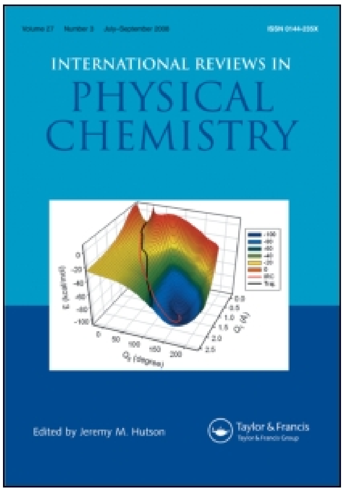

International Reviews in Physical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713724383

The minimum metallic conductivity

N. F. Mott[a]
[a] Cavendish Laboratory, Cambridge, England

To cite this Article Mott, N. F.(1985) 'The minimum metallic conductivity', International Reviews in Physical Chemistry, 4: 1, 1 — 18
To link to this Article: DOI: 10.1080/01442358509353351
URL: http://dx.doi.org/10.1080/01442358509353351

INTERNATIONAL REVIEWS IN PHYSICAL CHEMISTRY, 1985, VOL. 4, NO. 1, 1–18

The minimum metallic conductivity

by N. F. MOTT
Cavendish Laboratory, Madingley Road,
Cambridge CB3 0HE, England

Both experiment and theory show that the conductivity of a degenerate electron gas in a 'dirty metal' tends continuously to zero in the limit of low temperatures, as the Fermi energy approaches a mobility edge. A derivation is given in terms of the Kubo–Greenwood formula. This is not the case when the metal–insulator transition is induced by a magnetic field, and then a minimum metallic conductivity $\sigma_{\min}$ can be observed. Doped crystalline semiconductors and amorphous and liquid semiconductors are discussed in this context. A brief account is given of two-dimensional systems.

1. Introduction

The concept of a minimum metallic conductivity was introduced by the present author more than a decade ago (Mott 1972), though the concept is inherent in papers published earlier (Mott 1967). It deals with the conductivity of systems, such as amorphous metals, in which the mean free path of an electron is short. The concept stated that, if the conductivity is metallic, that is if it tends to a finite value as $T \to 0$, it cannot be less than a certain value $\sigma_{\min}$. $\sigma_{\min}$ is given by the equation

$$\sigma_{\min} = 0{\cdot}03\ e^2/\hbar a \tag{1}$$

the constant is approximate, and a in an amorphous metal would be the distance between atoms, so that $\sigma_{\min} \sim 250\,\Omega^{-1}\,\mathrm{cm}^{-1}$; in the impurity band of a doped semiconductor a is the distance between donors, and $\sigma_{\min}$ probably lies between 10 and $1\,\Omega^{-1}\,\mathrm{cm}^{-1}$.

In spite of a great deal of evidence that this prediction was correct (Mott *et al.* 1975, Mott and Davis 1979), we now know that it is not. Abrahams *et al.* (1979) used a scaling theory to show that the conductivity of a 'metal' at zero temperature would tend to zero continuously as the Fermi energy approached the mobility edge—that is the energy at which electron states there become localized in the sense first described by Anderson (1958). Experimental evidence that this was so was rapidly forthcoming from the extensive investigations of Thomas and co-workers on crystalline silicon doped with phosphorus (Rosenbaum *et al.* 1980. Paalanen *et al.* 1982, Thomas 1983). These authors used temperatures down to 2 mK, and claimed that older work which showed the existence of $\sigma_{\min}$ gave this impression because the temperatures used, above 1 K, were not low enough. A continuous drop to zero was later observed in several other systems, for instance amorphous Si : Nb (Hertel *et al.* 1983) where σ goes continuously to zero as the concentration of Nb is decreased.

A review of the position a year ago was given in this journal by Edwards and Sienko (1983). This article first outlines the arguments for a minimum metallic conductivity and than shows why they were wrong. It also discusses why the concept remains useful.

One reason is that, if the conductivity of an amorphous semiconductor is written

$$\sigma=\sigma_0 \exp\{-(E_c-E_F)/kT\} \tag{2}$$

σ_0 may in certain cases approximate to $\sigma_{\min}$. Here E_c is the mobility edge and E_F the Fermi energy. This is so, particularly, in liquid semiconductors which undergo a metal–insulator transition. Also in doped semiconductors where a metal–insulator transition is caused by the influence of a magnetic field, a minimum metallic conductivity given by (1) is observed. This can now be understood theoretically (Mott 1984 a, b), as we shall see.

2. The model of Anderson (1958)

In his paper which introduced the concept of localization, Anderson (1958) considered an electron moving with the three-dimensional potential energy illustrated in figure 1. A Kronig–Penney array of potential wells was envisaged, leading through the tight binding approximation to the band width B given by

$$B=2zI$$

where z is the coordination number and I the transfer integral

$$I=\int \psi_1 H \psi_2 \, d^3x$$

ψ_1, ψ_2 being wave functions on adjacent sites. In this situation the wave functions are of the form

$$\Psi=\sum_n \exp(i\mathbf{k}\mathbf{a}_n)\psi(\mathbf{r}-\mathbf{a}_n)$$

where a_n are the lattice sites. Anderson then introduced random depths for the wells, spread over a range V_0. This will introduce scattering, leading to a finite mean free path

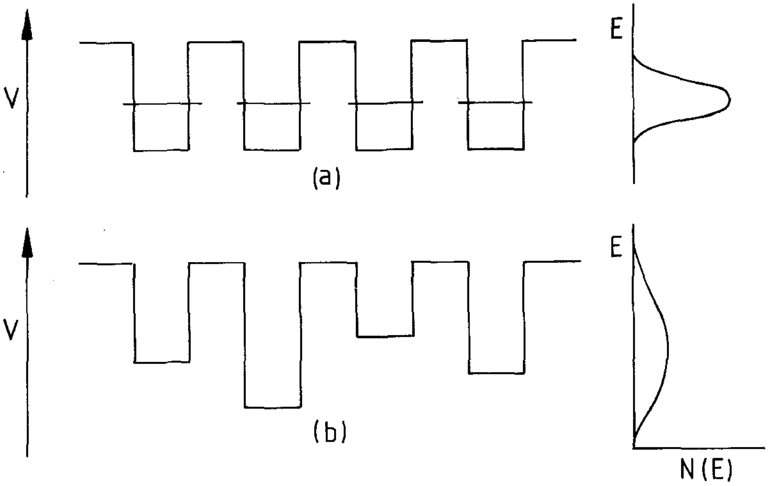

Figure 1. Potential energy function used by Anderson (1958). (*a*) Without random potential. (*b*) With random potential. The density of states is also shown.

l. When $V_0 \simeq B$, use of elementary quantum mechanics shows that $l \simeq a$, the distance between the wells. The wave functions should thus be of the form (with $\psi_n = \psi(\mathbf{r} - \mathbf{a}_n)$)

$$\Psi_{\text{ext}} = \sum_n c_n \exp(i\phi_n)\psi_n \tag{3}$$

where the ϕ_n are random phases and the c_n real numbers. According to a principle introduced by Ioffe and Regel (1960), $l = a$ is the smallest mean free path possible; it is in fact clear that the wave function cannot lose phase memory more quickly than it does according to equation (3).

It is interesting to ask what the conductivity would be if a band of this kind is half filled with a degenerate electron gas, that is with one electron per atom, in a model where the Coulomb repulsion between electrons is neglected. The conductivity of a metal can be written

$$\sigma = ne^2\tau/m$$

where n is the number of electrons per unit volume, $1/a^3$ in our case, and τ is the time of relaxation for electrons with the Fermi energy given by $\tau = l/v_F$; v_F is the velocity of such an electron. We write $v_F = \hbar k_F/m$, where k_F is the wave number and find, with $l = a$

$$\sigma = \frac{e^2}{\hbar a} \bigg/ (k_F a)$$

k_F is given by

$$\frac{4\pi k_F^3}{3} = \frac{1}{2a^3}(2\pi)^3$$

whence $k_F a = (3\pi^2) \simeq 3{\cdot}1$. So approximately

$$\sigma = e^2/3\hbar a \tag{4}$$

If a is a few Ångströms, this Ioffe–Regel conductivity is of order $3000\,\Omega^{-1}\,\text{cm}^{-1}$, though its deduction from the assumption of one electron per atom and a spherical Fermi surface means that in real materials it can be somewhat larger. It looks at first sight as if the conductivity cannot be smaller, and in some systems this seems to be so. Thus in amorphous metals for which the resistivity approaches this value, the temperature coefficient of resistivity is very small, and according to a rule first given by Mooij (1973) it will be *negative* if the resistivity is greater than $170\,\mu\Omega\,\text{cm}$ ($\sigma \sim 6000\,\Omega^{-1}\,\text{cm}^{-1}$). The reason for the negative value is discussed in § 7; but a saturation at a value in the range 3–$6000\,\Omega^{-1}\,\text{cm}^{-1}$ is a well established fact.

However, in the model of Anderson, the conductivity will continue to decrease as V_0 increases beyond the value B, although the mean free path cannot get any shorter. The reason is that the density of states will decrease by a factor g defined in the model by

$$g = N(E_F)/N(E_F)_{V_0=0}$$

This does not occur in liquid metals. For $V_0 > B$, $N(E_F) \simeq 1/a^3(B^2 + V_0^2)^{1/2}$. The conductivity will then be

$$\sigma = \frac{e^2}{3\hbar a} g^2 \tag{5}$$

The reason why g^2 occurs is clear if we calculate the conductivity from the Kubo–Greenwood formula. This amounts to calculating $\sigma(\omega)$, the a.c. conductivity at frequency ω, and letting ω then to zero. $\sigma(\omega)$ is equivalent to the optical absorption coefficient, and will be proportional to

$$\left\{\left|\int \Psi_1^* \frac{\partial}{\partial x} \Psi_2 \, d^3x\right|^2 N(E_1)\, N(E_2)\right\}_{\mathrm{av}}$$

where E_1 and Ψ_1 are energies and wave functions for electrons below the Fermi limiting energy and E_2, Ψ_2 above. The 'av' means an average over all pairs of occupied and empty states separated by an energy $\hbar\omega$. Letting ω tend to zero, this becomes

$$\left|\int \Psi_1^* \frac{\partial}{\partial x} \Psi_2 \, d^3x\right|^2_{\mathrm{av}} \{N(E_{\mathrm{F}})\}^2 \tag{6}$$

where the average is over all states Ψ_1, Ψ_2 at the Fermi surface. This is where the term g^2 comes from in equation (7).

There is some evidence for the reality of the term g^2, mainly from the conductivities of liquid semimetals for which the conductivity falls below the Ioffe–Regel value. There the drop of the density of states may be due to two overlapping bands as in figure 2. Thus in $\mathrm{Te}_{1-x}\mathrm{Tl}_{2+x}$ a plot of the Pauli paramagnetism (which should be proportional to g) against $\sigma^{1/2}$ gives a straight line (Cutler 1977). For liquid tellurium the conductivity in $\Omega^{-1}\,\mathrm{cm}^{-1}$ varies between 1300 at 675 K and 2750 at 1100 K, and the plot of $\sigma^{1/2}$ against the Knight shift gives a straight line. Similar results are obtained for metal-ammonia solutions (Acrivos and Mott 1971).

Anderson's paper of 1958 predicted that at a certain value of V_0/B the states throughout the band would become localized. By this is meant that the wave functions have the form

$$\exp(-r/\xi)\,\mathrm{Re}(\Psi_{\mathrm{ext}}) \tag{7}$$

where Ψ_{ext} is given by (3); and the real part must be taken since Ψ for a non-degenerate localized state must be real. So, if we think of a degenerate gas, than as for the extended (non-localized) case, at zero temperature states are occupied up to a limiting energy E_{F}, but are all localized; at zero temperature the 'gas'—or as Anderson called it, the Fermi glass, cannot carry a current.

The value of V_0/B for which localization will occur throughout the band depends on co-ordination number (z), but for $z = 6$ it is about 2 or rather less†. For this value, the

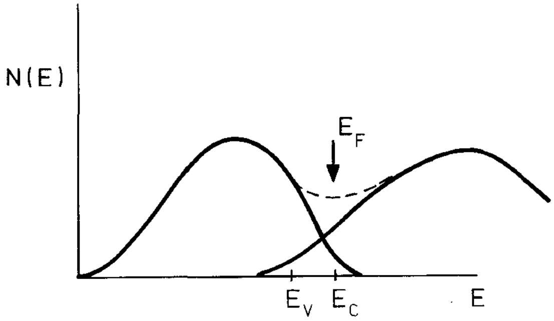

Figure 2. Showing two overlapping bands, pinning the Fermi energy E_{F}. States could be localized between E_{v} and E_{c}.

---

† About 1·6 according to Elyutin *et al.* (1984).

factor g in the middle of the band is estimated to be about $\frac{1}{3}$. Thus according to (5) the minimum metallic conductivity should be of order

$$0{\cdot}03e^2/\hbar a \tag{8}$$

If V_0/B is less than the critical value, it was first pointed out by the present author that states should none the less be localized up to a critical energy E_c, known as the mobility edge, as shown in figure 3. For an energy E_F above E_c, wave functions are extended, that is of form (3); for energies E_F below E_c (as shown in figure 3), they are of the form (7). As E_c-E tends to zero it is believed that ξ, the localization length, tends to infinity, as

$$1/\xi \propto (E_c-E)^s \tag{9}$$

There is an extensive literature on the value of s, but for a theory in which interaction between is neglected $s=1$ seems to us probable, though numerical work indicates slightly larger values.

For a degenerate electron gas, if the Fermi energy lies at E_c, the 'minimum metallic conductivity' should be modified to (Mott 1981)

$$\sigma_{\min}=0{\cdot}03e^2/\hbar a_E \tag{10}$$

where

$$(a/a_E)^3=N(E_c)/N(E)_{av}$$

an equation which would give sensibly the same as (8) unless E_c lies very near the band edge. The 'av' means an average over the whole band.

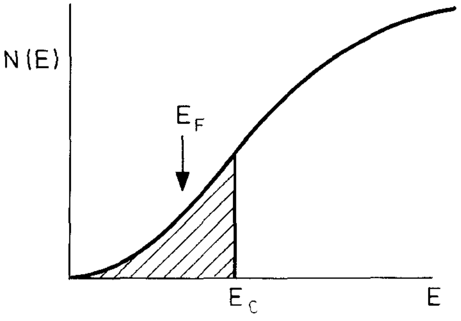

Figure 3. Density of states in a narrow band of a Fermi glass with the Fermi energy shown and localized states shaded.

3. The experimental situation

Although we now believe that the predictions of (8) and (10) are not in general correct, we next discuss the evidence in their favour. The experimental situation closest to the Anderson model is that of a doped and compensated n-type semiconductor. The donors represent the wells of figure 1; they have random depths because of the field of the charged acceptors. They are at random positions in space which introduces additional disorder, the effects of which are discussed in the literature (Debney 1977). An impurity band is formed, as in figure 1. The theory as presented up till now neglects the Coulomb interaction between electrons. But if the band is half full, this will have a major effect; the Hubbard U, defined as the intra-atomic Coulomb interaction, and given by

$$U=\langle e^2/\kappa r_{12}\rangle \tag{11}$$

where κ is the dielectric constant, can split the impurity band in two, the upper and lower Hubbard bands, or as they were called prior to Hubbard's work, the D^0 and D^- bands. In such a case the impurities form an antiferromagnetic insulator up to a certain concentration (n_0). As n increases, the Hubbard bands broaden and when they overlap each other a discontinuous metal–insulator transition (the Mott transition) is predicted, at a concentration

$$n_0^{1/3}a_{\mathrm{H}} \simeq 0{\cdot}26 \tag{12}$$

where a_{H} is the hydrogen radius. We shall return to the Mott transition in §8.

For compensated samples, on the other hand, if the compensation K ($=N_{\mathrm{A}}/N_{\mathrm{D}}$) is considerable, the Hubbard U should not play any major role. An approximation of non-interacting electrons is then a good starting point, though *long-range* Coulomb interactions can also play an important part (cf. §7).

The kind of evidence that, until the 1980s, supported the concept of a minimum metallic conductivity is illustrated schematically in figure 4. This shows the resistivity of a doped semiconductor plotted logarithmically against $1/T$; the different curves are for either a decreasing concentration of donors, or varying uniaxial stress, or increasing magnetic field, which is thought to shrink the orbits and so decrease B, the width of the impurity band. σ_{min} was introduced as both the smallest unactivated conductivity *and* the pre-exponential in (2); they are easily seen to be identical, since both should be given by $eN(E_{\mathrm{c}})\,kT\mu_{\mathrm{c}}$, where μ_{c} ($=eD/kT$) is the mobility for a carrier with energy E_{c}. D is here the diffusion coefficient at E_{c}.

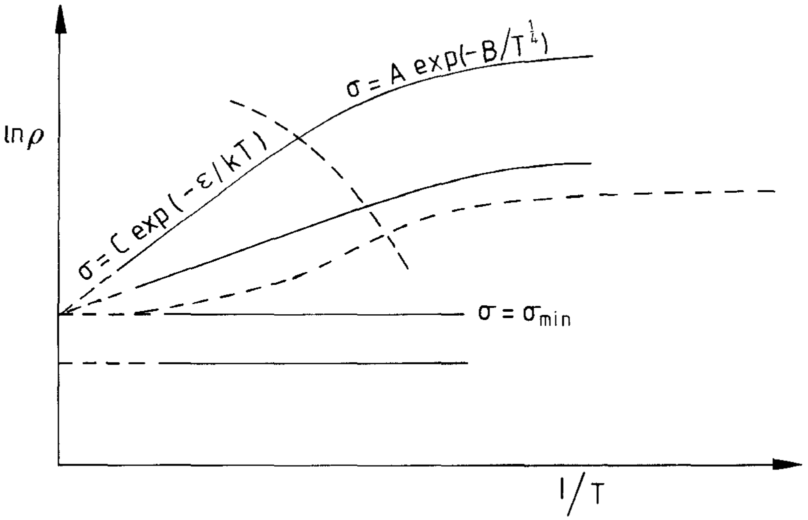

Figure 4. Resistivity of a non-crystalline metal as a function of $1/T$, when a metal–insulator transition is induced by change of composition, stress or magnetic field. In cases where no minimum metallic conductivity exists the behaviour should be as shown by the dotted line.

At low temperatures, conduction will be by hopping from one centre to another; the relationship

$$\sigma = A \exp(-B/T^{1/4}) \tag{13}$$

is deduced from a theory of non-interacting electrons and often observed, though important modifications are introduced when long-range Coulomb interactions are taken into account (Efros and Shklovskii 1975, Davies *et al.* 1982, 1984) which have not with certainty been observed. This article will not review the problem of hopping conduction.

Perhaps the most impressive example of this kind is the work of Biskupski (1982) and co-workers (Biskupski *et al.* 1981) on *n*-type compensated InP where the transition was induced by a magnetic field. These results are shown in figure 5. Since two variables, field and donor concentration, are involved, it was possible to plot $\sigma_{\min}$ against a; the results are shown in figure 6, giving excellent agreement with (8). The original work used temperatures only down to 1·3 K, but both Biskupski *et al.* (1984) and Long and Pepper (1984) have extended it down to 40 mK; the experiments show clearly that for this system $\sigma_{\min}$ does exist.

On the other hand the classic experiments of Thomas and co-workers on Si:P (Rosenbaum *et al.* 1980, Thomas 1983) using temperatures down to 2 mK, show equally clearly that in this system σ ($T=0$) does go continuously to zero, apparently as $(n-n_c)^s$ with $s \simeq \frac{1}{2}$; these experiments are reviewed by Edwards and Sienko (1983) in this journal. In a-Si–Nb, with decreasing concentration of Nb, σ ($T=0$) goes to zero at 11·5% Nb as $(n-n_c)^s$ with $s=1$ (Hertel *et al.* 1983) and there are now many examples of this behaviour. Here n_c is the critical concentration.

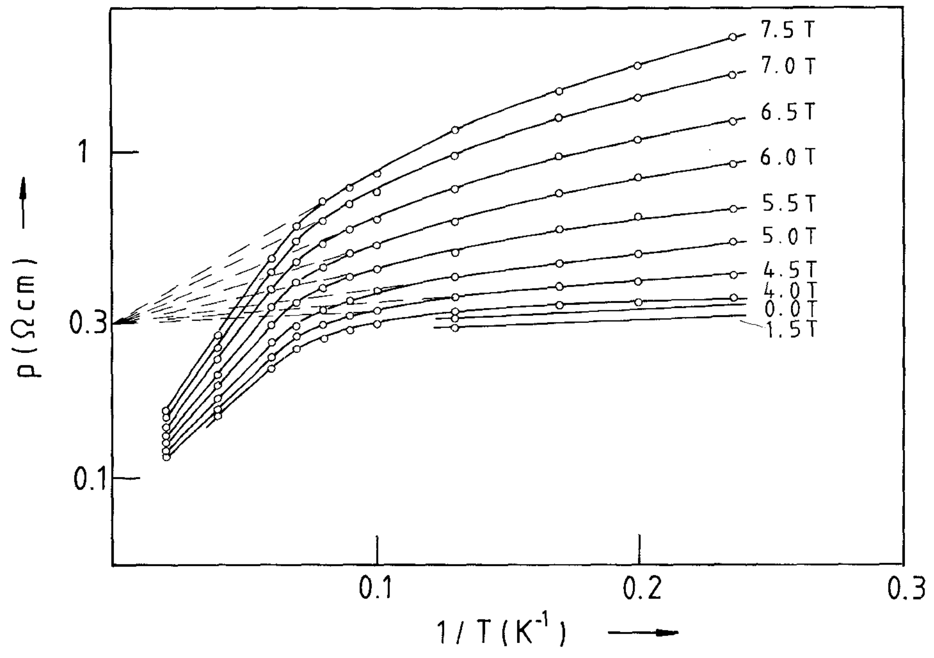

Figure 5. Resistivity of InP as a function of $1/T$, for various magnetic fields (in Tesla) (Biskupski 1982).

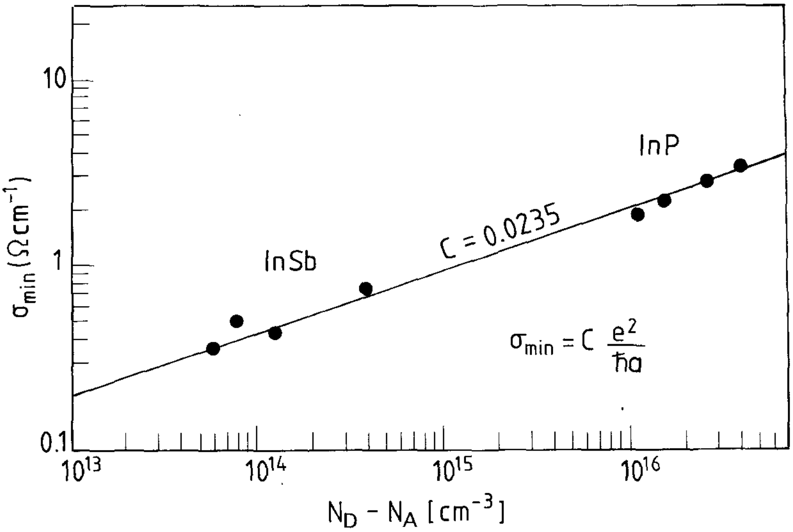

Figure 6. σ_{min} deduced from the results of figure 5 (Biskupski 1982) and similar work.

4. Theory of the conductivity at the Anderson transition

We have to explain, then, that σ ($T=0$) will in general go continuously to zero if composition or stress are changed, that σ_{min} is often a good approximation to the pre-exponential in (1), and that σ_{min} can exist when the transition is induced by a magnetic field.

The scaling theory uses the concept introduced by Thouless (1977) of the conductivity $\sigma(L)$ of a cube of side L As E_F is changed, no discontinuity would be expected in σ for finite L, and it was argued that $\sigma(\infty)$ cannot change discontinuously either. We believe that this is correct only in the absence of a magnetic field—as has been argued by others. This can be seen by examining the Kubo–Greenwood expression (6).

If the functions Ψ are localized, the conductivity (6) must vanish because the states are in general non-degenerate and if Ψ_1, Ψ_2 overlap, characteristic functions

$$a\Psi_1 + b\Psi_2$$

$$b\Psi_1 - a\Psi_2$$

will form, and these cannot have the same energy. But if Ψ_1 and Ψ_2 are extended and degenerate, it is difficult to see how the average of a squared term can vanish. Kaveh and Mott in a series of papers (Kaveh and Mott 1983, Mott and Kaveh 1983) proposed that the functions should be of the form

$$A\Psi_1^{ext} + B(a/r)^2\Psi_2^{ext} \tag{14}$$

that is made up of power-law localized *and* extended states; normalization gives $|A|^2 + |B|^2 = 1$ and at the transition $A \rightarrow 0$. The status of this theory is perhaps uncertain, but the present author believes that it is *not* relevant to the behaviour near n_c. Starting from the random phase wave function (3), it is possible to show, by considering the correlation between the phases ϕ_n in Ψ_1, Ψ_2, that $\sigma \to 0$. The argument (Mott 1984 a, b) is as follows:

As we have seen, for an energy ΔE just below the mobility edge, the wave functions in the absence of a magnetic field are real, and cannot overlap much. If two wave functions are centred at a distance R the transfer integral will be of the form

$$H_0 \exp(-R/\xi)$$

and this must be less than ΔE, so two states cannot be closer than a distance given by

$$R = \xi \ln(H_0/\Delta E)$$

Putting

$$\xi = a/(\Delta E/E)^s$$

this gives

$$R = \xi \ln\{(H_0/E)(a/\xi)^{1/s}\} \tag{15}$$

The essential point now is that just *above* the mobility edge, the functions within a volume R^3 will be little changed. Thus, within this volume:

(*a*) The functions Ψ_1, Ψ_2 remain real and all are identical with (7).

(*b*) They have long-range fluctuation of wavelength R given by (15), tending to infinity as $E \to E_c$. The 'participation ratio'

$$\left\{\Omega \int |\Psi|^4 \, d^3x\right\}^{-1}$$

tends to zero as $E \to E_c$. Since all the maxima are at the same positions in space, and σ according to (*b*) depends on the fourth power of Ψ, this property will multiply σ by

$$\{\exp(R/\xi)\}^2 \tag{16}$$

However, from (*a*) above we see that

$$\int_{R^3} \Psi_1^* \frac{\partial}{\partial x} \Psi_2 \, d^3x = \frac{1}{2} \int dy \, dz [\Psi^2]_x^{x+R}$$

apart from some constant factor $\exp(i\phi)$. x and $x+R$ are supposed to be the values at the minima of Ψ in figure 7. Here the phases are random, so that both for the y, z

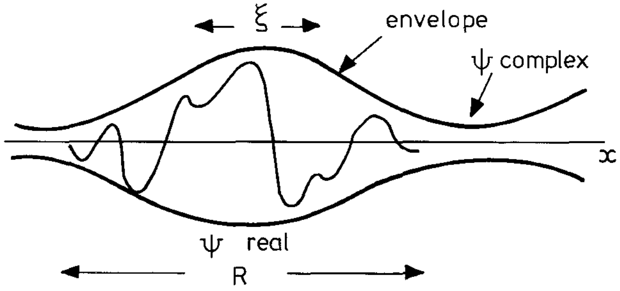

Figure 7. Sketch of extended wave function near the Anderson transition. For an energy in the range ΔE, all functions Ψ have peaks at same point, and differ only in the complex region.

integration and for the addition of the next term $[\Psi^2]_{x+R}^{x+2R}$ we can use the random phase approximation. So our conclusion that $\Psi_1^* = \Psi_2$ reduces the integral by $\surd(a/\xi)$, the region of space between $\frac{1}{2}\xi$ and $\frac{1}{2}R$ from the maximum contributing a negligible amount. Since $|\Psi|^2$ must be taken at a distance $\frac{1}{2}R$ from the maxima, a term is introduced which just cancels (16). Thus for energies ΔE above E_c

$$\begin{aligned}\sigma &= 0{\cdot}03(e^2/\hbar a)(a/\xi)\\ &= 0{\cdot}03e^2/\hbar\xi \qquad (17)\end{aligned}$$

where ξ is the localization length for electrons with energy Δ E *below* E_c. If ξ varies with $(E-E_c)^{-s}$ with $s=1$, a linear drop of σ with concentration can be explained.

In the presence of a magnetic field H, the localized wave functions are no longer real; there should be a phase change $\exp(i\pi a/L_H)$ on going from well to well, where

$$L_H = (c\hbar/He)^{1/2} \qquad (18)$$

The argument, therefore, is not applicable in a magnetic field.

We now ask what happens at finite temperature. The scaling theory introduces the concept of the conductivity $\sigma(L)$ for a cube of size L, and argues that this cannot show a discontinuity. Following Abrahams *et al.* (1979) we introduce the inelastic diffusion length L_i, which is the distance that an electron diffuses before an inelastic collision, so that

$$L_i = \surd(D\tau_i) \qquad (19)$$

where τ_i is the time before an inelastic collision. A finite value of τ_i introduces an uncertainty $\hbar/\tau_i$ into the energy, so we cannot say exactly whether there is a discontinuity or not. We now assert that for practical purposes, the conductivity when E lies at E_c is

$$\sigma = 0{\cdot}03e^2/\hbar L_i \qquad (20)$$

and as E drops below E_c there is a rapid drop in σ, though there will be no discontinuity. In a theory of non-interacting electrons (20) should give both the conductivity when E_F lies at E_c (or just above) and the preexponential factor in (2).

In the case of a magnetic field, we write

$$\sigma = 0{\cdot}03e^2/\hbar L_H \qquad (21)$$

there is no uncertainty in the energy and, we believe, a discontinuity in σ *does* occur at E_c. In terms of scaling theory, in a magnetic field there is no one-parameter scaling function. If $L_H \leqslant a$, then we expect

$$\sigma_{\min} = 0{\cdot}03e^2/\hbar a$$

the author's original expression. This accounts for the results of a Biskupski and of Long and Pepper on the conductivity of InP. This is in fact almost obvious; we have already seen that with the Anderson model there is a phase change of $\pi a/L_H$ in going from centre to centre, and for a random array with a *mean* distance a between each centre the assumption that the localized wave functions are real is completely invalid.

We now discuss the quantity L_i, the inelastic diffusion length. This can be a consequence *either* of electron–electron collisions or collisions with phonons. In low temperature work on doped semiconductors the former will be responsible. For a metallic sample, we suppose that the time τ before an electron–electron collision is given by (Ziman 1961).

$$\frac{1}{\tau}=\frac{1}{\tau_0}\left(\frac{kT}{E_F}\right)^2$$

where $\tau_0 \sim l/v_F$. The inelastic diffusion length is given by (19), namely

$$L_i=\sqrt{(D\tau)} \tag{22}$$

where D is the diffusion coefficient. Far from the mobility edge one might write $D=\frac{1}{6}l^2/\tau_0$, but if σ is reduced by the factor a/L_i, so must be D. We write then, with $l=a$

$$L_i=a(E_F/kT)(a/6L_i)^{1/2}$$

or

$$(L_i/a)^3=\tfrac{1}{6}(E_F/kT)^2 \tag{23}$$

This formula shows that L_i approaches a as the temperature rises, which perhaps could explain why the experiments prior to 1980 at temperatures above 1 K gave the lowest unactivated conductivity as σ_{min}—though we should expect the value to drop as $T^{2/3}$ at low T.

In a similar way we can consider the pre-exponential factor σ_0 in (2) when E_F lies below E_c. The lifetime of an electron at the mobility edge will be of the form

$$\frac{1}{\tau}\sim\frac{2\pi}{\hbar}|H|^2\{N(E_F)\}^2(E_c-E_F)$$

where $N(E)$ is the density of states per atom and H the matrix element of the screened electron–electron interaction. From a similar calculation (Mott 1984 a, b) we find

$$L_i^3=\tfrac{1}{6}a^3\tau/\tau_0,$$

where τ_0 is the elastic time of relaxation. We argue that as E_c-E_F increases, L_i will approach the value a. We do not, however, consider that the success of the assumption that σ_{min} is the pre-exponential factor is fully understood.

5. The pre-exponential factor in the conductivity of non-crystalline semiconductors

Our account of the theory up till now has been for non-interacting electrons. In discussions of the pre-exponential factor this should be sufficient. We have to consider the behaviour of a single electron (or hole) at a mobility edge; its interaction with electrons at the Fermi level can determine its lifetime, as we have seen in the last section, but apart from such Auger collisions electron–electron interaction is not important. In this section, then, we neglect this interaction and consider the pre-exponential factor in amorphous and liquid semiconductors.

We consider first the important case of hydrogenated amorphous silicon. Here the density of states at E_F is small, except perhaps for heavily doped specimens, so L_i will certainly be determined by interaction with phonons. The present author (1984 b) from experimental and theoretical considerations estimates it as

$$L_i \sim 30\ \text{Å}$$

weakly dependent on T. σ_0 should thus be about $16\ \Omega^{-1}\ \text{cm}^{-1}$. For the conductivity at a mobility edge the pre-exponential factor is greatly affected by the variation with temperature of $E_F - E_c$ (Overhof and Beyer 1983), and the main interest of this result is for the calculation of the drift mobility μ_D in photo cells. For this we expect

$$\mu_D = \mu_0 \exp(-\Delta E/kT)$$

where ΔE is an activation energy and

$$\mu_0 = \mu_c N(E_c)/N(E_A)$$

here E_A is the lowest energy in the conduction band and μ_c the mobility at E_c, related to σ_0 by

$$\sigma_0 = eN(E_c)kT\mu_c \tag{24}$$

Thus

$$\sigma_0 = eN(E_A)kT\mu_0$$

Data both from Spear (1983) and Tiedje and Rose (1981) give $\mu_0 \sim 100\,\mathrm{cm}^2/\mathrm{V\,s}$ (and $\Delta E = 0{\cdot}13\,\mathrm{eV}$); the value of $N(E_c)$ is more doubtful, but we take a value given by Spear ($0{\cdot}5 \times 10^{20}\,\mathrm{cm}^{-3}\,\mathrm{eV}^{-1}$), which leads for σ_0 to $\sim 50\,\Omega^{-1}\,\mathrm{cm}^{-1}$.

Many of these models are controversial. In particular, Cohen *et al.* (1983) have proposed that in the conduction band a polaron is formed, self-trapping occurring on every Si–Si bond. They also argue that self-trapping of this kind *must* occur if $\mu_c \to 0$ as $E \to E_c$. Arguments against this conclusion are given by Mott (1984 b).

The concept of $\sigma_{\min}$ has often been applied to liquids that undergo a metal–insulator transition, it being argued that if σ falls below $\sim 200\,\Omega^{-1}\,\mathrm{cm}^{-1}$, conduction is not metallic. Whether this is so or not will depend on our estimate of L_i, which will be determined by collisions with phonons *and* electron–electron collisions. As regards the former, although all collisions with electrons are inelastic, the momentum change is so small that it takes ~ 100 collisions to produce the incoherence between waves that determines L_i.

In liquid semiconductors E_F may vary rapidly with T, with a large effect on the pre-exponential factor. However, values of σ_0 can be obtained from the thermopower S, writing

$$\sigma = \sigma_0 \exp\{-(E_c - E_F)/kT\}$$

$$S = (k/e)\{(E_c - E_F)/kT + 1\}$$

so that, eliminating $E_c - E_F$, we have

$$\ln \sigma = \ln \sigma_0 + (k/e)S - 1 \tag{25}$$

A linear plot of σ against S with slope k/e, as the temperature is varied, is widely observed, and a value of σ_0 can be obtained. Thus, from observations of Schmutzler and Hensel (1972) for expanded fluid mercury the value is $\sim 20\,\Omega^{-1}\,\mathrm{cm}^{-1}$ but for caesium about $300\,\Omega^{-1}\,\mathrm{cm}^{-1}$. Also at the transition Mott and Davis (1979) predict that

$$|S| = (k/e)2\ln 2 \tag{26}$$

and using this result they find for caesium $\sigma_0 \sim 300\,\Omega^{-1}\,\mathrm{cm}^{-1}$.

We are led to conclude that $\sigma_{\min}$ is a fair approximation to the conductivity of fluids near the metal–insulator transition. In a paper to be published in *Phil. Mag.*, Mott has examined the problem and finds that $L_i \sim a$ is likely to be a good approximation for high resistivity liquids.

6. Quantum corrections far from the transition

For the conductivity of a metal the first effect of correlations between wave functions is given by the expression

$$\sigma=\sigma_{\text{Boltzmann}}\left[1-\frac{3}{(k_{\mathrm{F}}l)^2}\left\{1-\frac{l}{L}\right\}\right] \tag{27}$$

where L is the size of the specimen, or L_{i} or L_{H}, whichever is the smaller. This is valid for any mean free path l, even if $k_{\mathrm{F}}l$ is large. It was obtained by diagrammatic methods by Gorkov *et al.* (1979), Kawabata (1981) and Shapiro and Abrahams (1981), and a physically appealing method (with 4·5 instead of 3) by Bergmann (1983). It applies to electrons in a single-valley conduction band or an impurity band. The term l/L_{i} will always lead to a correction to the conductivity of the form σ_0+AT, if L_{i} results from electron–electron collisions.

It is of interest to extrapolate (27) to the metal–insulator transition. At the Ioffe–Regal limit for a half-full impurity band, $k_{\mathrm{F}}l=k_{\mathrm{F}}a\simeq\pi$ so σ cannot vanish. This is expected, because for Anderson localization the density of states must decrease, by about $g=\frac{1}{3}$. If g is not unity, (27) should be replaced by (with $\sigma_{\text{Boltzmann}}=\frac{1}{3}e^2/\hbar a$)

$$\sigma=(\tfrac{1}{3}e^2g^2/\hbar a)\left[1-\frac{c}{(k_{\mathrm{F}}a)^2g^2}\left(1-\frac{a}{L}\right)\right] \tag{28}$$

with c of order unity (Kaveh, *Phil. Mag.*, to be published). For $L=\infty$, this does in fact vanish for $g\simeq\frac{1}{3}$, the familiar Anderson criterion, which gives us some confidence in the formula. Also for finite L we find at $E=E_{\mathrm{c}}$

$$\sigma=e^2g^2/3\hbar L \tag{29}$$

which is the same formula that we obtained by other methods. Since for $L=\infty$ σ goes linearly to zero, we suppose $s=1$.

In the work of Thomas (1983) and co-workers on Si : P, in which the conductivity goes continuously to zero as the concentration n decreases to n_{c}, the measured specific heat (Sasaka 1980) conforms approximately to the value expected for a degenerate gas of electrons in the conduction band. One has to conclude that the transition takes place *in the conduction band.* This—we believe—could not occur in a single-valley semiconductor. Berggren (1982) first pointed out that for (spherical) valleys, k_{F}^2 would be reduced by $1/x^{2/3}$, so for six valleys ($x=6$) σ *can* drop to zero in the conduction band. We believe that this behaviour may be unique to silicon.

The linear increase in the conductivity described above has been observed in amorphous metals, as has the term in $T^{1/2}$ described in the next section, and may account for the Mooij rule discussed in §2.

7. Long-range Coulomb interactions

In this section we discuss the effects of *long-range* Coulomb interactions between electrons; our discussion is relevant to the impurity band of a compensated semiconductor, to a degenerate electron gas in a conduction band and in fact to any metal *except* a narrow half-filled band, where the Hubbard U is all-important. This is considered in our final section.

As one might expect, the Coulomb repulsion decreases the density of states at the Fermi level, and the effect is the greater the shorter the mean free path. Altshuler and

Aronov (1979) were the first to show that a cusp is to be expected in the plot of $N(E)$ against as in figure 8, the depth of the cusp being given by

$$\delta N(E)/N(E) \simeq 0{\cdot}025(mD/\hbar) \qquad (30)$$

where D is the diffusion coefficient. Since $D \to 0$ as $E \to E_c$ this expression diverges, and is only valid when $E - E_c$ is not too small. It leads:

(1) To a factor $(1-\delta N/N)^2$, reducing the conductivity near the transition.
(2) A temperature-dependent term in the conductivity

$$\sigma = \sigma_0(1 + AT^{1/2}) \qquad (31)$$

A is normally positive near the transition but can have either sign. Clearly at low temperatures it will predominate over terms varying as T or T^2, giving a larger correction than the linear term resulting from equation (27) and still more the Landau–Baber term CT^2 resulting from electron–electron collisions.

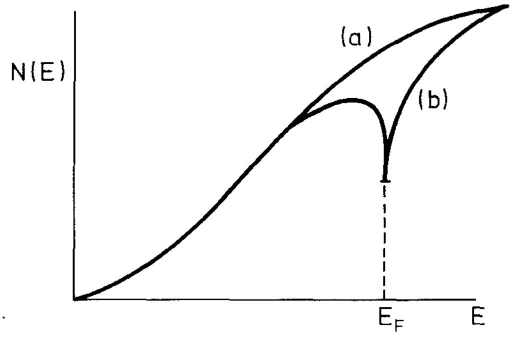

Figure 8. Cusp in the density of states of a 'dirty' metal resulting from long-range Coulomb interaction, (*a*) is without long-range Coulomb interaction, (*b*) taking this into account. The correction giving the cusp can have either sign.

The magnitude of the term varies from one material to another in a way that has not been satisfactorily explained. There is an interesting contrast between InP and InSb. As we have seen, a metal–insulator transition is observed in these materials under the influence of a strong magnetic field. In InP the term $AT^{1/2}$ is very small, and as we have seen a value of $\sigma_{min} \sim 0{\cdot}03e^2/\hbar a$ is observed. In InSb, for reasons not at present understood, the term is much bigger, and according to measurements by Mansfield *et al.* (1984), σ_{min}, if it exists, is much smaller. This is shown in figure 9.

In Si : P Thomas and co-workers (see Thomas 1983), inducing the transition by uniaxial stress, have shown that at very low temperatures

$$\sigma \propto (n - n_c)^{1/2}$$

instead of the linear behaviour observed in a-Si–Nb and other similar systems. At the same time they find that the dielectric constant κ increases rapidly as

$$1/\kappa \sim (n_c - n)^{\nu}, \quad \nu \simeq 1$$

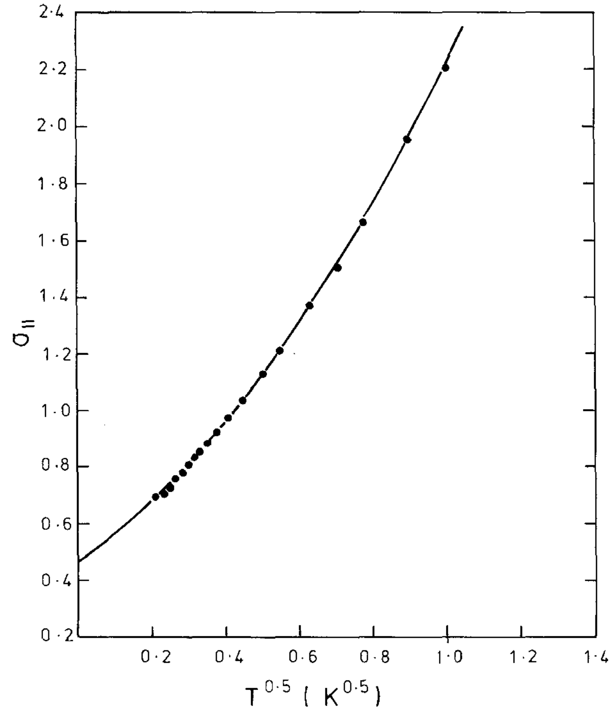

Figure 9. Longitudinal conductivity in mho cm^{-1} of just metallic InSb, plotted against $T^{1/2}$ at a magnetic field of 6 T. The MI transition occurs at 6·6 T. The solid line is a least squares fit to $\sigma(T)=\sigma(0)+AT^{1/2}$ (Mansfield *et al.* 1984).

Since for a Fermi glass we expect (see below)

$$\kappa \propto \xi^2$$

and, as we have shown, $\sigma \propto 1/\xi$, we have to deduce that

$$\xi \propto (E_c - E)^s \quad \text{and} \quad (E - E_c)^s$$

with $s=\frac{1}{2}$ for these systems. This is believed to be another result of long-range Coulomb interaction. An account of how this can happen is given by Grest and Lee (1983).

The expression $\kappa \propto \xi^2$ comes from the dimensional equation, which states that the polarizibility is proportional to $e^2N(E_F)\ \xi^2$. This neglects any effect of the Clausius–Mossotti equation

$$\kappa = 1 + \frac{4\pi N\alpha}{1-4\pi N\alpha/3} \tag{33}$$

on the dielectric constant near the transition; here α is the polarizibility. This term and its influence on metallic properties is discussed by Edwards and Sienko (1983). It should in our view be important only when the orbitals of atoms (or localized states) do not greatly overlap. As pointed out by Mott and Gurney (1940, p. 17) it is not successful in describing the dielectric constants of the alkali halides and this was ascribed to overlap. We think that in the non-crystalline systems discussed here, the denominator in (33) will never vanish, for the following reason. The hydrogen radius $\hbar^2\kappa/me^2$ will increase according to equation (12) as κ increases, but overlap must occur before κ becomes infinite, so that (33) breaks down.

The dielectric catastrophe just before the metal–insulator transition occurs in some fluid systems, for instance solutions of sodium in ammonia and liquid K–KCl (Edwards and Sienko 1981); it is suggested that the cause is similar.

8. The Mott transition

For an array of one-electron centres, where a transition occurs through the overlap of the two Hubbard bands, a discontinuous (first order) transition is predicted. This could disappear through the effect of disorder (Mott 1978), but clearly does not in systems such as metal–ammonia or expanded fluid caesium, if as postulated by the author the critical point is a consequence of the transition.

At temperatures above the critical point, two Hubbard bands will gradually overlap as the concentration increases until E_c and E_F coincide. Since a conductivity $\sigma_{\min} \sim 100\text{–}200\,\Omega^{-1}\,\text{cm}^{-1}$ has been quoted in favour of the transition being the cause of the critical point, we have to ask whether this is now justified. It should only be justified if $L_i \sim a$. Fortunately for caesium we have direct evidence that $\sigma \sim \sigma_{\min}$, from the method outlined in § 5; a small L_i could be a consequence of the strong interaction with spins. We think that, within a factor 2 or 3, $\sigma = \sigma_{\min}$ is likely.

In doped uncompensated semiconductors where the transition takes place in an impurity band (thus excluding Si : P), we think the disorder *must* be great enough to wipe out any discontinuity in σ. Two Hubbard bands begin to overlap at a concentration given by the Mott criterion (12), but the transition is now of Anderson type. If the transition is *not* produced by a magnetic field, we expect $\sigma\,(T=0)$ to increase as $n-n_c$ or $(n-n_c)^{1/2}$.

9. Two-dimensional systems

The inversion layer at the interface between crystalline silicon and silicon dioxide has been extensively investigated; the density of electrons and thus at low temperatures the Fermi energy can be changed in an MOS device by changing the gate voltage, and an 'Anderson transition' observed through variation of the source–drain current with temperature. Early work (reviewed by Mott *et al.* 1975, Pepper 1977) showed an Anderson transition between activated conduction and a value of the conductance almost temperature-independent and equal under suitable conditions to the theoretical minimum metal conductance, for which theory gives $0{\cdot}1e^2/\hbar$. The constant depends on the condition for Anderson localization in two dimensions. However Abrahams *et al.* (1979) were the first to point out that in two dimensions there can be no true metal and all states are weakly localized. This leads to a correction to the 'metallic' conductivity of the form

$$\sigma = \sigma_0 + A \ln T$$

which is widely observed. The behaviour at very low temperatures is controversial. According to the scaling theory all states are exponentially localized, with a localization length

$$\xi = l \exp\left(\tfrac{1}{2}\pi k_F l\right)$$

and the transition from strongly to weakly activated conduction, though rapid, is not sharp. Kaveh and Mott (1981) on the other hand suggest that there is a sharp mobility edge at which $\xi \to \infty$, and that for energies above this the functions tend to zero as $1/r$. This does of course conflict with the scaling theory, but there is experimental evidence that for two-dimensions the theory is incorrect (Davies *et al.* 1983).

Acknowledgments

I am indebted to M. Kaveh and M. Pepper for many discussions about the subject matter of this paper, and to R. Mansfield for permission to reproduce the results of figure 9.

References

Abrahams, E., Anderson, P. W., Licciardello, D. C., and Ramakrishnan, T. W., 1979, *Phys. Rev. Lett.*, **42,** 693.

Acrivos, J. V., and Mott, N. F., 1971, *Phil. Mag.*, **24,** 19.

Altshuler, B. L., and Aronov, A. G., 1979, *Solid St. Commun.*, **36,** 115.

Anderson, P. W., 1958, *Phys. Rev.*, **109,** 1492.

Berggren, K. F., 1982, *J. Phys.* C, **15,** L49.

Bergmann, G., 1983, *Phys. Rev.* B, **28,** 2914.

Biskupski, G., 1982, Thesis, Lille.

Biskupski, G., Dubois, H., and Ferré, G., 1981, *Phil Mag.* B, **43,** 183.

Biskupski, G., Dubois, H., Wojkiewicz, J. L., Briggs, H., and Remanyi, G., 1984, *J. Phys.* C, **17,** L411.

Cabane, B., and Froidevaux, C., 1969, *Physics Lett.* A, **29,** 572.

Cohen, M. H., Economou, E. N., and Soukoulis, C. M., 1983, *J. non-crystalline Solids*, **59–60,** 15.

Cutler, M., 1977, *Liquid Semiconductors* (New York: Academic Press).

Davies, J. H., Lee, P. A., and Rice, T. M., 1982, *Phys. Rev. Lett.*, **49,** 758.

Davies, J. H., Lee, P. A., and Rice, T. M., 1984, *Phys. Rev.* B, **29,** 4260.

Davies, R., Pepper, M., and Kaveh, M., 1983, *J. Phys.* C, **16,** L285.

Debney, B. T., 1977, *J. Phys.* C, **10,** 4719.

Edwards, P. P., and Sienko, M. J., 1981, *J. Am. chem. Soc.*, **103,** 2967.

Edwards, P. P., and Sienko, M. J., 1983, *Int. Rev. phys. Chem.*, **3,** 83.

Efros, A. L., and Shklovskii, B. I., 1975, *J. Phys.* C, **8,** L49.

Elyutin, P. V., Hickey, B., Morgan, G. J., and Wein, G. F., 1984, *Phys. Stat. Sol.* (in the press).

Gorkov, L. P., Larkin, A. J., and Khmelnitzkii, D. E., 1979, *JETP Lett.*, **30,** 32.

Grest, G. S., and Lee, P. A., 1983, *Phys. Rev. Lett.*, **50,** 693.

Hertel, G., Bishop, D. J., Spencer, E. G., Rowell, J. M., and Dynes, R. C., 1983, *Phys. Rev. Lett.*, **50,** 743.

Ioffe, A. F., and Regal, A. R., 1960, *Prog. Semiconduct.*, **4,** 237.

Kaveh, M., and Mott, N. F., 1981, *J. Phys.* C, **14,** L177.

Kaveh, M., and Mott, N. F., 1983, *Phil. Mag.* B, **47,** L9.

Kawabata, A., 1981, *Solid St. Commun.*, **38,** 823.

Long, A. P., and Pepper, M., 1984, *J. Phys.* C, **17,** 3391.

Mansfield, R., Abdul-Gader, M., and Fozooni, P., 1984 (private communication).

Mooij, J. H., 1973, *Phys. Stat. Sol.*, (*a*), **17,** 521.

Mott, N. F., 1967, *Adv. Phys.*, **16,** 47.

Mott, N. F., 1972, *Phil. Mag.*, **26,** 1015.

Mott, N. F., 1978, *Phil Mag.* B, **37,** 377.

Mott, N. F., 1981, *Phil. Mag.* B, **44,** 265.

Mott, N. F., 1984a, *Phil. Mag.* B, **49,** L75.

Mott, N. F., 1984b, *Phil. Mag.* B (in the press).

Mott, N. F., and Davis, E. E., 1979, *Electronic Processes in Non-crystalline Materials* (Oxford University Press).

Mott, N. F., and Gurney, R. H., 1940, *Electronic Processes in Ionic Crystals* (Oxford University Press), Chap. 1.

Mott, N. F., and Kaveh, M., 1983, *Phil. Mag.* B, **47,** 577.

Mott, N. F., Pepper, M., Pollitt, S., Wallis, R. H., and Adkins, C. J., 1975, *Proc. R. Soc.* A, **345,** 169.

OVERHOF, H., and BEYER, W., 1983, *Phil Mag.* B, **47**, 377.

PAALANEN, M. A., ROSENBAUM, T. F., THOMAS, G. A., and BHATT, R. N., 1982, *Phys. Rev. Lett.*, **48**, 1284.

PEPPER, M., 1977, *Proc. R. Soc.* A, **353**, 225.

ROSENBAUM, T. F., ANDRES, K., THOMAS, G. A., and BHATT, R. N., 1980, *Phys. Rev. Lett.*, **45**, 1723.

SASAKI, W., 1980, *J. phys. Soc. Japan*, **49**, 31.

SCHMUTZLER, R. W., and HENSEL, F., 1972, *J. non-crystalline Solids*, **8–10**, 718.

SHAPIRO, B., and ABRAHAMS, E., 1981, *Phys. Rev.* B, **24**, 4025.

SPEAR, W. E., 1983, *J. non-crystalline Solids*, **59–60**, 1.

THOMAS, G. A., 1983, *Physica.* B, **117**, 81.

THOULESS, D., 1977, *Phys. Rev. Lett.*, **39**, 1167.

TIEDJE, T., and ROSE, A., 1981, *Solid St. Commun.*, **37**, 48.

ZIMAN, J. M., 1961, *Principles of the Theory of Solids* (Cambridge University Press).