.

Takegahara and Wang (1977) evaluated (8.235) and (8.215) for metallic rubidium and cesium. Their results are shown in Fig. 8.11. In each case the solid line is calculated assuming that the ratio Γ/Γ_t is a constant, while the dashed line is calculated by solving (8.235) for the ε dependence of $\Lambda(\varepsilon)$. The latter curve is in very good agreement with the experiments, which are indicated by the points. The differences between the solid and dashed curves are similar to the differences between the Ziman formula (8.198) and the Holstein formula (8.235). Note that there is no region with a well-defined T^5 law for the resistivity in these metals.

8.4.3. Mass enhancement

The electron–phonon mass enhancement factor λ was introduced in Sec. 7.4. It is from the real part of the electron self-energy due to the electron–phonon interaction. Since this self-energy was found to be energy dependent but not very wave vector dependent, the electron effective mass m^* is approximated by

$$\lambda(\omega) = -\frac{\partial}{\partial \omega} \Sigma(k_F, \omega) \quad (8.239)$$

$$\frac{m^*}{m} = 1 + \lambda(\omega) = \frac{1}{Z(\omega)} \quad (8.240)$$

The mass enhancement factor is also related to the quasiparticle renormalization factor $Z(\omega)$. The values of these quantities at the Fermi energy $\omega = 0$ are λ and Z_F . The mass renor-

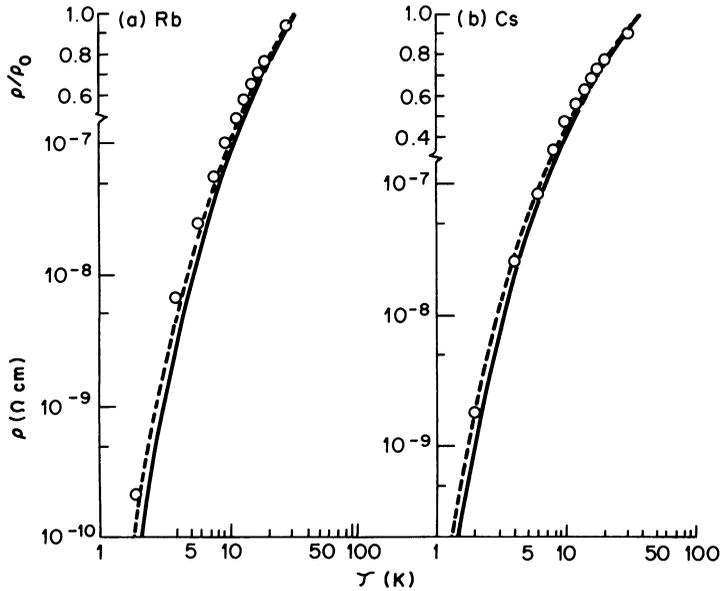


FIGURE 8.11 Calculated constant volume phonon limited electrical resistivity for (a) Rb and (b) Cs. Solid lines are calculations that treat $\Lambda(\varepsilon)$ as a constant. Dashed lines are full solution to integral equation for $\Lambda(\varepsilon)$. Points are experimental values. *Source:* Takegahara and Wang (1977) (used with permission).

malization factor λ can also be expressed as an average over the Fermi surface of a weighted average over the phonon density of states. This definition is given in Sec. 7.4. Values of λ in real metals range from 0.1 to 3.

An important question is the role which λ plays in the dc transport properties. The most obvious approximation is to use the effective mass $m^* = m(1 + \lambda)$ in transport formulas whenever classical theory says to use m . This substitution would make sense, since m^* is the effective mass which governs the motion of electrons on the Fermi surface, and these are involved in dc transport properties. However, this sensible procedure is wrong in most cases. The important point is that the factor $1 + \lambda$ can enter the final formulas in several ways. It changes the effective mass, relaxation time, and quasiparticle renormalization factor Z_F . The formula for the transport coefficient will have a number of factors of $1 + \lambda$. Often they all cancel, which is the case for the electrical conductivity.

Prange and Kadanoff (1964) investigated which transport measurements were influenced by the electron–phonon mass enhancement factor $1 + \lambda$. They concluded that the enhancement did affect the following measurements: specific heat, low-field cyclotron resonance, and the amplitude of the deHaas–van Alphen effect. The following measurements are not affected: dc electrical conductivity, thermoelectric power, thermal conductivity, the period of the deHaas–van Alphen effect, spin susceptibility, and the electron tunneling rate. Their conclusion on the thermoelectric power was challenged by Opsal *et al.* (1976), who detected a dependence on $1 + \lambda$. The list of quantities which are affected is much shorter than the list of quantities which are not affected. The usual case is that the transport property is not influenced by the mass enhancement factor; see Grimvall (1981) for a further discussion.

8.4.4. Thermoelectric power

The thermoelectric power is a subject which is seldom discussed in most solid-state courses, as the lecturer is busy treating subjects which are more fashionable. This tendency is regrettable, since it is an important measurement. Experiments show that diverse behaviors are found for simple metals, and even the sign of this quantity shows no regularity. The low-temperature theory in metals is poorly understood, although there has been extensive work.

The thermoelectric power was introduced briefly in Sec. 3.8. It is a simple measurement, at least conceptually. A conducting bar is insulated so electrical currents cannot exit from its ends, and then a temperature difference ΔT is maintained along the length of the bar. The two ends of the bar are found to have a voltage difference ΔV which is proportional to ΔT . The constant of proportionality is the thermoelectric power S :

$$S = -\frac{\Delta V}{\Delta T} = -\frac{1}{T} \frac{L^{(12)}}{L^{(11)}} \quad (8.241)$$

$$L^{(11)} = \frac{k_B T}{3i\omega v} \int_0^\beta d\tau e^{i\omega_n \tau} \langle T_\tau \mathbf{j}(\tau) \cdot \mathbf{j}(0) \rangle \quad (8.242)$$

$$L^{(12)} = \frac{k_B T}{3i\omega v} \int_0^\beta d\tau e^{i\omega_n \tau} \langle T_\tau \mathbf{j}_Q(\tau) \cdot \mathbf{j}(0) \rangle \quad (8.243)$$

The thermopower is defined theoretically as the ratio of two correlation functions. One is just proportional to the dc electrical conductivity $\sigma = \beta L^{(11)}$, which has already been evaluated for several models. The other correlation function has the argument of heat current \mathbf{j}_Q and the electrical current \mathbf{j} , where $\mathbf{j}_Q = \mathbf{j}_E - (\mu/e)\mathbf{j}$. Equation (8.241) differs by a factor of charge e from its earlier version in Sec. 3.9, since earlier \mathbf{j} meant the particle current, whereas now it is the electrical current. They differ by the unit of charge, which causes the change in (8.241).

The correlation function $L^{(12)}$ is now evaluated for its dc value. Both correlation functions are evaluated for values of $i\omega$; then analytically continue $i\omega \rightarrow \omega + i\delta$. Finally take the limit $\omega \rightarrow 0$ of the imaginary part of the retarded function.

In the evaluation of most correlation functions, there is usually a leading term which provides the dominant part of the answer. There are numerous small correction terms which can usually be ignored. In calculating the correlation function $L^{(12)}$ for the thermoelectric power, the dominant term vanishes, and one is left with obtaining all the numerous small correction terms. This feature makes it difficult to obtain an accurate answer.

The heat current operator has many terms. The one which is expected to provide the dominant term is from the kinetic energy of the electron: $\mathbf{j}_Q = \sum \mathbf{v}_p \xi_p C_{p\sigma}^\dagger C_{p\sigma}$. This heat current operator is used in most theories. Similarly, for the electrical current the operator is $\mathbf{j} = e \sum v_p C_{p\sigma}^\dagger C_{p\sigma}$. The correlation function is called $L^{(12a)}$:

$$L^{(12a)} = \frac{1}{v\beta i\omega} \frac{e}{3m^2} \sum_{\mathbf{k}p\sigma} \mathbf{k} \cdot \mathbf{p} \xi_p \int_0^\beta d\tau e^{i\omega_n \tau} \langle T_\tau C_{\mathbf{k}\sigma}^\dagger(\tau) C_{\mathbf{k}\sigma}(\tau) C_{\mathbf{p}\sigma}^\dagger C_{\mathbf{p}\sigma} \rangle$$

The important feature of this correlation function is the bracket containing four electron operators. This type of operator sequence was encountered in earlier sections of this chapter. It is evaluated as

$$L^{(12a)} = \frac{1}{v\beta^2 i\omega} \frac{2e}{3m^2} \sum_{\mathbf{k}} \xi_{\mathbf{k}} \sum_{ik} \mathbf{k} \cdot \Gamma(\mathbf{k}, ik, ik + i\omega) \mathcal{G}(k) \mathcal{G}(k + i\omega)$$

where $\Gamma(\mathbf{k}, ik, ik + i\omega)$ is the vector vertex function of the bubble diagram. This same vertex function enters into the correlation function for the conductivity. Write it again as $\Gamma = \mathbf{k}\gamma$. There is the same summation over Matsubara frequency ik which was done in the previous section for the correlation function $L^{(11)}$. The result was given in (8.204). The retarded form of this correlation function is

$$\begin{aligned} \text{Im}[L_{\text{ret}}^{(12a)}] &= -\frac{2e}{3\beta m^2 \omega} \int \frac{d^3 k}{(2\pi)^3} k^2 \xi_{\mathbf{k}} \int_{-\infty}^{\infty} \frac{d\varepsilon}{2\pi} [n_F(\varepsilon + \omega) - n_F(\varepsilon)] \\ &\quad \times \{G_{\text{adv}}(\mathbf{k}, \varepsilon) G_{\text{ret}}(\mathbf{k}, \varepsilon + \omega) \gamma(\mathbf{k}, \varepsilon - i\delta, \varepsilon + \omega + i\delta) \\ &\quad - \text{Re}[G_{\text{ret}}(\mathbf{k}, \varepsilon) G_{\text{ret}}(\mathbf{k}, \varepsilon + \omega) \gamma(\mathbf{k}, \varepsilon + i\delta, \varepsilon + \omega + i\delta)]\} \end{aligned}$$

The next step is to take the limit $\omega \rightarrow 0$, which causes the electron occupation factors to become $dn_F(\varepsilon)/d\varepsilon$. The integral over wave vector is changed to an integral over $k^2 d^3 k = 12\pi^3 m n_0 d\xi_{\mathbf{k}}$:

$$\begin{aligned} \text{Im}[L_{\text{ret}}^{(12a)}] &= \frac{en_0}{\beta m} \int_{-\infty}^{\infty} \frac{d\varepsilon}{2\pi} \left[-\frac{dn_F(\varepsilon)}{d\varepsilon} \right] \int_{-\infty}^{\infty} d\xi_{\mathbf{k}} \xi_{\mathbf{k}} \{ |G_{\text{ret}}(\mathbf{k}, \varepsilon)|^2 \\ &\quad \times \gamma(\mathbf{k}, \varepsilon) - i\delta, \varepsilon + i\delta) - \text{Re}[G_{\text{ret}}(\mathbf{k}, \varepsilon)^2 \gamma(\mathbf{k}, \varepsilon + i\delta, \varepsilon + i\delta)] \} \end{aligned} \quad (8.244)$$

Again it is assumed that the retarded and advanced Green's functions are significantly dependent only on ξ in their kinetic energy term and that the self-energy and vertex functions have negligible dependence on ξ . Then the following integrals are evaluated by a contour integration in analogy with (8.213):

$$\begin{aligned} \int_{-\infty}^{\infty} d\xi_{\mathbf{k}} \xi_{\mathbf{k}} G_{\text{ret}}(\xi, \varepsilon)^2 &= i\pi \\ \int_{-\infty}^{\infty} d\xi_{\mathbf{k}} \xi_{\mathbf{k}} |G_{\text{ret}}(\xi, \varepsilon)|^2 &= \frac{\pi \Omega(\varepsilon)}{\Gamma(\varepsilon)} \\ \int_{-\infty}^{\infty} d\xi_{\mathbf{k}} \xi_{\mathbf{k}} G_{\text{adv}}(\xi, \varepsilon)^2 &= -i\pi \end{aligned} \quad (8.245)$$

Two of the integrals equal $\pm i\pi$, which comes from the semicircle closing the contour at infinity. These integrals give zero when taking the real part. The important contribution must arise from the combination $|G_{\text{ret}}(\xi, \varepsilon)|^2$, since it has a nontrivial contribution from the integral over $d\xi_{\mathbf{k}}$:

$$\text{Im}[L_{\text{ret}}^{(12a)}] = \frac{en_0}{2\beta m} \int_{-\infty}^{\infty} d\varepsilon \left[-\frac{dn_F(\varepsilon)}{d\varepsilon} \right] \frac{\Omega(\varepsilon)\Lambda(\varepsilon)}{\Gamma(\varepsilon)} = 0 \quad (8.246)$$

This integral is zero after one evaluates the integral $d\varepsilon$. It vanishes because the integrand is an antisymmetric function of ε . The quantities $dn_F(\varepsilon)/d\varepsilon$, $\Lambda(\varepsilon)$, and $\Gamma(\varepsilon)$ are all symmetric functions of ε , while $\Omega(\varepsilon)$ is an antisymmetric function. The correlation function $L^{(12a)}$ is zero.

The integral vanishes because of the single power of ξ in the $d\xi$ integral. It makes the single power of $\Omega(\varepsilon)$ in the integral over $d\varepsilon$. The single power of ξ comes from the heat current operator.

A nonzero result for the correlation function is obtained by repeating this derivation and retaining all the correction terms. One important term is from the argument of the wave vector integration:

$$\begin{aligned} k^2 d^3 k &= 2\pi k^3 dk^2 = 3\pi m d\xi_k [k_F^2 + (k^2 - k_F^2)]^{3/2} \\ &\approx 4\pi k_F^3 d\xi_k \left(1 + \frac{3\xi k}{2E_F}\right) \end{aligned} \quad (8.247)$$

The second term in parentheses makes a nonzero contribution to the thermopower. This term is now evaluated. Kinetic energy integrals such as (8.245) must now be done, except there is a factor of ξ^2 in the integrand which multiplies the Green's functions. The evaluation of the integrand is tricky, since technically the integral diverges. At large values of ξ , then $G \sim 1/\xi$ and $(\xi G)^2 \sim 1$. The integrand does not fall off at large values of ξ , and taking the limits to $\pm\infty$ gives an infinite integral. Usually this problem is solved by ignoring it. The product $|G_{\text{ret}}|^2 = A/2\Gamma \approx \pi\delta(\xi - \varepsilon)/\Gamma$ is replaced by a delta function for energy conservation. The integral over $(\xi G_{\text{ret}})^2$ is set equal to zero:

$$\int d\xi \xi^2 |G_{\text{ret}}(\xi, \varepsilon)|^2 \approx \frac{\pi \varepsilon^2}{\Gamma(\varepsilon)} \quad (8.248)$$

The term in ε^2 is thermally averaged according to (8.244) using G&R 3.531(3);

$$\int_{-\infty}^{\infty} d\varepsilon \varepsilon^2 \left[-\frac{dn_F(\varepsilon)}{d\varepsilon} \right] = \frac{\pi^2}{3} \quad (8.249)$$

$$\text{Im}[L_{\text{ret}}^{(12a)}] = \frac{\pi^2 e n_0 \tau_t}{3\beta m E_F} \eta(T) \quad (8.250)$$

$$S = -\left(\frac{k_B}{\sigma}\right) \text{Im}[L_{\text{ret}}^{(12a)}] = -\frac{\pi^2 k_B^2 T}{3e E_F} \eta(T) \quad (8.251)$$

where the conductivity is $\sigma = n_0 e^2 \tau_t / m$. The dimensions of S are volts per degree. The parameter $\eta(T)$ is dimensionless. So far $\eta = \frac{3}{2}$ from the coefficient of ξ/E_F in (8.247). Other contributions to $\eta(T)$ are derived below.

Taylor and MacDonald (1986) evaluated this expression for the alkali metals at high temperature. Rather good agreement is obtained, as shown below. First it is necessary to find more contributions to $\eta(T)$. At high temperature it is a good approximation to set $\tau_t = \tau_{pt}$, where the transport lifetime is entirely from phonons, and is given by the transport form of alpha-squared- f in (8.226):

$$\frac{1}{\tau_{pt}(k)} = \frac{m}{(2\pi)^2 k^3} \int_0^{2k} q^3 dq \int_0^{2\pi} d\phi \sum_{\lambda} M_{\lambda}^2(\mathbf{q}) \{n_B[\omega_{\lambda}(\mathbf{q})] + n_F[\omega_{\lambda}(\mathbf{q})]\}$$

Contributions to $\eta(T)$ are obtained by expanding k about k_F and keeping the first-order terms in $(k - k_F)/k_F \approx \xi/2E_F$. The prefactor of k^{-3} gives another contribution of $\eta = \frac{3}{2}$ to η . The integration limit of $2k$ gives a contribution to η of

$$2q(T) = \frac{2m}{\pi^2} \tau_{pt} \int_0^{2\pi} d\phi \sum_{\lambda} M_{\lambda}^2(2k_F) \{n_B[\omega_{\lambda}(2k_F)] + n_F[\omega_{\lambda}(2k_F)]\}$$

Another contribution to $\eta(T)$ comes from the matrix element $M_{\lambda}(\mathbf{q})$. It is usually calculated using a screened pseudopotential for the electron-ion interaction. The better pseudopotentials are nonlocal, which means they depend upon $M_{\lambda}(\mathbf{k}, \mathbf{k} + \mathbf{q})$ rather than just on \mathbf{q} . This k

TABLE 8.2 Thermoelectric parameters in the alkali metals.
 $\eta = 3 - 2q - r/2$ (Taylor and MacDonald, 1986, used with permission).

| Metal | $T(K)$ | $2q(T)$ | $r(T)/2$ | η | |
|-------|--------|---------|----------|--------|------------|
| | | | | Theory | Experiment |
| Li | 424 | 9.26 | -1.43 | -5.33 | -6.3 |
| Na | 300 | 0.04 | -0.09 | 3.05 | 2.9 |
| K | 200 | 0.83 | -1.87 | 4.04 | 4.0 |
| Rb | 100 | 4.78 | -4.49 | 2.71 | 2.8 |
| Cs | 100 | 9.32 | -7.15 | 0.83 | 0.0 |

dependence can also be expanded around the point k_F . The pseudopotential gives another contribution to $\eta(T)$ which is called $r(T)$. The contributions to $\eta(T)$ are

$$\eta(T) = 3 - 2q(T) - \frac{1}{2}r(T) \tag{8.252}$$

Table 8.2 shows the evaluation of these terms for the alkali metals at various temperatures. The comparison with the experimental data is good for Na, K, and Rb. The high-temperature thermopower seems to be understood in these cases. Both $q(T)$ and $r(T)$ are small for Na because the electron-ion pseudopotential is nearly zero at $q = 2k_F$. The thermopower is easier to evaluate at high rather than at low temperature. At high temperature the resistance is dominated by the electron-phonon interaction, which is well approximated by using the transport form of alpha-squared- f . At low temperatures one also has to include the k dependence of impurity scattering, as well as the ordinary form of alpha-squared- f while solving (8.235). In addition, phonon drag is important at low temperature.

8.5. QUANTUM BOLTZMANN EQUATION

There are several different methods of doing transport theory. The theory used in the preceding sections uses the Kubo relation for the conductivity and is called “linear response.” One assumes that currents are proportional to fields. The proportionality constants can be evaluated in equilibrium. This method works because one assumes that the applied fields are small, and the system is only infinitesimally disturbed from equilibrium.

A second method of transport theory is discussed in this section. One assumes the existence of a distribution function f , which describes the behavior of the particles. One writes a differential equation for the motion of f through phase space. The differential equation is a Boltzmann equation. One then tries to solve the Boltzmann equation for a system out of equilibrium. For fields that are small, the system is only slightly out of equilibrium, and one reproduces the linear response solutions described earlier. The advantage of the Boltzmann equation method is that one can also try to solve the equation when the system is far from equilibrium.

The original Boltzmann equation described the behavior of a distribution function $f(\mathbf{v}, \mathbf{R}, t)$ of three variables: velocity, position, and time. The Wigner distribution function (WDF) was introduced in Sec. 3.7. It is equivalent to a distribution function $f(\mathbf{k}, \omega; \mathbf{R}, t)$ with four variables: wave vector \mathbf{k} , energy ω , position \mathbf{R} , and time t . This latter distribution function is the one needed for many-particle systems. Since $f(\mathbf{k}, \omega; \mathbf{R}, t)$ is not positive

definite, calling it a “distribution function” is probably misleading. This phrase is used since it is widespread, but the warning should be kept in mind.

The differential equation obeyed by $f(\mathbf{k}, \omega; \mathbf{R}, t)$ is called the *quantum Boltzmann equation*, which is abbreviated QBE. It is derived rigorously in the following sections. Here a quick derivation is provided using semiclassical arguments. According to the Liouville theorem a distribution function $f(q_i, t)$ is stationary when it obeys the equation

$$\delta f = 0 = \frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial q_i} \dot{q}_i + \left(\frac{\partial f}{\partial t} \right)_s \quad (8.253)$$

where the last term is from scattering. For the WDF, $\dot{\mathbf{k}} = \mathbf{F}$ (force), $\dot{\mathbf{R}} = \mathbf{v}$ (velocity), and $\dot{\omega} = \mathbf{v} \cdot \mathbf{F}$ (Joule heating). The QBE is

$$0 = \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{R}} + \mathbf{F} \cdot \nabla_{\mathbf{k}} + \mathbf{v} \cdot \mathbf{F} \frac{\partial}{\partial \omega} \right) f + \left(\frac{\partial f}{\partial t} \right)_2 \quad (8.254)$$

This equation has nearly the right form. It is derived rigorously in the next section, which produces a few more terms from self-energy contributions. That derivation also provides a prescription for obtaining the scattering term.

Equation (8.254) has one feature that is important. The additional variable ω also causes a new driving term on the left of the form $\mathbf{v} \cdot \mathbf{F} \partial / \partial \omega$. This term was first derived by Mahan and Hansch (1983). The semiclassical distribution function $f(\mathbf{v}, \mathbf{R}, t)$ lacks this driving term since it lacks the energy variable ω . The QBE for the WDF is a different equation from the traditional Boltzmann equation.

8.5.1. Derivation of the QBE

The QBE is the equation of motion for the Green's function $G^<$. The method of deriving transport equations was pioneered by Kadanoff and Baym (1962). Recall from Sec. 3.7 that $f(\mathbf{r}, t; \mathbf{R}, T) = -iG^<(\mathbf{r}, t; \mathbf{R}, T)$, where (\mathbf{r}, t) are the relative variables and (\mathbf{R}, T) are the position and time in center-of-mass. The Green's function was defined in terms of the field operator as

$$G^<(\mathbf{r}, t; \mathbf{R}, T) = i \langle \psi^\dagger(\mathbf{R} - \frac{1}{2}\mathbf{r}, T - \frac{1}{2}t) \psi(\mathbf{R} + \frac{1}{2}\mathbf{r}, T + \frac{1}{2}t) \rangle \quad (8.255)$$

The next step is to Fourier transform the relative variables (\mathbf{r}, t) into (\mathbf{k}, ω) :

$$G^<(\mathbf{k}, \omega; \mathbf{R}, T) = \int d^3r e^{-i\mathbf{k} \cdot \mathbf{r}} \int dt e^{i\omega t} G^<(\mathbf{r}, t; \mathbf{R}, T) \quad (8.256)$$

The QBE will be derived for a particle in a weak electric field. The intent is to describe interacting many-particle systems that have a small current flowing in response to a small electric field. The derivation will be sufficiently general to include any kind of particles and nearly any kind of interactions.

The electric field can be introduced as either a scalar or a vector potential. The QBE is independent of this choice, as required by gauge invariance. Here both are included, in order to provide the most general derivation. There will be an electric field \mathbf{E}_v , which is from a vector potential, and another electric field \mathbf{E}_s , from a scalar potential. The final version of the

QBE will include only the total electric field $\mathbf{E} = \mathbf{E}_v + \mathbf{E}_s$. The scalar potential is introduced through the interaction term

$$H_E = -e\mathbf{E}_s \cdot \sum_j \mathbf{r}_j \quad (8.257)$$

The vector potential is introduced by changing the momentum of each charged particle to $(\mathbf{p} - e\mathbf{A}/c)$, where the vector potential is $\mathbf{A} = -c\mathbf{E}_v t$. A vector potential proportional to time could occur in a wire loop with a slowly varying magnetic flux through the center. The present theory is for a dc electric field.

Equations (3.331) and (3.332) are equations of motion for the four Green's functions in the 2×2 matrix for \bar{G} . The two electric field terms are added to the left-hand side of these equations—they are included in H_0 . The derivation of the QBE involves several algebraic manipulations on the left-hand side of these equations. In order to avoid a lot of cumbersome notation, the right-hand side of these equations are not written during these steps. On the left, the steps are the same for all six of the real-time Green's functions. The generic symbol G applies to any one of them.

The two equations for G in (3.331) and (3.332) are first added and then subtracted:

$$\left[i \left(\frac{\partial}{\partial t_1} - \frac{\partial}{\partial t_2} \right) - H_1 - H_2 \right] G = \quad (8.258)$$

$$\left[i \left(\frac{\partial}{\partial t_1} + \frac{\partial}{\partial t_2} \right) - H_1 + H_2 \right] G = \quad (8.259)$$

where $H_1 = H_0(\mathbf{r}_1, \mathbf{p}_1)$ and $H_2 = H_0(\mathbf{r}_2, -\mathbf{p}_2)$. The two equations contain time derivatives that relate either to the relative or center-of-mass motion

$$\frac{\partial}{\partial t_1} + \frac{\partial}{\partial t_2} = \frac{\partial}{\partial T}, \quad \frac{\partial}{\partial t_1} - \frac{\partial}{\partial t_2} = 2 \frac{\partial}{\partial t} \quad (8.260)$$

For particles with parabolic band dispersion, the sum and difference of the two Hamiltonians produce simple expressions in relative coordinates. As a first step, consider what happens to the two momentum terms which are in the form of $(\mathbf{p} - e\mathbf{A}/c)$

$$\mathbf{p}_1 + e\mathbf{E}_v t_1 = (\mathbf{p} + e\mathbf{E}_v T) + \frac{1}{2}(\mathbf{P} + e\mathbf{E}_v t) \quad (8.261)$$

$$\mathbf{p}_2 - e\mathbf{E}_v t_2 = -(\mathbf{p} + e\mathbf{E}_v T) + \frac{1}{2}(\mathbf{P} + e\mathbf{E}_v t) \quad (8.262)$$

where \mathbf{p} and \mathbf{P} are the relative and center-of-mass momentum. It is important to understand all of the various plus and minus signs. Since $\mathbf{A} = -c\mathbf{E}_v t$ then $(\mathbf{p} - e\mathbf{A}/c) = (\mathbf{p} + e\mathbf{E}_v t)$ for \mathbf{p}_1 . However, in H_2 the momentum enters as $-\mathbf{p}_2$ which explains the sign change on the bottom. On the right one uses $t_{1,2} = T \pm t/2$ and $\mathbf{p}_{1,2} = \pm\mathbf{p} + \mathbf{P}/2$. These results make it easy to see the form of $H_1 \pm H_2$:

$$H_1 + H_2 = \frac{1}{m}(\mathbf{p} + e\mathbf{E}_v T)^2 + \frac{1}{4m}(\mathbf{P} + e\mathbf{E}_v t)^2 - 2e\mathbf{E}_s \cdot \mathbf{R} \quad (8.263)$$

$$H_1 - H_2 = \frac{1}{m}(\mathbf{p} + e\mathbf{E}_v T) \cdot (\mathbf{P} + e\mathbf{E}_v t) - e\mathbf{E}_s \cdot \mathbf{r} \quad (8.264)$$

Divide the top equation by two, and obtain the following two equations for the Green's function:

$$2 \left[i \frac{\partial}{\partial t} - \frac{1}{2m} (\mathbf{p} + e\mathbf{E}_v T)^2 - \frac{1}{m} (\mathbf{P} + e\mathbf{E}_v t)^2 + e\mathbf{E}_s \cdot \mathbf{R} \right] G = \quad (8.265)$$

$$\left[i \frac{\partial}{\partial T} - \frac{1}{m} (\mathbf{p} + e\mathbf{E}_v T) \cdot (\mathbf{P} + e\mathbf{E}_v t) + e\mathbf{E}_s \cdot \mathbf{r} \right] G = \quad (8.266)$$

These two equations describe the relative and center-of-mass motion of the function $G(\mathbf{r}, t; \mathbf{R}, T)$. The goal is to derive the QBE for the WDF $G^<(\mathbf{k}, \omega; \mathbf{R}, T)$. Fourier transform the variables (\mathbf{r}, t) to the set (\mathbf{q}, Ω) as in (8.256). This transform changes \mathbf{p} to \mathbf{q} , \mathbf{r} to $i\nabla_{\mathbf{q}}$, d/dt to $-i\Omega$, and t to $-i\partial/\partial\Omega$. The two transformed equations are

$$2 \left[\Omega + e\mathbf{E}_s \cdot \mathbf{R} - \frac{1}{2m} (\mathbf{q} + e\mathbf{E}_v T)^2 + \frac{1}{8m} \left(\nabla_R + e\mathbf{E}_v \frac{\partial}{\partial\Omega} \right)^2 \right] G(\mathbf{q}, \Omega; \mathbf{R}, T) = \quad (8.267)$$

$$i \left[\frac{\partial}{\partial T} + \frac{1}{m} (\mathbf{q} + e\mathbf{E}_v T) \cdot \left(\nabla_R + e\mathbf{E}_v \frac{\partial}{\partial\Omega} \right) + e\mathbf{E}_s \cdot \nabla_{\mathbf{q}} \right] G(\mathbf{q}, \Omega; \mathbf{R}, T) =$$

The second of these equations has, on the left-hand side, exactly the same terms as are found in the Boltzmann equation. This similarity suggests that the QBE is the same as the BE. However, this conclusion is incorrect. The above set of equations have several things wrong with them

1. There is the term $e\mathbf{E}_s \cdot \mathbf{R}$. This term seems to combine with Ω to produce a center-of-mass energy $\omega = \Omega + e\mathbf{E}_s \cdot \mathbf{R}$. Together they suggest that the energy of a particle depends upon its location. The energy is different at one end of a sample than at the other. However, this behavior is contrary to common sense. When there is a small electric field along the sample, and a small current flowing, we expect the system to be uniform. There is the same particle density, current density, etc. at each point in the solid. There is no dependence upon \mathbf{R} . This undesirable term has to be eliminated.
2. The relative momentum seems to enter in the combination of $\mathbf{q} + e\mathbf{E}_v T$. It depends upon the center-of-mass time T . This feature is also unphysical, and needs to be eliminated.
3. The result is not gauge invariant, since the two fields \mathbf{E}_s and \mathbf{E}_v enter differently.

All of these problems can be eliminated through a variable transformation. Of course, this transformation also causes some derivatives to change:

$$\begin{aligned} \Omega + e\mathbf{E}_s \cdot \mathbf{R} &\Rightarrow \omega \\ \mathbf{q} + e\mathbf{E}_v T &\Rightarrow \mathbf{k} \\ \nabla_R &\Rightarrow \nabla_R + e\mathbf{E}_s \frac{\partial}{\partial\omega} \\ \frac{\partial}{\partial T} &\Rightarrow \frac{\partial}{\partial T} + e\mathbf{E}_v \cdot \nabla_{\mathbf{k}} \end{aligned} \quad (8.268)$$

These transformations cause the two equations in (8.267) for the Green's function to now have the form

$$\begin{aligned} & \left[\omega - \varepsilon_{\mathbf{k}} + \frac{1}{8m} \left(\mathbf{V}_R + e\mathbf{E} \frac{\partial}{\partial \omega} \right)^2 \right] G(\mathbf{k}, \omega; \mathbf{R}, T) = \\ & i \left[\frac{\partial}{\partial T} + \mathbf{v}_{\mathbf{k}} \cdot \mathbf{V}_R + e\mathbf{E} \cdot \left(\mathbf{V}_{\mathbf{k}} + \mathbf{v}_{\mathbf{k}} \frac{\partial}{\partial \omega} \right) \right] \mathbf{G}(\mathbf{k}, \omega; \mathbf{R}, T) = \end{aligned} \quad (8.269)$$

The notation on the Green's function has been changed again. The arguments (\mathbf{q}, Ω) have been changed to (\mathbf{k}, ω) . The left-hand sides of these two equations are now in the form that is useful. The lower equation has exactly the same terms as in (8.254), and is the quantum Boltzmann equation. The electric field is $\mathbf{E} = \mathbf{E}_v + \mathbf{E}_s$. Both electric fields contribute in the same way, and the result is now gauge invariant.

The above variable transformation makes the QBE gauge invariant. The results are valid for dc electric fields. A more complicated variable transformation is required for ac electric fields (Levanda and Fleurov, 1994).

Now it is time to restore the scattering terms to the right-hand side of the equal sign. The scattering terms for $G^<$ and G_{ret} are in (3.341). They are the most important Green's functions in applications using real time. The self-energy terms on the right are found by following the same steps to bring the left-hand side to (8.269). In doing these operations, keep in mind that the scattering terms are of the form

$$\int dx_3 \Sigma(x_1, x_3) G(x_3, x_2)$$

First add and subtract the two equations for each Green's function. There follows a series of variable transformations. They can be understood by examining just one term in the right-hand side of Eqn. (8.269). The various variable arguments yield a fairly complicated expression, which is derived by the following steps:

1. Change the two sets of variables $(x_1, x_3)(x_3, x_2)$ to the center-of-mass grouping $(x_1 - x_3, (x_1 + x_3)/2)(x_3 - x_2, (x_3 + x_2)/2)$.
2. Change the integration variable from x_3 to $y = x_1 - x_3$, so the variable grouping become $(y, x_1 - y/2)(x_1 - x_2 - y, (x_1 + x_2 - y)/2)$.
3. Change to center of mass variables $x = (x_1 - x_2)$, $X = (x_1 + x_2)/2$, which produces the variable grouping $(y, X + (x - y)/2)(x - y, X - y/2)$.
4. Change the integration variable x to $z = x - y$ which produces the final arguments of $(y, X + z/2)(z, X - y/2)$.

Below are the full scattering equations for $G^<$ and G_{ret} . The explicit variables in the scattering terms are only written out in the first term, but are identical for the other three. The order of the terms in the scattering integral is important. The first one always has argument $(y, X + z/2)$ while the second one has argument $(z, X - y/2)$. The four-vector in the exponent

is $q = (\mathbf{k}, \omega)$:

$$\begin{aligned} & \left[\omega - \varepsilon_{\mathbf{k}} + \frac{1}{8m} \left(\mathbf{V}_R + e\mathbf{E} \frac{\partial}{\partial \omega} \right)^2 \right] G_{\text{ret}}(\mathbf{k}, \omega; \mathbf{R}, T) \\ &= 1 + \frac{1}{2} \int dze^{-iq} \int dye^{-iq} [\Sigma_{\text{ret}}(y, X + z/2) G_{\text{ret}}(z, X - y/2) \\ & \quad + G_{\text{ret}}(y, X + z/2) \Sigma_{\text{ret}}(z, X - y/2)] \end{aligned} \quad (8.270)$$

$$\begin{aligned} & i \left[\frac{\partial}{\partial T} + \mathbf{v}_{\mathbf{k}} \cdot \mathbf{V}_R + e\mathbf{E} \cdot \left(\mathbf{V}_{\mathbf{k}} + \mathbf{v}_{\mathbf{k}} \frac{\partial}{\partial \omega} \right) \right] G_{\text{ret}}(\mathbf{k}, \omega; \mathbf{R}, T) \\ &= \int dze^{-iq} \int dye^{-iq} [\Sigma_{\text{ret}} G_{\text{ret}} - G_{\text{ret}} \Sigma_{\text{ret}}] \end{aligned} \quad (8.271)$$

$$\begin{aligned} & \left[\omega - \varepsilon_{\mathbf{k}} + \frac{1}{8m} \left(\mathbf{V}_R + e\mathbf{E} \frac{\partial}{\partial \omega} \right)^2 \right] G^<(\mathbf{k}, \omega; \mathbf{R}, T) \\ &= \frac{1}{2} \int dze^{-iq} \int dye^{-iq} [\Sigma_t G^< - \Sigma^< G_t + G_t \Sigma^< - G^< \Sigma_t] \end{aligned} \quad (8.272)$$

$$\begin{aligned} & i \left[\frac{\partial}{\partial T} + \mathbf{v}_{\mathbf{k}} \cdot \mathbf{V}_R + e\mathbf{E} \cdot \left(\mathbf{V}_{\mathbf{k}} + \mathbf{v}_{\mathbf{k}} \frac{\partial}{\partial \omega} \right) \right] G^<(\mathbf{k}, \omega; \mathbf{R}, T) \\ &= \int dze^{-iq} \int dye^{-iq} [\Sigma_t G^< - \Sigma^< G_t - G_t \Sigma^< + G^< \Sigma_t] \end{aligned} \quad (8.273)$$

These four equations are the important ones for nonequilibrium calculations. Although one primarily wants to find $G^<$, it is always necessary to first find G_{ret} . These equations were first derived by Fleurov and Kozlov (1978).

So far no approximations have been made, and the equations are exact. The QBE has a linear term in the electric field \mathbf{E} . However, the equation is exact to all powers of \mathbf{E} , not just to the first power.

8.5.2. Gradient expansion

Equations (8.271) are usually too hard to solve because the scattering terms on the right have a complicated form. Some sort of approximation has to be introduced to simplify the right side. The approximation described below is only valid to first order in the field. Kadanoff and Baym (1962) introduced an approximation for evaluating these scattering terms, which is called the “gradient expansion.” They assume that the center-of-mass time T is very large, and take the limit that $T \rightarrow \infty$. At large values of T , they assume that the system is approaching its asymptotic limit, so that variations with respect to T are small. Obviously the gradient expansion is not suitable for studying transients, since it is poor at small values of T . Neither is it useful for steady-state ac phenomena (Mahan, 1987). Indeed, the T dependence is so poorly described in the gradient approximation that it should not be used and the T derivative terms should be dropped from the QBE. Nevertheless, the gradient expansion is used since it is applicable for homogeneous (small R dependence) steady state (small T dependence) systems.

The center-of-mass variables are all in the form $(\mathbf{R} + \Delta\mathbf{R}, T + \Delta T)$, which are expanded in a Taylor series about the point $X \equiv (\mathbf{R}, T)$. The integrals can be done for each term in the

series. These integrals usually cause further derivatives. The gradient expansion is shown below for a typical scattering term:

$$I = \int dze^{-iq} \int dye^{-iq} \Sigma(y, X + z/2) G(z, X - y/2) \quad (8.274)$$

$$= \int dze^{-iq} \int dye^{-iq} \left\{ \Sigma(y, X) G(z, X) + \frac{z}{2} \frac{\partial \Sigma(y, X)}{\partial X} G(z, X) - \Sigma(y, X) \frac{y}{2} \frac{\partial G(z, X)}{\partial X} + \dots \right\} \quad (8.275)$$

$$= \Sigma(q, X) G(q, X) + \frac{i}{2} [(\nabla_q \Sigma) \nabla_X G - (\nabla_X \Sigma) \nabla_q G] + \dots \quad (8.276)$$

The first term on the right are usually the largest. It is customary to retain the first derivative terms and ignore higher derivatives. They can be expressed using a Poisson bracket notation:

$$[\Sigma, G] = (\nabla_q \Sigma) \nabla_X G - (\nabla_X \Sigma) \nabla_q G \quad (8.277)$$

$$= \frac{\partial \Sigma}{\partial \Omega} \frac{\partial G}{\partial T} - \frac{\partial G}{\partial \Omega} \frac{\partial \Sigma}{\partial T} - \mathbf{v}_q \Sigma \cdot \nabla_{\mathbf{R}} G + \mathbf{v}_q G \cdot \nabla_{\mathbf{R}} \Sigma \quad (8.278)$$

These frequency derivatives are with respect to the Ω variables, since the variable change ($\mathbf{k} = \mathbf{q} + e\mathbf{E}_v T$, $\omega = \Omega + e\mathbf{E}_s \cdot \mathbf{R}$) has not yet been made. If this step is taken now, the derivatives get altered according to Eqn. (8.268), which changes the Poisson brackets to

$$[C, D] \Rightarrow [C, D] + e\mathbf{E} \cdot \left[\left(\frac{\partial C}{\partial \omega} \right) \nabla_{\mathbf{k}} D - \left(\frac{\partial D}{\partial \omega} \right) \nabla_{\mathbf{k}} C \right] \quad (8.279)$$

where \mathbf{E} is again the total electric field. This analysis finally derives from (8.271) the following expression for the nonequilibrium retarded Green's function in an electric field, when G_{ret} and Σ_{ret} depend upon $(\mathbf{k}, \omega; \mathbf{R}, T)$:

$$\left[\omega - \varepsilon_{\mathbf{k}} + \frac{1}{8m} \left(\mathbf{v}_R + e\mathbf{E} \frac{\partial}{\partial \omega} \right)^2 - \Sigma_{\text{ret}} \right] G_{\text{ret}}(\mathbf{k}, \omega; \mathbf{R}, T) = 1 \quad (8.280)$$

$$i \left\{ \frac{\partial}{\partial T} + \mathbf{v}_{\mathbf{k}} \cdot \nabla_{\mathbf{R}} + e\mathbf{E} \cdot \left[\left(1 - \frac{\partial \Sigma_{\text{ret}}}{\partial \omega} \right) \nabla_{\mathbf{k}} + (\mathbf{v}_{\mathbf{k}} + \nabla_{\mathbf{k}} \Sigma_{\text{ret}}) \frac{\partial}{\partial \omega} \right] \right\} G_{\text{ret}} = i[\Sigma_{\text{ret}}, G_{\text{ret}}], \quad (8.281)$$

The additional terms from the Poisson bracket, which are linear in the field \mathbf{E} , have been transferred to the left of the equal sign.

These equations simplify for nonequilibrium systems which are both homogeneous ($\nabla_{\mathbf{R}} = 0$) and steady state ($\partial/\partial T = 0$). The Poisson brackets vanish, as do several terms on the left. Also ignored are terms nonlinear in the electric field $O(\mathbf{E}^2)$, and find for the above two equations

$$[\omega - \varepsilon_{\mathbf{k}} - \Sigma_{\text{ret}}] G_{\text{ret}} = 1 \quad (8.282)$$

$$ie\mathbf{E} \cdot \left[\left(1 - \frac{\partial \Sigma_{\text{ret}}}{\partial \omega} \right) \nabla_{\mathbf{k}} + (\mathbf{v}_{\mathbf{k}} + \nabla_{\mathbf{k}} \Sigma_{\text{ret}}) \frac{\partial}{\partial \omega} \right] G_{\text{ret}} = 0 \quad (8.283)$$

The first equation is easily solved, to yield

$$G_{\text{ret}}(\mathbf{k}, \omega) = \frac{1}{\omega - \varepsilon_{\mathbf{k}} - \Sigma_{\text{ret}}(\mathbf{k}, \omega)} + O(\mathbf{E}^2) \quad (8.284)$$

The retarded Green's function appears to have no first-order term in the electric field. It actually does, since the self-energy Σ_{ret} has a term linear in the field due to the electron–phonon interaction. This term is small and seems to have little effect. If it is ignored then the retarded Green's function is unchanged from its value in equilibrium. This result considerably simplifies the solution to the QBE. The solution (8.284) also satisfies the equation (8.283). Related quantities such as the advanced function $G_{\text{adv}} = G_{\text{ret}}^*$ and the spectral function $A(\mathbf{k}, \omega)$ are also unchanged to first order in the electric field. This completes the discussion of the retarded Green's function in a static electric field.

Next the gradient expansion is applied to (8.271) for the equations for $G^<$. Again keeping only first-order derivatives, the final results are the quantum Boltzmann equation (QBE):

$$\begin{aligned} \left[\omega - \varepsilon_{\mathbf{k}} + \frac{1}{8m} \left(\mathbf{V}_R + e\mathbf{E} \frac{\partial}{\partial \omega} \right)^2 \right] G^< &= G^< \text{Re}[\Sigma_{\text{ret}}] + \Sigma_{\text{ret}} \text{Re}[G_{\text{ret}}] \\ &+ \frac{i}{4} [\Sigma^>, G^<] - \frac{i}{4} [\Sigma^<, G^>] + \frac{i}{4} e\mathbf{E} \cdot \left[\left(\frac{\partial \Sigma^>}{\partial \omega} - \frac{\partial \Sigma^<}{\partial \omega} \right) \mathbf{V}_{\mathbf{k}} G^< \right. \\ &\left. + \left(\frac{\partial G^>}{\partial \omega} - \frac{\partial G^<}{\partial \omega} \right) \mathbf{V}_{\mathbf{k}} \Sigma^< \right] \end{aligned} \quad (8.285)$$

$$\begin{aligned} i \left\{ \frac{\partial}{\partial T} + \mathbf{v}_{\mathbf{k}} \cdot \mathbf{V}_R + e\mathbf{E} \cdot \left[\left(1 - \frac{\partial \Sigma_{\text{ret}}}{\partial \omega} \right) \mathbf{V}_{\mathbf{k}} + (\mathbf{v}_{\mathbf{k}} + \mathbf{V}_{\mathbf{k}} \Sigma_{\text{ret}}) \frac{\partial}{\partial \omega} \right] \right\} G^< \\ - ie\mathbf{E} \cdot \left[\frac{\partial \Sigma^<}{\partial \omega} \mathbf{V}_{\mathbf{k}} \text{Re}[G_{\text{ret}}] - \frac{\partial G_{\text{ret}}}{\partial \omega} \mathbf{V}_{\mathbf{k}} \Sigma^< \right] = \Sigma^> G^< - \Sigma^< G^> \\ + i[\text{Re}[\Sigma_{\text{ret}}], G^<] + i[\Sigma^<, \text{Re}[G_{\text{ret}}]] \end{aligned} \quad (8.286)$$

Standard relationships have been used in deriving this equation, such as $G_t - G_{\bar{t}} = 2 \text{Re}[G_{\text{ret}}]$ and $G_t + G_{\bar{t}} = G^< + G^>$.

Equation (8.286) is the quantum Boltzmann equation. It is rather formidable. It is also difficult to solve, since it is usually an integral equation. Sometimes it is nonlinear in the particle density because the self-energy functions $\Sigma^<$ and $\Sigma^>$ are also functions of $G^<$ and $G^>$.

The QBE also contains the functions $G^>$ and $\Sigma^>$. Similar equations for these functions can be derived by starting from the general equations (3.331) and (3.332). This derivation shows that the equation for $G^>$ is almost identical to the one for $G^<$. In fact, one can prove that the following identities are valid:

$$G^> = G^< - iA \quad (8.287)$$

$$\Sigma^> = \Sigma^< + i \text{Im}[\Sigma_{\text{ret}}] = \Sigma^< - 2i\Gamma \quad (8.288)$$

These relations are trivial to show for equilibrium, but they are also valid for nonequilibrium situations. These identities will be used often to simplify expressions. For example, the main scattering term in the QBE can be immediately simplified to

$$\Sigma^> G^< - \Sigma^< G^> = -i\{2\Gamma G^< - \Sigma^< A\} \quad (8.289)$$

This result will be employed in the calculations. The quantities $G^<$ and $\Sigma^<$ are generally proportional to the density of particles, while retarded functions are only indirectly dependent upon the density of particles—only through the self-energy function Σ_{ret} . In the QBE, each

term has one factor that is either $G^<$ or $\Sigma^<$, so each term is proportional to the density of particles. This equation does have the character of a transport equation.

The QBE simplifies for the treatment of systems that are homogeneous ($\mathbf{V}_R = 0$) and steady ($\partial/\partial T = 0$). These derivatives are dropped, as well as the Poisson brackets, since the latter contain similar derivatives. In this case the QBE is

$$e\mathbf{E} \cdot \left[\left(1 - \frac{\partial \Sigma_{\text{ret}}}{\partial \omega} \right) \mathbf{v}_k + (\mathbf{v}_k + \mathbf{v}_k \Sigma_{\text{ret}}) \frac{\partial}{\partial \omega} \right] G^< - ie\mathbf{E} \cdot \left[\frac{\partial \Sigma^<}{\partial \omega} \mathbf{v}_k \text{Re}[G_{\text{ret}}] - \frac{\partial G_{\text{ret}}}{\partial \omega} \mathbf{v}_k \Sigma^< \right] = \Sigma^< A - 2\Gamma G^< \quad (8.290)$$

This equation can be simplified. The QBE is only valid to first power in the electric field, since terms of $O(\mathbf{E}^2)$ have systematically been ignored. For example, in the gradient expansion the second derivative terms would give contributions of $O(\mathbf{E}^2)$ were they retained. So the equation is exact to first order in the electric field, but is not exact to higher orders in the field. This fact is utilized to simplify the left-hand side of the equation. Since the field multiplies each term, on this side of the equal sign the Green's functions can be taken to have $\mathbf{E} = 0$. Of course, these expressions are just the equilibrium quantities in Sec. 3.7.

Consider the frequency derivatives of the left-hand side of the equation. Write the spectral function in the shorthand notation

$$A = \frac{2\Gamma}{\sigma^2 + \Gamma^2}, \quad \Gamma = -2 \text{Im}[\Sigma_{\text{ret}}], \quad \sigma = \omega - \varepsilon_k - \text{Re}[\Sigma_{\text{ret}}] \quad (8.291)$$

The left-hand side will have three types of frequency derivatives: $\partial n_F/\partial \omega$, $\partial \gamma/\partial \omega$, $\partial \sigma/\partial \omega$. The coefficients of the latter two vanish, which leaves only terms in $\partial n_F/\partial \omega$. All of the terms proportional to n_F vanish, leaving only terms multiplied by $\partial n_F/\partial \omega$:

$$A(\mathbf{k}, \omega)^2 \frac{\partial n_F}{\partial \omega} e\mathbf{E} \cdot \{ (\mathbf{v}_k + \mathbf{v}_k \text{Re}[\Sigma_{\text{ret}}]) \Gamma + \sigma \mathbf{v}_k \Gamma \} = \Sigma^> G^< - \Sigma^< G^> \quad (8.292)$$

The factor of $A(\mathbf{k}, \omega)^2$ appears in each term, and was taken outside. The left-hand side of this equation now contains only known quantities, which can be calculated in equilibrium. The scattering terms remain on the right-hand side. Finding them still involves work, usually in the form of an integral equation. This final form for the QBE is exact for transport which is linear in the field, and for steady state, homogeneous systems. It is quite analogous to the similar expression for the classical BE, which is

$$-e\mathbf{E} \cdot \mathbf{v}_k \frac{\partial f^{(0)}}{\partial \omega} = \Sigma^> G^< - \Sigma^< G^> \quad (8.293)$$

The classical equation has $\partial f^{(0)}/\partial \omega$, while (8.292) has $\partial n_F/\partial \omega$.

Equation (8.292) is the steady state, homogeneous form of the QBE. It should be the starting point for many transport calculations. It is exact, and is an alternative to using the Kubo formalism, which is also exact. The derivation of this equation has been complicated, and has entailed some work. However, once derived, it is often the easiest starting point for deriving the transport coefficients. Calculations using Eqn. (8.292) entail less work in getting to the answer than any other formalism.

8.5.3. Electron scattering by impurities

The quantum Boltzmann equation (QBE) (8.292) will be solved for the electron scattering by impurities. This case was solved in Sec. 8.1 using the Kubo formula. This exercise is useful for two reasons: (1) producing the known result demonstrates the correctness of the QBE; and (2) this case is the easiest one to solve, and provides an introduction to the techniques for solving the QBE. The present example makes two assumptions that are intended to make the solution as easy as possible: (1) the impurities are dilute, so that the simultaneous scattering from several impurities can be neglected; and (2) the impurities have no internal degrees of freedom, such as spin or vibrations, which can be altered by the electron scattering. The second assumption implies that impurity is a simple potential that elastically scatters the electron. The impurities are randomly located in the solid. The method of Sec. 4.1.5 is employed for averaging over the random distributions.

The first step in solving the QBE is to find the retarded functions. For impurity scattering they are independent of the electric field, at least for small fields. The retarded self-energy for scattering from impurities of density n_i is given in Sec. 4.1

$$\Sigma_{\text{ret}}(\mathbf{k}, \omega) = n_i T_{\mathbf{k}\mathbf{k}}(\omega) \quad (8.294)$$

The T matrix is energy dependent, and hence off-shell.

The next step is to express the self-energies $\Sigma^<$ and $\Sigma^>$ in terms of the Green's functions $G^<$ and $G^>$. The retarded functions G_{ret} and Σ_{ret} are known, since they are unchanged by the electric field. However, $G^<$ and $G^>$ are, at this point, unknown since they are affected by the electric field. The self-energy functions are

$$\Sigma^{<,>}(\mathbf{k}, \omega) = n_i \int \frac{d^3 p}{(2\pi)^3} |T_{\mathbf{p}\mathbf{k}}(\omega)|^2 G^{<,>}(\mathbf{p}, \omega) \quad (8.295)$$

This result is derived below. The off-diagonal T matrix is the one in Eqn. (4.111), which also depends upon the energy ω .

Equation (8.295) is now derived. An impurity at $\mathbf{R} = 0$ is represented by an electron potential $V(\mathbf{r})$, whose Fourier transform is $V(\mathbf{q})$. The self-energy Σ from a single scattering event is

$$\tilde{\Sigma}_0(x_1, x_2) = V(\mathbf{r}_1) \delta^4(x_1 - x_2) \tilde{I} \quad (8.296)$$

where \tilde{I} is the unit tensor. This self-energy is inserted into (2.157). That equation is iterated in order to find the effects of repeated scattering from the same impurity. The resulting self-energy series is rewritten in a symbolic notation, where the product of two functions implies an integral over dx . Iteration of these equations gives the series for $\tilde{\Sigma}$:

$$\tilde{\Sigma} = \tilde{\Sigma}_0[\tilde{I} + \tilde{G}_0 \tilde{\Sigma}_0 + \tilde{G}_0 \tilde{\Sigma}_0 \tilde{G}_0 \tilde{\Sigma}_0 + \dots] \quad (8.297)$$

After summing this series, the matrix $\tilde{\Sigma}$ is examined for its individual components. The ones for $\Sigma^{<,>}$ are

$$\Sigma^{<,>} = [1 + G_{\text{ret}} \Sigma_{\text{ret}}] \Sigma_0^{<,>} [1 + \Sigma_{\text{adv}} G_{\text{adv}}] + \Sigma_{\text{ret}} G^{<,>} \Sigma_{\text{adv}} \quad (8.298)$$

Note the analogy with the equation for $G^<$ in (2.159). The resemblance is expected, since the series for $\tilde{\Sigma}$ has the same mathematical structure as the one for G .

For impurity scattering, the unperturbed self-energies $\Sigma_0^{<,>} = 0$ are zero. There is only the last term in (8.298). Since the self-energies Σ_{ret} and Σ_{adv} are T matrices, (8.298) gives the result in (8.295). The damping function $\Gamma(\mathbf{k}, \omega)$ and other self-energy $\Sigma^>(\mathbf{k}, \omega)$ are

$$2\Gamma(\mathbf{k}, \omega) = n_i \int \frac{d^3p}{(2\pi)^3} |T_{\mathbf{pk}}|^2 A(\mathbf{p}, \omega) \quad (8.299)$$

$$\Sigma^> = \Sigma^< - 2i\Gamma \quad (8.300)$$

For dilute impurities, it is sufficient to retain only those terms that are first order in the impurity concentration n_i . The broadening due to the spectral function $A(\mathbf{p}, \omega)$ is from impurity scattering. Since the self-energy is multiplied by n_i , one can replace the spectral function $A(\mathbf{p}, \omega) \approx 2\pi\delta(\omega - \varepsilon_{\mathbf{p}})$. This expression is then equal to the imaginary part of the T matrix: $2\Gamma = n_i v_k \sigma_T$, where σ_T is the total cross section from impurity scattering in Sec. 4.1.

The starting point for solving the QBE is (8.292). On the left of the equals sign is the factor of

$$[(\mathbf{v}_{\mathbf{k}} + \mathbf{V}_{\mathbf{k}} \text{Re}[\Sigma_{\text{ret}}])\Gamma + \sigma \mathbf{V}_{\mathbf{k}} \Gamma] A(\mathbf{k}, \omega)^2 \quad (8.301)$$

This expression is simplified by neglecting terms of $O(n_i^2)$ such as $\Gamma \text{Re}[\Sigma_{\text{ret}}]$. The term in $\sigma = \omega - \varepsilon_{\mathbf{k}} - \text{Re}[\Sigma_{\text{ret}}]$ is small since the factor of $A(\mathbf{k}, \omega)^2$ tends to force $\sigma \approx 0$. The above expressions can be approximated by $v_{\mathbf{k}} \Gamma A(\mathbf{k}, \omega)^2$:

$$A(\mathbf{k}, \omega)^2 \frac{dn_F(\omega)}{d\omega} e\mathbf{E} \cdot \mathbf{v}_{\mathbf{k}} \Gamma - 2i\Gamma G^< - i\Sigma^< A \quad (8.302)$$

These steps complete the derivation of the left-hand side of the QBE.

On the right side of (8.302) the scattering terms vanish in equilibrium since $\Sigma^< = 2in_F\Gamma$ and $G^< = in_F A$. Since current is flowing in response to the field, the system is slightly out of equilibrium. The right-hand side is expected to have factors similar to those on the left. This discussion suggests the following ansatz for the nonequilibrium Green's function

$$G^< = iA(\mathbf{k}, \omega) \left[n_F(\omega) - \left(\frac{dn_F(\omega)}{d\omega} \right) e\mathbf{E} \cdot \mathbf{v}_{\mathbf{k}} \Lambda(\mathbf{k}, \omega) \right] \quad (8.303)$$

The function $\Lambda(\mathbf{k}, \omega)$ is unknown, and needs to be determined by solving the QBE. The factors that multiply Λ are for later convenience. The above choice does not make any assumptions for the value of Λ . Using this ansatz in the self-energy function gives

$$\begin{aligned} \Sigma^< &= in_i \int \frac{d^3p}{(2\pi)^3} |T_{\mathbf{kp}}|^2 A(\mathbf{p}, \omega) \left[n_F(\omega) - \left(\frac{dn_F(\omega)}{d\omega} \right) e\mathbf{E} \cdot \mathbf{v}_{\mathbf{p}} \Lambda(\mathbf{p}, \omega) \right] \\ &= 2i\Gamma n_F - in_i \frac{dn_F}{d\omega} n_i \int \frac{d^3p}{(2\pi)^3} |T_{\mathbf{kp}}|^2 A(\mathbf{p}, \omega) e\mathbf{E} \cdot \mathbf{v}_{\mathbf{p}} \Lambda(\mathbf{p}, \omega) \end{aligned} \quad (8.304)$$

Putting these two expressions into the right-hand side of (8.302), the equilibrium terms cancel, and the remaining terms each have the common factor of $(dn_F/d\omega)A$:

$$\begin{aligned} A^2 \frac{dn_F(\omega)}{d\omega} e\mathbf{E} \cdot v_{\mathbf{k}} \Gamma &= - \left(\frac{dn_F(\omega)}{d\omega} \right) A 2\Gamma e\mathbf{E} \cdot \mathbf{v}_{\mathbf{k}} \Lambda \\ &\quad - n_i \int \frac{d^3p}{(2\pi)^3} |T_{\mathbf{kp}}|^2 A(\mathbf{p}, \omega) e\mathbf{E} \cdot \mathbf{v}_{\mathbf{p}} \Lambda(\mathbf{p}, \omega) \end{aligned} \quad (8.305)$$

After canceling all of the common factors, there is an integral equation for the unknown function $\Lambda(\mathbf{k}, \omega)$:

$$\mathbf{v}_k \Lambda(\mathbf{k}, \omega) = \frac{1}{2} A(\mathbf{k}, \omega) \mathbf{v}_k + \frac{ni}{2\Gamma} \int \frac{d^3p}{(2\pi)^3} |T_{\mathbf{k}\mathbf{p}}|^2 A(\mathbf{p}, \omega) \mathbf{v}_p \Lambda(\mathbf{p}, \omega) \quad (8.306)$$

This integral equation for $\Lambda(\mathbf{K}, \omega)$ is nearly identical to the one found in solving the conductivity from the Kubo relation. The quantity $\mathbf{v}_k \Lambda(\mathbf{k}, \omega)$ is similar to the factor of $\Gamma(\mathbf{k}, \omega - i\delta, \omega + i\delta)$ in Sec. 8.1.3. As the derivation proceeds, the differences between these quantities disappears, and the two approaches give the same resistivity.

After solving this equation for $\Lambda(\mathbf{k}, \omega)$, the Green's function $G^<$ in (8.303) is known. It is used in (3.340) for the current. The first term in (8.303) for $G^<$ gives a zero current. The second term gives a current proportional to the field \mathbf{E} , and this proportionality defines the electrical conductivity

$$\sigma_{\mu\nu} = e^2 \int \frac{d^3p}{(2\pi)^3} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} v_{p\mu} v_{p\nu} \left(-\frac{dn_F(\omega)}{d\omega} \right) A(\mathbf{p}, \omega) \Lambda(\mathbf{p}, \omega) \quad (8.307)$$

The two equations (8.306) and (8.307) provide the solution for the conductivity from the QBE.

For electrons in metals, with a spherical Fermi surface, the conductivity from impurity scattering is found easily from these equations. First, do the integral over $d\omega$ in (8.307). At low temperature the factor $(-dn_F/d\omega)$ is nearly a delta function which sets $\omega \approx 0$. At low temperature

$$\sigma_{\mu\nu} = e^2 \int \frac{d^3p}{(2\pi)^3} v_{p\mu} v_{p\nu} A(\mathbf{p}, 0) \Lambda(\mathbf{p}, 0) \quad (8.308)$$

The equation is reduced as far as possible. Next one must solve the equation (8.306) for Λ . Adopt a vector coordinate system where the \hat{z} direction is \hat{k} . Then the various scalar products of vectors can be found from the law of cosines:

$$\hat{k} \cdot \hat{E} = \cos(\theta_0), \quad \hat{k} \cdot \hat{p} = \cos(\theta) \quad (8.309)$$

$$\hat{p} \cdot \hat{E} = \cos(\theta_0) \cos(\theta) + \sin(\theta) \sin(\theta_0) \cos(\phi) \quad (8.310)$$

Equation (8.306) is multiplied by the vector \hat{E} . The integral over the $d\phi$ part of d^3p makes the $\cos(\phi)$ term vanish. Each of the remaining terms has a factor of $\cos(\theta_0)$ which can be canceled. Then one finds the scalar equation

$$\Lambda(k, 0) = \frac{1}{2} A(k, 0) + \frac{n_i}{2\Gamma} \int \frac{d^3p}{(2\pi)^3} |T_{\mathbf{k}\mathbf{p}}|^2 A(\mathbf{p}, 0) \Lambda(\mathbf{p}, 0) \cos(\theta) \quad (8.311)$$

The spectral functions $A(k,0)$ and $A(p,0)$ force $k = p = k_F$. The factor of $\Lambda(p, 0)$ under the integral can be set equal to $\Lambda(k_F, 0)$ and taken out of the integral. Then the above equation can be solved to find

$$\Lambda(k_F, 0) = \frac{\tau_t}{2\tau} A(k_F, 0) \quad (8.312)$$

where $\tau = 1/2\Gamma$ is the time between scattering from impurities, while τ_i is the lifetime which is important for resistivity; it has the factor of $(1 - \cos \theta')$:

$$\sigma = \frac{ne^2\tau_i}{m} \quad (8.313)$$

$$\frac{1}{\tau_i} = n_i \int \frac{d^3p}{(2\pi)^3} |T_{\mathbf{k}\mathbf{p}}|^2 A(\mathbf{p}, \omega) [1 - \cos(\theta)] \quad (8.314)$$

This formula for the resistivity is identical to (8.25). The QBE gives the same formula for the resistivity as found earlier from the Kubo formula. The two methods are also in exact agreement for much more complex cases, such as electron–phonon scattering at nonzero temperatures and frequency. Most of the steps in the derivation were spent in getting to (8.302). This equation is the starting point for any calculation for homogeneous systems. The number of steps between this equation and the final resistivity is small. The QBE is an efficient method of finding the resistivity or other transport coefficients.

8.6. QUANTUM DOT TUNNELING

8.6.1. Electron tunneling

Cohen *et al.* (1962) introduced the concept of the tunneling Hamiltonian, which has become universally adopted for the discussion of tunneling. Their idea was to write the Hamiltonian as three terms:

$$H = H_R + H_L + H_T \quad (8.315)$$

$$H_T = \sum_{\mathbf{k}\mathbf{p}\sigma} (T_{\mathbf{k}\mathbf{p}} C_{\mathbf{k}\sigma}^\dagger C_{\mathbf{p}\sigma} + \text{hc}) \quad (8.316)$$

The first term H_R is the Hamiltonian for particles on the right side of the tunneling junction. It contains all many-body interactions. Similarly, H_L has all the physics for particles on the left side of the junction. These two are considered to be strictly independent. Not only do these two operators commute, $[H_L, H_R] = 0$, but they commute term by term. The Hamiltonian on the right can be expressed in terms of one set of operators $C_{\mathbf{k}\sigma}$ and those on the left by another set $C_{\mathbf{p}\sigma}$, and these operators are independent $\{C_{\mathbf{k}\sigma}, C_{\mathbf{p}\sigma}^\dagger\} = 0$. This assumption is probably reasonable. They further assumed that the tunneling is caused by the term H_T in (8.316). The tunneling matrix element $T_{\mathbf{k}\mathbf{p}}$ can transfer particles through an insulating junction. This transfer rate is assumed to depend only on the wave vectors on the two sides \mathbf{k} and \mathbf{p} and not on other variables, such as the energy or spin of the particles.

The theory of electron tunneling was mainly applied to superconductors. It developed very rapidly and was entirely based on the tunneling Hamiltonian. The theory showed excellent agreement with the many experiments. The history books were written describing this satisfactory situation, and the scientists in this field wandered off to do something else. About this time there began a serious investigation, starting with Zawadowski (1967), Caroli *et al.* (1975), and Feuchtwang (1975), about the validity of the tunneling Hamiltonian. Of course it was found to be a poor approximation, since the tunneling rate depends on the energy of the particle as well as its wave vector. What does this turn of events do to the lovely agreement between theory and experiment for tunneling in superconductors? Actually, it probably changes none of it. The tunneling in superconductors takes place over a very narrow

span of energies in the metal, i.e., within a Debye energy of the Fermi surface. Also, all the electrons involved have their wave vector very near the Fermi wave vectors k_F and p_F on the two sides of the junction. It is an adequate approximation to treat the transfer rate $T_{\mathbf{k}\mathbf{p}}$ as a constant T_0 which is evaluated at k_F and p_F , because the variations in $T_{\mathbf{k}\mathbf{p}}$ with energy must be on the scale E/E_F which are negligible for $E \approx \Delta \approx 1$ meV. Similarly, the variation of $T_{\mathbf{k}\mathbf{p}}$ with \mathbf{p} or \mathbf{k} is on the scale of the Fermi wave vectors. One can treat the transfer rate T_0 as a constant if the energies involved are small. The tunneling Hamiltonian is believed to be an improper formalism only when the applied voltages are large, say 1 eV.

For quantum dots the energy scales are also rather small, and are near to the chemical potentials of the metal electrodes. Here the tunneling Hamiltonian is also thought to be a valid formalism.

The general model of a tunneling junction is shown in Fig. 8.12. It describes a non-equilibrium situation, since the chemical potential on the left-hand side μ_L is not the same as μ_R on the right. They differ by the applied voltage $eV = \mu_L - \mu_R$. The potential drop of eV occurs in the insulating region between the electrodes.

The tunneling Hamiltonian (8.316) is used to derive a correlation function for electron tunneling currents. This correlation function has the form of a Kubo formula, except for an important difference. The Kubo formula for the conductivity in Sec. 3.7 expresses the ratio between the current and the voltage (actually electric field). In tunneling theory, the correlation function gives the current as a function of voltage.

The tunneling current through the insulating region is expressed as the rate of change of the number of particles on, for example, the left-hand side of the junction N_L . This rate is found from the commutator of $N_L = \sum C_{\mathbf{p}\sigma}^\dagger C_{\mathbf{p}\sigma}$ with the tunneling Hamiltonian. Only the term H_T fails to commute with N_L ,

$$\dot{N}_L = i[H, N_L] = i[H_T, N_L] = i \sum_{\mathbf{k}\mathbf{p}\sigma} [T_{\mathbf{k}\mathbf{p}} C_{\mathbf{k}\sigma}^\dagger C_{\mathbf{p}\sigma} - T_{\mathbf{k}\mathbf{p}}^* C_{\mathbf{p}\sigma}^\dagger C_{\mathbf{k}\sigma}] \quad (8.317)$$

The total current I through the tunneling interface is defined as the average value of this operator:

$$I(t) = -e\langle \dot{N}_L(t) \rangle \quad (8.318)$$

The average value of $\langle \dot{N}_L(t) \rangle$ is obtained by following the same steps used to derive the Kubo formula in Sec. 3.8. The total Hamiltonian is written as $H = H' + H_T$, where $H' = H_R + H_L$. Go to the interaction representation, where the tunneling term H_T is treated as the interaction and everything else H' is H_0 . Then the S matrix is expanded in terms of the perturbation H_T .

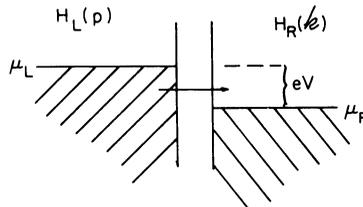


FIGURE 8.12 Tunneling between two normal metals. The arrow shows the electron path through the oxide interface.

The objective is to obtain a formula where $I \propto |T_{\mathbf{k}\mathbf{p}}|^2$, so only the first term needs to be retained in the expansion of the S matrix. These steps bring us to the formula

$$I(t) = -ei \int_{-\infty}^t dt' \langle [\dot{N}_L(t), H_T(t')] \rangle \quad (8.319)$$

$$H_T(t') = e^{iH't'} H_T e^{-iH't'} \quad (8.320)$$

$$\dot{N}_L(t) = e^{iH't} \dot{N}_L e^{-iH't} \quad (8.321)$$

where the time dependence of $H_T(t)$ and $\dot{N}_L(t)$ is governed by H' .

An important step in the calculation is to insert the chemical potentials μ_L and μ_R for the two sides of the junction. This insertion must be done with more care than usual, because the chemical potential is not the same on the two sides of the system. The initial Hamiltonian (8.316) has been written to not include the chemical potentials, so the energy is measured on an absolute scale rather than relative to the chemical potentials. However, now insert the chemical potentials into the time developments, so that the energy can be measured with respect to the different chemical potentials on each side of the tunnel junction. The symbols K_R and K_L denote the Hamiltonian with respect to the respective chemical potentials:

$$K_R = H_R - \mu_R N_R \quad (8.322)$$

$$K_L = H_L - \mu_L N_L \quad (8.323)$$

$$K' = K_R + K_L \quad (8.324)$$

For a free-particle system, $K_R = \sum \xi_{\mathbf{k}} C_{\mathbf{k}\sigma}^\dagger C_{\mathbf{k}\sigma}$, while $H_R = \sum \varepsilon_{\mathbf{k}} C_{\mathbf{k}\sigma}^\dagger C_{\mathbf{k}\sigma}$ since $\xi_{\mathbf{k}} = \varepsilon_{\mathbf{k}} - \mu_R$. Since the number operators commute with H' , it is possible to write $H' = K' + \mu_R N_R + \mu_L N_L$ and $\exp(iH't) = \exp(iK't) \exp[it(\mu_L N_L + \mu_R N_R)]$ since the exponentials can be separated when the operators commute. The time development of H_T is

$$\begin{aligned} H_T(t) &= e^{iH't} H_T e^{-iH't} = e^{iK't} [e^{it(\mu_L N_L + \mu_R N_R)} H_T e^{-it(\mu_R N_R + \mu_L N_L)}] e^{-itK'} \\ &= e^{iK't} \sum_{\mathbf{k}\mathbf{p}\sigma} \left[T_{\mathbf{k}\mathbf{p}} e^{it(\mu_R - \mu_L)} C_{\mathbf{k}\sigma}^\dagger C_{\mathbf{p}\sigma} + e^{it(\mu_L - \mu_R)} T_{\mathbf{k}\mathbf{p}}^* C_{\mathbf{p}\sigma}^\dagger C_{\mathbf{k}\sigma} \right] e^{-itK'} \end{aligned}$$

The commutator of H_T with the number operators produces the factor $\mu_L - \mu_R = eV$, which is identified as the applied voltage, as in Fig. 8.12. The applied voltage appears in the correlation function. The correlation function will now be evaluated by assuming that both sides of the junction are in separate thermodynamic equilibrium.

The current operator in (8.319) now becomes

$$\begin{aligned} I(t) &= e \int_{-\infty}^t dt' \left\langle \sum_{\mathbf{k}\mathbf{p}\sigma} [T_{\mathbf{k}\mathbf{p}} e^{-ieVt'} C_{\mathbf{k}\sigma}^\dagger(t) C_{\mathbf{p}\sigma}(t) - T_{\mathbf{k}\mathbf{p}}^* e^{ieVt'} C_{\mathbf{p}\sigma}^\dagger(t) C_{\mathbf{k}\sigma}(t)], \right. \\ &\quad \left. \sum_{\mathbf{k}'\mathbf{p}'\sigma'} [T_{\mathbf{k}'\mathbf{p}'} e^{-ieVt'} C_{\mathbf{k}'\sigma'}^\dagger(t') C_{\mathbf{p}'\sigma'}(t') + T_{\mathbf{k}'\mathbf{p}'}^* e^{ieVt'} C_{\mathbf{p}'\sigma'}^\dagger(t') C_{\mathbf{k}'\sigma'}(t')] \right\rangle \end{aligned}$$

From now on the time development of $C_{\mathbf{k}\sigma}$ operators is governed by $C_{\mathbf{k}\sigma}(t) = e^{iK_R t} C_{\mathbf{k}\sigma} e^{-iK_R t}$ and $C_{\mathbf{p}\sigma}$ operators by $C_{\mathbf{p}\sigma}(t) = e^{iK_L t} C_{\mathbf{p}\sigma} e^{-iK_L t}$. Define the operator A as

$$A(t) = \sum_{\mathbf{k}\mathbf{p}\sigma} T_{\mathbf{k}\mathbf{p}} C_{\mathbf{k}\sigma}^\dagger(t) C_{\mathbf{p}\sigma}(t) \quad (8.325)$$

and the current is written as the summation of two currents

$$I = I_S + I_J \quad (8.326)$$

$$I_S(t) = e \int_{-\infty}^{\infty} dt' \Theta(t-t') \{ e^{ieV(t'-t)} \langle [A(t), A^\dagger(t')] \rangle - e^{ieV(t-t')} \langle [A^\dagger(t), (t')] \rangle \} \quad (8.327)$$

$$I_J(t) = e \int_{-\infty}^{\infty} dt' \Theta(t-t') \{ e^{-ieV(t'+t)} \langle [A(t), A(t'0)] \rangle - e^{-ieV(t+t')} \langle [A^\dagger(t), A^\dagger(t')] \rangle \} \quad (8.328)$$

The term I_S is for single-particle tunneling, which is important for quantum dots. The other term I_J describes the tunneling currents associated with the Josephson effect in superconductors. It is evaluated in Chapter 10.

The terms in I_S have just the right combination of factors to be a retarded Green's function. The integrand depends only on the difference of $(t-t')$, so set $t' = 0$:

$$\bar{U}_{\text{ret}}(t) = -i\Theta(t) \langle [A(t), A^\dagger(0)] \rangle \quad (8.329)$$

$$U_{\text{ret}}(-eV) = \int_{-\infty}^{\infty} dt e^{-ieVt} \bar{U}_{\text{ret}}(t) \quad (8.330)$$

The second term in I_S is the Hermitian conjugate of $U_{\text{ret}}(-eV)$ except for a sign. But the Hermitian conjugate of the retarded function is just the advanced function. The single-particle tunneling current is written as;

$$I_S = ie[U_{\text{ret}}(-eV) - U_{\text{adv}}(-eV)] \quad (8.331)$$

$$= -2e \text{Im}[U_{\text{ret}}(-eV)] \quad (8.332)$$

It is twice the imaginary part of a retarded correlation function, which has the form of a spectral density function. Using the relationship between the Matsubara and the retarded correlation functions discussed in Sec. 3.3, the way to calculate the single-particle tunneling is to evaluate in the Matsubara formalism the correlation function

$$\begin{aligned} U(i\omega) &= - \int_0^\beta d\tau e^{i\omega_n \tau} \langle T_\tau A(\tau) A^\dagger(0) \rangle \\ &= - \sum_{\mathbf{k}\mathbf{p}\sigma} \sum_{\mathbf{k}'\mathbf{p}'\sigma'} T_{\mathbf{k}\mathbf{p}} T_{\mathbf{k}'\mathbf{p}'}^* \int_0^\beta d\tau e^{i\omega_n \tau} \langle T_\tau C_{\mathbf{k}\sigma}^\dagger(\tau) C_{\mathbf{p}\sigma}(\tau) C_{\mathbf{p}'\sigma'}^\dagger C_{\mathbf{k}'\sigma'} \rangle \end{aligned} \quad (8.333)$$

The single-particle tunneling current is just the spectral function of this operator evaluated at the real frequency $-eV/\hbar$, as shown in (8.330). The Matsubara frequency $\omega_n = 2n\pi/\beta$ is boson, since the correlation function has pairs of fermion operators. In the tunneling Hamiltonian, the right- and left-hand sides of the tunneling junction are independent. Then the correlation function factors into a product of the Green's functions for the right and left sides of the junction:

$$\begin{aligned} U(i\omega) &= \sum_{\mathbf{k}\mathbf{p}\sigma} |T_{\mathbf{k}\mathbf{p}}|^2 \int_0^\beta d\tau e^{i\omega_n \tau} \langle T_\tau C_{\mathbf{k}\sigma}(0) C_{\mathbf{k}\sigma}^\dagger(\tau) \rangle \langle T_\tau C_{\mathbf{p}\sigma}(\tau) C_{\mathbf{p}\sigma}^\dagger(0) \rangle \\ &= \sum_{\mathbf{k}\mathbf{p}\sigma} |T_{\mathbf{k}\mathbf{p}}|^2 \int_0^\beta d\tau e^{i\omega_n \tau} \mathcal{G}_R(\mathbf{k}, -\tau) \mathcal{G}_L(\mathbf{p}, \tau) \\ &= \sum_{\mathbf{k}\mathbf{p}\sigma} |T_{\mathbf{k}\mathbf{p}}|^2 \frac{1}{\beta} \sum_{ip} \mathcal{G}_R(\mathbf{k}, ip - i\omega) \mathcal{G}_L(\mathbf{p}, ip) \end{aligned} \quad (8.334)$$

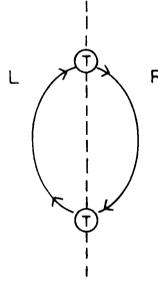


FIGURE 8.13 Feynman diagram for tunneling, where the T s are the vertices which link the right (R) and left (L) sides of the junction.

The only terms which enter are those for $\mathbf{k} = \mathbf{k}'$, $\mathbf{p} = \mathbf{p}'$ and $\sigma = \sigma'$. A type of Feynman diagram is shown in Fig. 8.13. The symbols T in circles are the tunneling vertices. The vertical dashed line is meant to divide the right from the left side of the junction. The solid lines are the interacting Green's functions \mathcal{G}_R and \mathcal{G}_L .

The correlation function is the product of the two Green's functions for the two sides of the junction. Such a simple result is obtained only when one can neglect the vertex diagrams of the correlation function. The argument for neglecting them in tunneling is that vertex corrections require some interaction between electrons on the two sides of the junction, which is improbable. Actually, there is another case where vertex corrections are needed in tunneling, and that is to account for processes where the tunneling electron can excite vibrations or other excitations in the interface. That happens in quantum dot tunneling.

The first example of solving (8.334) will be for the tunneling between two normal metals. The experiments show the tunneling is linear in voltage V for small values of V , and becomes slowly nonlinear at larger values of V . The latter behavior is a failure of the tunneling Hamiltonian, which predicts only linear behavior. The present discussion is limited to small voltages, for which the tunneling Hamiltonian formalism is valid.

For a normal system, consider the electrons to be simple quasiparticles, and approximate the Green's functions by $\mathcal{G}_R^{(0)}$ and $\mathcal{G}_L^{(0)}$. The Matsubara summation is then familiar:

$$\frac{1}{\beta} \sum_{ip} \mathcal{G}_L^{(0)}(\mathbf{p}, ip) \mathcal{G}_R^{(0)}(\mathbf{k}, ip - i\omega) = \frac{n_F(\xi_k) - n_F(\xi_p)}{i\omega + \xi_k - \xi_p} \quad (8.335)$$

$$U(i\omega) = \sum_{\mathbf{kp}\sigma} |T_{\mathbf{kp}}|^2 \frac{n_F(\xi_k) - n_F(\xi_p)}{i\omega + \xi_k - \xi_p} \quad (8.336)$$

$$\begin{aligned} I &= -2e \operatorname{Im}[U_{\text{ret}}(eV)] \\ &= 4\pi e \sum_{\mathbf{kp}} |T_{\mathbf{kp}}|^2 [n_F(\xi_k) - n_F(\xi_p)] \delta(eV + \xi_k - \xi_p) \end{aligned}$$

The energies ξ_k and occupation factor $n_F(\xi_k)$ refer to the right side of the junction, and ξ_p and $n_F(\xi_p)$ refer to the left side. An additional factor of two in front is for the summation over spins. The spin is usually preserved in tunneling. The summations over \mathbf{k} and \mathbf{p} are just over wave vectors. For small voltages, only electrons very near to the Fermi energy are involved in

tunneling. For these small energies, the density of states on both sides is assumed to be a constant:

$$\sum_{\mathbf{k}} \rightarrow \int \frac{d^3k}{(2\pi)^3} \rightarrow N_R \int d\xi_R \quad (8.337)$$

$$\sum_{\mathbf{p}} \rightarrow \int \frac{d^3p}{(2\pi)^3} \rightarrow N_L \int d\xi_L \quad (8.338)$$

$$\begin{aligned} I &= 4\pi e N_R N_L |T|^2 \int d\xi_L \int d\xi_R [n_F(\xi_R) - n_F(\xi_L)] \delta(eV + \xi_R - \xi_L) \\ &= 4\pi e N_R N_L |T|^2 \int d\xi_R [n_F(\xi_R) - n_F(\xi_R + eV)] \end{aligned} \quad (8.339)$$

At zero temperature, the occupation numbers are step functions, so (8.339) becomes

$$I = 4\pi e N_R N_L |T|^2 \int_{-eV}^0 d\xi_R = \sigma_0 V \quad (8.340)$$

$$\sigma_0 = 4\pi e^2 N_R N_L |T|^2 \quad (8.341)$$

The tunneling device behaves as a simple resistor, with a conductance (inverse resistance) given by σ_0 . With a little more care, one can show that the integral in (8.339) also equals eV at nonzero temperatures, since the thermal smearing cancels between $n_F(\xi)$ and $n_F(\xi + eV)$. The result is temperature independent, at least for a range of low temperatures $k_B T \ll E_F$.

A formal expression for the tunneling current can also be derived when interacting Green's functions \mathcal{G}_R , \mathcal{G}_L are retained in the Matsubara summation. This type of summation was evaluated in Sec. 7.1 using the Lehmann representation

$$\begin{aligned} X(i\omega) &= \frac{1}{\beta} \sum_{ip} \mathcal{G}_L(\mathbf{p}, ip) \mathcal{G}_R(\mathbf{k}, ip - i\omega) \\ &= \int \frac{d\varepsilon_L}{2\pi} A_L(\mathbf{p}, \varepsilon_L) \int \frac{d\varepsilon_R}{2\pi} A_R(\mathbf{k}, \varepsilon_R) S \end{aligned} \quad (8.342)$$

$$S = \frac{1}{\beta} \sum_{ip} \frac{1}{ip - \varepsilon_L} \frac{1}{ip - i\omega - \varepsilon_R} = \frac{n_F(\varepsilon_R) - n_F(\varepsilon_L)}{i\omega + \varepsilon_R - \varepsilon_L} \quad (8.343)$$

The retarded function is obtained by the analytical continuation $i\omega \rightarrow eV + i\delta$ and then take the imaginary part.

$$-2 \operatorname{Im}[X_{\text{ret}}(eV)] = \int \frac{d\varepsilon_R}{2\pi} A_L(\mathbf{p}, \varepsilon_R + eV) A_R(\mathbf{k}, \varepsilon_R) [n_F(\varepsilon_R) - n_F(\varepsilon_R + eV)]$$

These steps bring us to the formula for the tunneling current of Schrieffer *et al.* (1963):

$$I = 2e \sum_{\mathbf{kp}} |T_{\mathbf{kp}}|^2 \int \frac{d\varepsilon_R}{2\pi} A_L(\mathbf{p}, \varepsilon_R + eV) A_R(\mathbf{k}, \varepsilon_R) [n_F(\varepsilon_R) - n_F(\varepsilon_R + eV)] \quad (8.344)$$

Again a factor of two is added for the spin summation. The tunneling current is expressed in terms of the spectral functions on the two sides of the junction. Equation (8.344) is the exact formula for I within the model of the tunneling Hamiltonian. The earlier expression (8.339) is recovered with the free-quasiparticle approximation $A_R = 2\pi\delta(\varepsilon_R - \xi_{\mathbf{k}})$ and $A_L = 2\pi\delta(\varepsilon_L - \xi_{\mathbf{p}})$. The virtue of (8.344) is that it contains all many-body effects on the two sides of the junction. Its drawback is that it is based on the tunneling Hamiltonian formalism.

If the tunneling matrix element $T_{\mathbf{k}\mathbf{p}}$ can be approximated as a constant, then the summations over wave vectors give the interacting density of states

$$\sum_{\mathbf{k}} A_R(\mathbf{k}, \varepsilon_R) = 2\pi N_R(\varepsilon_R) \quad (8.345)$$

$$\sum_{\mathbf{p}} A_L(\mathbf{p}, \varepsilon_L) = 2\pi N_L(\varepsilon_L) \quad (8.346)$$

$$I = 4\pi e |T|^2 \int d\varepsilon N_R(\varepsilon) N_L(\varepsilon + eV) [n_F(\varepsilon) - n_F(\varepsilon + eV)]$$

which is a useful formula for describing many-body effects in tunneling which occur on a small energy scale near to the Fermi surface.

A useful formula for quantum dots is to calculate the rate of electron tunneling in each direction: left-to-right (“ LR ”) and right-to-left (“ RL ”):

$$\begin{aligned} I_{RL} &= 2e \sum_{\mathbf{k}\mathbf{p}} |T_{\mathbf{k}\mathbf{p}}|^2 \int \frac{d\varepsilon_R}{2\pi} A_L(\mathbf{p}, \varepsilon_R + eV) A_R(\mathbf{k}, \varepsilon_R) n_F(\varepsilon_R) [1 - n_F(\varepsilon_R + eV)] \\ &\approx 4\pi e |T|^2 N_R N_L \int d\varepsilon n_F(\varepsilon) [1 - n_F(\varepsilon + eV)] = \frac{\sigma_0 V}{1 - e^{-\beta eV}} \end{aligned} \quad (8.347)$$

$$\begin{aligned} I_{LR} &= 2e \sum_{\mathbf{k}\mathbf{p}} |T_{\mathbf{k}\mathbf{p}}|^2 \int \frac{d\varepsilon_R}{2\pi} A_L(\mathbf{p}, \varepsilon_R + eV) A_R(\mathbf{k}, \varepsilon_R) n_F(\varepsilon_R + eV) [1 - n_F(\varepsilon_R)] \\ &\approx 4\pi e |T|^2 N_R N_L \int d\varepsilon n_F(\varepsilon + eV) [1 - n_F(\varepsilon)] = \frac{\sigma_0 V}{e^{\beta eV} - 1} \end{aligned} \quad (8.348)$$

$$I_{RL} - I_{LR} = \sigma_0 V \quad (8.349)$$

The previous result is obtained by subtracting the currents in the two directions.

8.6.2. Quantum dots

Figure 8.14 shows the typical geometry with a quantum dot (QD). There are two metal electrodes on the right and left, and a spherical metal particle between them. The sphere is imbedded in an insulating material through which the electron can tunnel. The tunneling of electrons from the right to the left electrode is facilitated by the QD. The tunneling becomes a two-step process: the electron tunnels from one electrode to the QD, and then tunnels from the QD to the other electrode. There is also coherent tunneling from one electrode to the other through the states of the QD, which act as virtual intermediate states. The first experiments of this type were done by Giaever and Zeller (1968), and they have become popular recently. Now the QD and electrodes might be composed of semiconductors, where the nanostructure is made by lithography.

First examine the properties of the QD. Assume it is a sphere of radius a in atomic units: the dimensional radius $R = aa_0$ where $a_0 = 0.05292$ nm is the Bohr radius. The volume of the sphere is $V_0 = 4\pi R^3/3$. The number of electrons is $N = n_0 V_0$, where the density $n_0 = k_F^3/3\pi^2$ is assumed to be the same as in the bulk metal. The density of states per spin is $N_F = mk_F/(2\pi^2\hbar^2)$. The quantity $N_F V_0$ has the units of inverse energy: it is the number of

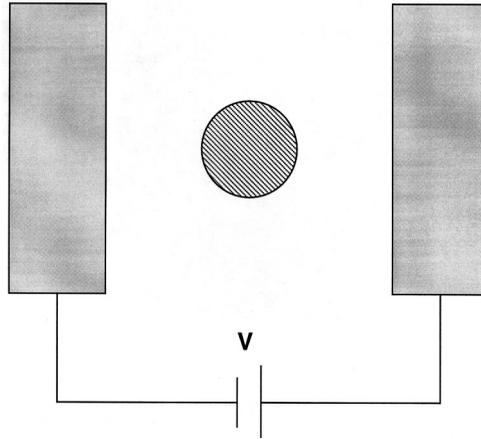


FIGURE 8.14 Geometry for quantum dot tunneling.

electrons per unit energy. Conversely, its inverse is the average separation between energy levels on the QD

$$\Delta E = \frac{1}{N_F V_0} = 3\pi \frac{E_{Fy}}{(k_F a_0)^3} \quad (8.350)$$

$$U = eV_c = \frac{e^2}{2R} = \frac{E_{Fy}}{a} \quad (8.351)$$

The quantity U is called the *charging energy*. It is the energy required to add a charge of e to a neutral metal sphere. All of the charge collects on the surface, and the energy is just $U = e^2/2R$. Table 8.3 compares these two energies as a function of sphere radius. The energies are given in terms of electron volts. The two energy scales are interesting. The charging energy is large compared to the thermal energy $k_B T$, regardless of which temperature is used. If \bar{n} is the number of electrons on the QD for it to be charge neutral, then the energy for n electrons on the QD is

$$E_n = \Delta E(n - \bar{n}) + U(n - \bar{n})^2 \quad (8.352)$$

Table 8.3 Number of electrons N , charging energy U , and energy level separation ΔE of a quantum dot of radius a in atomic units. Energies in units of electron volts. A value of $r_s = 3.96$ (sodium) is assumed

| a | N | U | ΔE |
|-----|------|------|------------|
| 5 | 2 | 2.72 | 2.116 |
| 10 | 16 | 1.36 | 0.265 |
| 15 | 54 | 0.91 | 0.078 |
| 20 | 129 | 0.68 | 0.033 |
| 25 | 252 | 0.54 | 0.017 |
| 30 | 435 | 0.45 | 0.010 |
| 35 | 690 | 0.39 | 0.006 |
| 40 | 1030 | 0.34 | 0.004 |

The system behaves as an Anderson model, in that the energy of the QD depends quadratically upon the total number of electrons n . The energy is for a many-electron system. The energy U is due to electron–electron interactions. A many-electron description is required for treating the QD.

The other energy ΔE is also interesting. It is much smaller than the charging energy for typical systems with $N \sim 10$ –1000 electrons. Here the relevant question is whether ΔE is greater or smaller than the thermal energy $k_B T$. Of course, this comparison depends upon the temperature of the experiment: whether it is at 300 K, 1 K, or 1 mK. If $\beta \Delta E \ll 1$ then the energy levels on the QD can be treated as a continuum, and no error is made by replacing the summation over states by a continuous integral. However, if $\beta \Delta E \gg 1$, the states are well separated in energy compared to $k_B T$. The summation over states must remain a discrete summation. Either condition could apply, depending upon the size of the QD and the temperature of the experiment. Different theories in the literature assume one case or the other. Some simple statistical calculations are presented of the fluctuation of the number of electrons on quantum dots. These fluctuations are important for controlling, or else not controlling, the flow of current onto the quantum dot. These fluctuations play a role in the phenomena called “Coulomb Blockade”. The two different formulas in the literature for the tunneling rate are obtained as two limits of the fluctuations.

The statistical averaging over the states n of the quantum dot includes both the single-electron states $(n - \bar{n})\Delta E$ and the charging energy $U(n - \bar{n})^2$. Sometimes they can be decoupled. There are, in fact, many configurations of one-electron energy states. Denote their quantum numbers by α , so that the energy without the charging energy is E_α .

Calculate the partition function for the electrons on the quantum dot. Assume there are a number of single-electron energy levels with energy E_α as well as a charging energy $U(n - \bar{n})^2$. Let n_α be zero or one, and it denotes whether an electron is, or is not, in the state α . Then the total energy of the system is

$$n = \sum_{\alpha} n_{\alpha} \quad (8.353)$$

$$E = \sum_{\alpha} n_{\alpha} E_{\alpha} + U(n - \bar{n})^2 \quad (8.354)$$

The label α includes the spin and other quantum numbers. Many combinations of states may have the same value of n . In this case, the partition function can be written exactly as

$$Z = \sum_n e^{-\beta U(n - \bar{n})^2} \int_0^{2\pi} \frac{d\theta}{2\pi} e^{in\theta} \prod_{\alpha} \sum_{n_{\alpha}=0}^1 e^{-n_{\alpha}(\beta(E_{\alpha} - \mu) + i\theta)} \quad (8.355)$$

$$= \sum_n e^{-\beta U(n - \bar{n})^2} \int_0^{2\pi} \frac{d\theta}{2\pi} e^{in\theta} S(\theta) \quad (8.356)$$

$$S(\theta) = \prod_{\alpha} [1 + e^{-\beta(E_{\alpha} - \mu) - i\theta}] \quad (8.357)$$

For $\theta = 0$ then $S(0)$ is just the usual Fermi–Dirac distribution function. The integral over θ selects out the terms with a certain value of n . If $U = 0$ then the summation over n gives a delta function at $\theta = 0$ and the partition function reverts to being a usual product of Fermi–Dirac distributions for noninteracting electrons.

The role of the fluctuations is determined by the function $S(\theta)$. This function is evaluated by breaking the summation over α into two parts depending upon whether $\xi_\alpha = E_\alpha - \mu$ is positive or negative, i.e., whether $\exp(-\beta\xi_\alpha)$ is less than or larger than one.

$$\begin{aligned} \ln(S(\theta)) &= \sum_{\xi_\alpha < 0} \ln[e^{-i\theta}(e^{i\theta} + e^{-\beta\xi_\alpha})] + \sum_{\xi_\alpha > 0} \ln \left[(1 + e^{-\beta\xi_\alpha}) \left(1 + \frac{e^{-\beta\xi_\alpha}(e^{-i\theta} - 1)}{1 + e^{-\beta\xi_\alpha}} \right) \right] \\ &= \ln[S(0)] - i\theta\bar{n} + \delta S(\theta) \end{aligned} \quad (8.358)$$

$$\bar{n} = \sum_{\xi_\alpha < 0} 1 \quad (8.359)$$

$$\tilde{n}_\alpha = \frac{1}{e^{\beta\xi_\alpha} + 1} = n_F(\xi_\alpha) \quad (8.360)$$

$$\delta S(\theta) = \sum_{\xi_\alpha < 0} \ln[1 + (1 - \tilde{n}_\alpha)(e^{i\theta} - 1)] + \sum_{\xi_\alpha > 0} \ln[1 + \tilde{n}_\alpha(e^{-i\theta} - 1)] \quad (8.361)$$

where \bar{n} is the number of electrons below the chemical potential, and \tilde{n}_α is the Fermi–Dirac distribution of electrons for the state α .

First evaluate δS for the case that the energy separation between the states of energy E_α is less than the thermal energy $k_B T = 1/\beta$. Then the summation over α can be replaced by a continuous integral. Let $N_F(\xi)V_0 = \rho_D(\xi)$, $\rho_D(0) = 1/\Delta E$ be the density of energy states on the dot. Then one gets that

$$\sum_{\xi_\alpha} = V_0 \int N_F(\xi) d\xi = \int d\xi \rho_D(\xi) \quad (8.362)$$

$$\delta S = \int_{-W}^0 d\xi \rho_D(\xi) \ln[1 + n_F(-\xi)(e^{i\theta} - 1)] + \int_0^W d\xi \rho_D(\xi) \ln[1 + n_F(\xi)(e^{-i\theta} - 1)] \quad (8.363)$$

$$\delta S = - \sum_{l=1} \alpha_l [1 - \cos(l\theta)] \quad (8.364)$$

$$\delta S \approx -2\gamma [1 - \cos(\theta)] \quad (8.365)$$

$$\gamma = k_B T \rho_D(0) \ln 2 = \frac{k_B T}{\Delta E} \ln 2 \quad (8.366)$$

The integral has been approximated by expanding the log function in a Taylor series and keeping the first term. Higher terms ($l > 1$) in the series for δS are much smaller: the series converges rapidly. In general $\gamma > 1$ means that $k_B T > \Delta E$.

With this approximation the partition function can be expressed in closed form

$$Z = Z_{FD} Z_C \quad (8.367)$$

$$Z_{FD} = \prod_\alpha [1 + e^{-\beta\xi_\alpha}] \quad (8.368)$$

$$Z_C = \sum_{\nu} \mathcal{J}_\nu(\gamma) e^{-\beta U \nu^2} \quad (8.369)$$

$$\nu = n - \bar{n} \quad (8.370)$$

$$\mathcal{J}_\nu(\gamma) = \int_0^{2\pi} \frac{d\theta}{2\pi} e^{i\nu\theta - 2\gamma(1 - \cos(\theta))} \quad (8.371)$$

$$= e^{-2\gamma} I_\nu(2\gamma) \approx \frac{1}{2\sqrt{\pi\gamma}} e^{-\nu^2/4\gamma} \quad (8.372)$$

This formula is an important result. It shows that the Fermi–Dirac statistics of the states labeled α are decoupled from the charging of the QD, which is given by the last term Z_C . When $\gamma > 1$, then for small values of v the factor of $\mathcal{J}_v(\gamma)$ is independent of v , and the averaging over the value of n involves only the term $\exp(-\beta U v^2)$. This result is assumed by most theoretical papers. The charging is treated as a separate statistical process from the Fermi–Dirac statistics of the electrons on the QD.

The other case is to evaluate $\delta S(\theta)$ while assuming that the separation between energy levels is greater than the thermal energy. In that case this function is evaluated as

$$\delta S \approx -\gamma^<(1 - e^{i\theta}) - \gamma^>(1 - e^{-i\theta}) \quad (8.373)$$

$$\gamma^< = \sum_{\xi_\alpha < 0} n_F(-\xi_\alpha) \quad (8.374)$$

$$\gamma^> = \sum_{\xi_\alpha > 0} n_F(\xi_\alpha) \quad (8.375)$$

$$\mathcal{J}_v = e^{-(\gamma^< + \gamma^>)} \left(\frac{\gamma^>}{\gamma^<} \right)^{v/2} I_v(2\sqrt{\gamma^< \gamma^>}) \quad (8.376)$$

$$Z = Z_{FD} Z_C \quad (8.377)$$

$$Z_C = \sum_v \mathcal{J}_v e^{-\beta U v^2} \quad (8.378)$$

Here $\gamma < 1$. The formulas for the two cases become identical when $\gamma = \gamma^< = \gamma^>$.

However, if $\gamma < 1$ then \mathcal{J}_v does depend upon γ and this factor should be included in the analysis. Our derivation shall assume that $\gamma > 1$, so that the energy levels on the QD can be considered to be continuous. The thermodynamic averaging over the charging energy is accomplished by including the factor of $\exp(-\beta U n^2)$. In this case, the system is a type of Hubbard model.

8.6.3. Rate equations

The Coulomb blockade is derived using a set of rate equation, which describe how an electron hops on or off the QD. Define f_n as the probability of there being n electrons on the QD. If the number is \bar{n} in equilibrium, then we expect the probability to be something like

$$f_n = \frac{1}{Z_{QD}} \exp[-\beta(n - \bar{n})\Delta - \beta(n - \bar{n})^2 U] \quad (8.379)$$

$$1 = \sum_n f_n \quad (8.380)$$

where Δ is the energy to put a single electron on the QD, irrespective of the charging energy. It is convenient to change the notation and to define $n' = n - \bar{n}$ as the number of excess electrons ($n' > 0$) or holes ($n' < 0$) on the QD. Similarly, the label on the site probability is changed to $f_{n'}$. The subscript denotes the deviations from neutrality.

Define $R_L(m, n)$ as the probability that the number of electrons on the QD changes from n to m by a jump to the electrode to the left. Similarly, $R_R(m, n)$ is the probability of going

from n to m by a jump to the electrode on the right. Since only single jumps are considered, then $m = n \pm 1$. These quantities are given earlier as $I_{RL,LR}$ in Eqn. (8.347) and (8.348).

$$R_L(n+1, n) = \frac{\xi_L E_L(n)}{e^{\beta E_L(n)} - 1} \quad (8.381)$$

$$R_L(n-1, n) = \frac{\xi_L E_L(n-1)}{1 - e^{-\beta E_L(n-1)}} \quad (8.382)$$

$$R_R(n+1, n) = \frac{\xi_R E_R(n)}{e^{\beta E_R(n)} - 1} \quad (8.383)$$

$$R_R(n-1, n) = \frac{\xi_R E_R(n-1)}{1 - e^{-\beta E_R(n-1)}} \quad (8.384)$$

$$\xi_{R,L} = 4\pi |T_{L,R}|^2 N_{L,R} N_{QD} \quad (8.385)$$

$$E_L(n) = eV_{QD} + U(2n+1) \quad (8.386)$$

$$E_R(n) = e(V_{QD} - V) + U(2n+1) \quad (8.387)$$

The left electrode is taken to have zero voltage, while the right one has a voltage of V . The quantum dot has a voltage of V_{QD} . The voltage differences V_{QD} and $V_{QD} - V$ are the differences in the chemical potential between the left electrode and the QD, and between the QD and the right electrode. The factor of $(2n+1) = (n+1)^2 - n^2$ is the difference between the charging energies before and after a hop that adds an electron to the QD. The factors of N_L, N_R, N_{QD} are the density of states at the chemical potential for the left and right electrodes, and for the QD.

As an example, consider the subsystem of just the QD and the left electrode. The rate equation for the process of having electrons hop between the QD and the electrode is

$$\frac{d}{dt} f_n = -f_n [R_L(n+1, n) + R_L(n-1, n)] + f_{n+1} R_L(n, n+1) + f_{n-1} R_L(n, n-1) \quad (8.388)$$

The first term on the right of the equals sign is from events where the system starts in state n and changes by having an electron hop either on to or off of the QD. The remaining two terms are events where the QD starts with either $n \pm 1$ electrons and gets to n by a hop. In equilibrium, the average rate of change is zero. Set the right-hand side of the above equation to zero. This step can be accomplished with the identities

$$f_n R_L(n+1, n) = f_{n+1} R_L(n, n+1) \quad (8.389)$$

$$f_{n-1} R_L(n, n-1) = f_n R_L(n-1, n) \quad (8.390)$$

Adding these two equations gives the right-hand side of (8.388). However, the above two equations are actually identical. The second is the same as the first by changing $n \rightarrow n-1$ everywhere. Only consider the solution to the first equation, which can be manipulated to write it as

$$\frac{f_{n+1}}{f_n} = \frac{R_L(n+1, n)}{R_L(n, n+1)} = \exp[-\beta E_L(n)] \quad (8.391)$$

$$f_n = f_0 \exp[-eV_{QD}n - \beta U n^2] \quad (8.392)$$

The second equation above is the solution deduced from the first. The result is expected for thermal equilibrium.

The case for the QD tunneling is different. There is a voltage V across the device, and a current flowing, so that the system is not in thermal equilibrium. The relationship among the factors of f_n must be derived anew. The rate equation for this case, including both electrodes, is

$$\begin{aligned} \frac{df_n}{dt} = & -f_n[R_L(n+1, n) + R_L(n-1, n) + R_R(n+1, n) \\ & + R_R(n-1, n) + f_{n+1}[R_L(n, n+1) + R_R(n, n+1)] \\ & + f_{n-1}[R_L(n, n-1) + R_R(n, n-1)] \end{aligned} \quad (8.393)$$

In steady state, the average derivative is zero, and the right-hand side vanishes. As in the above case, this constraint is accomplished by

$$\begin{aligned} f_n[R_L(n+1, n) + R_R(n+1, n)] = f_{n+1}[R_L(n, n+1) + R_R(n, n+1)] \\ \frac{f_{n+1}}{f_n} = \frac{R_L(n+1, n) + R_R(n+1, n)}{R_L(n, n+1) + R_R(n, n+1)} \end{aligned} \quad (8.394)$$

The above equation determines the occupation numbers f_n for the quantum dot. Another important equation is for the current. One can count either the electrons leaving the first electrode, or else those arriving at the second one. These two expressions for the current are

$$\begin{aligned} I_L = e \sum_n f_n [R_L(n+1, n) - R_L(n-1, n)] \\ = e \sum_n [f_n [f_n R_L(n+1, n) - f_{n+1} R_L(n, n+1)] \end{aligned} \quad (8.395)$$

$$\begin{aligned} I_R = e \sum_n f_n [R_R(n-1, n) - R_R(n+1, n)] \\ = e \sum_n [f_{n+1} R_R(n, n+1) - f_n R_R(n+1, n)] \end{aligned} \quad (8.396)$$

In the second equal sign we changed $n \rightarrow n+1$ in one of the terms, which is permitted since there is a summation over n . In steady state the two currents are identical: $I_L = I_R$. This identity is automatically satisfied by Eqn. (8.394).

As a simple example, consider the case that the charging energy $U = 0$. In that case the solution to the rate equations is that

$$\xi_L E_L = -\xi_R E_R \quad (8.397)$$

$$\xi_L e V_{QD} = -\xi_R e (V_{QD} - V) \quad (8.398)$$

$$V_{QD} = V \frac{\xi_R}{\xi_R + \xi_L} \quad (8.399)$$

Putting this result into Eqn. (8.394) gives that the ratio $f_{n+1}/f_n = 1$. All of the occupation numbers f_n are the same. An evaluation of the current gives that

$$I = -e^2 V \frac{\xi_L \xi_R}{\xi_L + \xi_R} \quad (8.400)$$

This result is the same as that obtained by two resistors in series. A simple tunnel junction acts like a resistor. A series of two junction behaves as two resistors in series. The net resistance is the sum of the two resistances. In the present problem, the resistance for each connection are $r_{L,R} = 1/(e^2 \xi_{L,R})$. The above formula shows that the net resistance of the two-step tunneling

is the same as adding the two resistances. This result is expected once the energy levels in the QD are treated as continuous, and the charging energy U is neglected.

The Coulomb blockade is obtained by including the charging energy. The example is simplified by assuming that the two electrodes are identical, so $\xi_L = \xi_R$. Furthermore, it is assumed that the QD is exactly midway between the two electrodes. Then one must choose $V_{QD} = V/2$ by symmetry. For each value of V the ratio (8.394) is solved for the relevant number of n -values. Define this ratio as $r(2n + 1)$. By normalizing all occupation numbers to f_0 , we find for the partition function and the current

$$Z_{QD} = f_0 \left\{ 1 + r(1)[1 + r(3)(1 + \dots)] \right. \\ \left. + \frac{1}{r(-1)} \left[1 + \frac{1}{r(-3)}(1 + \dots) \right] \right\} \quad (8.401)$$

$$I = \frac{f_0}{Z_{QD}} \{ R_L(1, 0) - R_L(-1, 0) + r(1)[R_L(2, 1) - R_L(0, 1)] \\ + \frac{1}{r(-1)} [R_L(0, -1) - R_L(-2, -1)] + \dots \} \quad (8.402)$$

A typical result from this formula is shown in Fig. 8.15. The current is an asymmetric function of the voltage. The onset of current begins at $e|V| = 2U$ since the current can only occur if $e|V|/2 > U$. Electrons can only hop from either electrode on to the QD if the voltage difference supplies the charging energy U . Since the QD is in the middle of the space between the electrodes, a voltage of V gives only $V/2$ between the QD and either electrode. Note that there is a second weak threshold when $e|V|/2 = 3U$ at $e|V| = 0.6$ eV. This point is the threshold for having two extra electrons on the QD: the factor of three is $2^2 - 1^2$.

Since the geometry is perfectly symmetric, the current can also proceed by the generation of holes ($n < 0$). In this case, the first step is for an electron to hop from the QD to

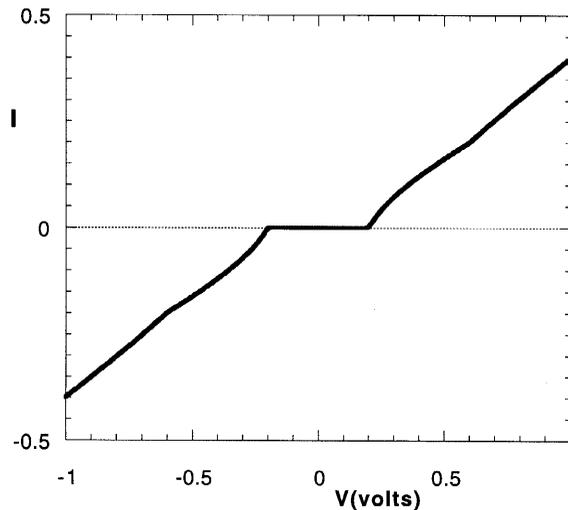


FIGURE 8.15 Current–voltage characteristic for a quantum dot placed symmetrically between two electrodes. $U = 0.1$ eV and $T = 10$ K.

the left electrode, making a hole on the QD. The second step is for an electron to tunnel from the right electrode to the QD, which fills this hole state.

This example is typical for the case that the QD is large enough that its electron states can be considered to be continuous. A much different case, for a small QD, is when the energy levels are large compared to the thermal energy. Then the conductance is a series of steps, where each step is another available energy state on the QD.

An interesting experiment has a magnetic impurity in the QD (Goldhaber-Gordon *et al.*, 1998). The tunneling becomes a type of spectroscopy which can measure the density of states of the magnetic impurity. The experiment allows a test of solutions to the Anderson model.

8.6.4. Quantum conductance

Ballistic transport is the case where an electron carries current without any scattering (Landauer, 1981, 1989). There are several relevant length scales: the length L of the sample, and the mean-free-path λ of the electron. Earlier in this chapter it was shown there are separate values of the mfp for: momentum scattering, energy scattering, electron–electron interactions, etc. Here the most relevant mfp is that for momentum scattering (λ_l), since it enters into the definition of the electrical conductivity. If $L < \lambda_l$ then the current is carried across the material without much chance of scattering. In this case the transport is called *ballistic*.

A special formula is needed for the current from ballistic transport. One cannot employ the Boltzmann equation or the Kubo formula, since they all assume a system with quasi-equilibrium, which means that the mfp is smaller than the dimensions of the sample. Instead, the correct picture is from simple quantum mechanics. Consider a particle moving to the right in one dimension, with a wave vector k . The number of particles per unit distance with wave vector k is

$$\frac{dk}{2\pi} n_F(\xi_k) \quad (8.403)$$

The Fermi–Dirac occupation number is for the source of electron on the left electrode. Assume there is another electrode on the right which is at a voltage V with respect to the one on the left. Then the number of electrons going from right to left is given by

$$\frac{dk}{2\pi} n_D(\xi_k + eV) \quad (8.404)$$

The net current is obtained by subtracting these two formula, and multiplying by ev_k , where v_k is the velocity

$$I = \frac{e}{2\pi} \int dk v_k [n_f(\xi_k) - n_F(\xi_k + eV)] \quad (8.405)$$

$$= \frac{e}{2\pi\hbar} \int d\xi_k [n_f(\xi_k) - n_F(\xi_k + eV)] \quad (8.406)$$

$$I = \sigma_0 V \quad (8.407)$$

$$\sigma_0 = \frac{e^2}{h} \quad (8.408)$$

The quantity e^2/h is the quantized conductance. Its units are interesting. In cgs units, $e^2/\hbar c$ is dimensionless, so that e^2/h has the units of velocity. That is not the correct result here.

Instead, in SI units e has the units of Coulombs, and Planck's constant has the units of Joule-second, which gives that σ_0 has the units of Siemens, which is the inverse of the Ohm. The approximate value is $\sigma_0 = 38.8 \mu\text{S}$.

The above derivation applies to a single channel of electrons. Consider the conduction down a quantum wire. Treat the wire as a type of wave guide, which has many conducting modes, although each new one occurs at higher energy. The number of conducting modes as a function of energy is defined as $N_c(E)$. This quantity is a series of step functions as each new mode opens up along the quantum wire. Then the current is actually given by the expression

$$I = 2\sigma_0 N_c(eV)V \quad (8.409)$$

The factor of two in front is for spin degeneracy. Each orbital mode usually has spin degeneracy. This symmetry is broken in a magnetic field. Some writers prefer to define $2\sigma_0$ as the quantum of conductance.

PROBLEMS

1. Use the piezoelectric electron–Phonon interaction in Sec. 1.3 to calculate the temperature dependence of the electrical conductivity in a semiconductor.

2. Consider the self-energy of an electron from unscreened exchange,

$$\Sigma(\mathbf{p}) = \frac{1}{\beta v} \sum_{\mathbf{q}} v_q \sum_{ip} \mathcal{G}(\mathbf{p} + \mathbf{q}, ip) \quad (8.410)$$

which was evaluated in Sec. 5.1.6. Derive the Ward identity for this self-energy—the equivalent of (8.81). Use this result to show that the Coulomb ladder diagrams have negligible effect upon the basic polarization $P^{(1)}(\mathbf{q}, i\omega)$ in the limit where $\mathbf{q} \rightarrow 0$.

3. The correlation function

$$\chi(\mathbf{q}, i\omega) = - \int_0^\beta d\tau e^{i\omega\tau} \langle T_\tau \rho(\mathbf{q}, \tau) \rho(-\mathbf{q}, 0) \rangle \quad (8.411)$$

vanishes when $\mathbf{q} \rightarrow 0$. Show that the bare bubble $P^{(1)}(\mathbf{q}, i\omega)$ has this feature. Use the Ward identity to show that it still vanishes when self-energy functions and vertex functions are included in the evaluation.

4. Derive the rate of temperature relaxation in a semiconductor due to the deformation potential scattering of electrons by acoustical phonons. Give the answer for high and low temperatures. Evaluate the numerical value of the lifetimes using the data in Table 8.1.

5. Use polar coupling to optical phonons to calculate the rates of relaxation for an electron in a semiconductor: (a) for momentum, (b) for energy, and (c) temperature.

6. Use deformation coupling to optical phonons to calculate the rates of relaxation for an electron in a semiconductor: (a) for momentum, (b) for energy, and (c) for temperature.

7. Use (8.228) to evaluate $\gamma(\varepsilon + i\delta, \varepsilon + i\delta)$, and show that it equals unity, in agreement with the Ward identity.

8. Derive the Ward identity (8.219) for the phonon ladder diagrams.

9. Solve the scalar vertex equation (8.219) using the same techniques used to solve $\gamma(\mathbf{p}, \varepsilon + i\delta, \varepsilon + i\delta)$. Show that $\Gamma(\varepsilon + i\delta, \varepsilon + i\delta) = 1 - \partial\Sigma/\partial\varepsilon$ but that $\Gamma(\varepsilon + i\delta, \varepsilon - i\delta)$ does not obey the Ward identity. Furthermore, show for every solution $R(\varepsilon) = \Gamma(\varepsilon + i\delta, \varepsilon - i\delta)$ to its vertex equation that $R(\varepsilon) + a \text{Im}[\Sigma(\varepsilon)]$ is also a solution, where a is an arbitrary constant.

10. Write down the collision rate $(\partial f/\partial t)$ for electron–electron interactions. Evaluate it for a semiconductor where the few conduction electrons obey Maxwell–Boltzmann statistics. Use Debye screening for the dielectric function.

11. Show that the leading term in the thermopower for a free-electron gas is

$$S = \frac{\pi^2 k_B^2 T}{2eE_F} \quad (8.412)$$

12. Evaluate the correlation function $L^{(22)}$ for the thermal conductivity of a metal. Use the free-electron heat current operator, and show that the integral diverges for the first term in the vertex summation.

13. Solve Eqns. (8.215) and (8.235) in the limit of small temperature and show that the resistivity $\rho \sim T^5$ from phonons. Assume that $\alpha_r^2 F(u) = gu^2$ at small u , where g is a constant.

14. Compare the relaxation time for transport deduced from the force-balance theory (8.198) with that given by (8.236). How do they differ? Show they become identical in the limit of high T .

15. Write down the rate equations for current between two electrodes separated by two identical quantum dots in series. What is the minimum voltage for current to flow at zero temperature?