

The mobility edge since 1967

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1987 J. Phys. C: Solid State Phys. 20 3075

(<http://iopscience.iop.org/0022-3719/20/21/008>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 128.42.202.150

The article was downloaded on 15/07/2012 at 17:25

Please note that [terms and conditions apply](#).

REVIEW ARTICLE

The mobility edge since 1967

Sir Nevill Mott

The Cavendish Laboratory, Cambridge, UK

Received 23 March 1987

Abstract. A mobility edge is defined as the energy separating localised and non-localised states in the conduction or valence bands of a non-crystalline material, or the impurity band of a doped semiconductor. This review is limited to three-dimensional systems, since in one or two dimensions a mobility edge in this sense does not exist, because all states are localised. We distinguish between the properties of electrons in the conduction bands of non-crystalline semiconductors, notably hydrogenated amorphous silicon (a-Si-H), and those in a degenerate electron gas, such as that in amorphous Si-Nb alloys or impurity bands in doped crystalline semiconductors. In the former the use of a one-electron model is legitimate, but a consideration of the interaction with phonons is essential; even at the absolute zero of temperature this leads to a broadening of the mobility edge. Our main purpose here is to review recent work on the effects of this interaction on the pre-exponential factor σ_0 in the conductivity expressed as

$$\sigma = \sigma_0 \exp[-(E_c - E_F)/k_B T]$$

and the pre-exponential factor in the drift mobility. In the final section we also give a brief review of some of the recent work on the effects of the electron-electron interaction in metallic systems, and also spin-orbit scattering.

Contents

1. Introduction
2. The Anderson model of localisation
3. The mobility edge according to the Anderson model
4. Effects of multiple scattering
5. A deduction of $\sigma(E)$ from the Kubo-Greenwood formula
6. The pre-exponential factor in the conductivity of a-Si-H
7. Liquid semiconductors
8. Spin-orbit scattering
9. Some effects of electron-electron interactions

1. Introduction

In a non-crystalline material, it is generally assumed that the lowest states in a conduction band are localised, that is to say they are traps. This seems to have been first proposed by Fröhlich (1947). The density of states is then as in figure 1; the highest energy at which

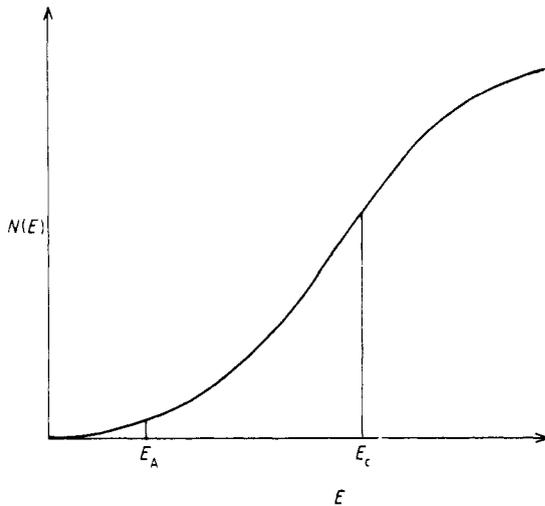


Figure 1. Density of states $N(E)$ in a non-crystalline material showing the mobility edge E_c . E_A is the energy at which injected electrons are in equilibrium at a given temperature.

states are localised is called the mobility edge, denoted by E_c . The term was first used by Cohen *et al* (1969), but the concept of a sharp energy separating localised from non-localised states was discussed in some detail two years earlier by the present author (Mott 1967). As far as we know, the first paper proposing a sharp energy separating localised from non-localised states was that of Banyai (1964). In the early days (round 1970) there was much discussion about whether the mobility edge should be sharp, and whether localised and non-localised (extended) states could coexist at the same energy. In general we believe they cannot, because a localised state will hybridise with the extended states, if they exist, so that only extended states remain. As regards sharpness, if the model is one in which the atoms are held rigidly in position, the mobility edge is sharp. In reality, in the conduction band of an amorphous semiconductor, the electron will always suffer *inelastic* collisions with phonons, even at zero temperature. We can thus introduce the inelastic diffusion length L_i , given by

$$L_i = (D\tau_i)^{1/2} \quad (1)$$

where D is the diffusion coefficient resulting from elastic collisions and τ_i is the time between inelastic collisions. These will broaden the mobility edge; according to the present author (1985a) the width is

$$\Delta E/E_0 \approx 0.03(a_E/L_i)^3. \quad (2)$$

Here E_0 is as marked in figure 1; E_A is the energy in the conduction band tail where the density of electrons is a maximum and (Mott and Davies 1979, p 31, Mott 1981b)

$$(a_E/a)^3 = \langle N(E) \rangle / N(E_c). \quad (3)$$

We shall return to equation (2) in § 5, and come to the conclusion that the numerical factor may be somewhat too small.

On the other hand, when we consider a *degenerate* electron gas in the limit of low temperatures, that is to say, an amorphous metal such as a-Si_{1-x}Nb_x (Hertel *et al* 1983),

or a heavily doped semi-conductor† then if the Fermi energy moves through E_c in figure 1 as a result of changes in the pressure, composition or magnetic field, a *sharp* metal-insulator transition is expected. Neither the mobility edge nor the Fermi energy is broadened, for the reason first pointed out for the Fermi energy by Jones *et al* (1934) and later by Landau (1957) in his discussion of the Fermi surface. Our discussion will separate the two cases; in the conduction band of semiconductors the mobility edge is always broadened by phonons, but electron-electron interaction is rarely important; in metals, however, the latter is extremely important, much more so than in crystals, and a theory which neglects it is subject to serious errors. The mobility edge is, however, sharp at zero temperature.

2. The Anderson model of localisation

The model introduced by Anderson (1958) has been extensively used in the discussion of the mobility edge, in ways that will first be outlined here, together with some of the consequences to be drawn from it. Figure 2 shows the Anderson potential; a three-

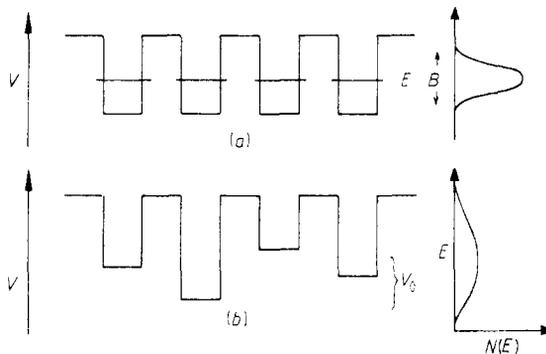


Figure 2. Potential energy function introduced by Anderson (1958). (a), In the absence of a random potential; (b), with random potential. The density of states is also shown.

dimensional array of potential wells is envisaged, with coordination number z . A random potential V such that $-\frac{1}{2}V_0 \leq V \leq \frac{1}{2}V_0$ is applied to each well, and a tight-binding approximation assumed. If V_0 is zero, the wavefunctions are of the form

$$\Psi = \sum_n \exp(ika_n)\psi(r - a_n) \tag{4}$$

† In discussions here of heavily doped semiconductors, we suppose that near the transition conduction is in an impurity band (see Mott and Kaveh 1985b, p 388, and also evidence drawn from ESR by Jerome *et al* 1985 and calculations by Ghazali and Serre 1982). Other points of view have been expressed, see for example Meyer *et al* (1985). We suppose also that the material is compensated and that the random positions of the centres in space, together with random fields produced by charged centres, gives rise to a mobility edge. For uncompensated samples the intra-site repulsion (the Hubbard U , cf § 9), is important, but the present author has argued that the transition is nonetheless of Anderson type, as assumed here (for a review see Mott and Kaveh 1985b, and § 9 of this paper.)

with a band width B given by

$$B = 2zI$$

where I is the transfer integral. A finite value of V_0 introduces an elastic mean free path, l . If $V_0 \sim B$, then $l \sim a$; this we call the Ioffe–Regel (Ioffe and Regel 1960) limit; l cannot be smaller than a and at the limit the wavefunctions, instead of (4), take the form

$$\Psi_{\text{IR}} = \sum_n c_n \exp(i\varphi_n) \psi(r - a_n) \quad (5)$$

where c_n, φ_n are random real numbers. If we write the conductivity in the form

$$\sigma = S_F e^2 l / 12\pi^3 \hbar \quad (6)$$

for a half-full band, putting $l = a$ and S_F (the Fermi surface area) equal to $4\pi k_F^2$ and $k_F = \pi/a$, we find

$$\sigma_{\text{IR}} = \frac{1}{3} e^2 / \hbar a. \quad (7)$$

As V_0 increases, the density of states will decrease. The present author has introduced the factor g , defined as

$$g = N(E_F) / N(E_F)_{V_0=0} \quad (8)$$

to describe the reduction. One can then show, (e.g. Mott and Davis 1979 ch 2), using the Kubo–Greenwood formula, that

$$\sigma = \sigma_{\text{IR}} g^2 \quad (9)$$

provided that interference effects resulting from multiple scattering are neglected. Anderson's work shows that, for a certain value of V_0 , states throughout the band became localised. According to recent calculations (Elyutin *et al* 1984), this occurs† when $V_0/B = 1.6$ and for this value the present author finds

$$1/g = 1.74(1 + V_0^2/B^2)^{1/2} \sim 3.$$

When the Fermi energy lies at the mobility edge, equation (9) gives a conductivity σ_{min} , where

$$\sigma_{\text{min}} = 0.03 e^2 / \hbar a. \quad (10)$$

3. The mobility edge according to the Anderson model

When V_0/B is less than the critical value, 1.6 according to Elyutin *et al* (1984) for $z = 6$, mobility edges will separate localised from extended states as in figure 1. The present author has proposed that for energies on the localised side of the mobility edge, wavefunctions should be of the form

$$\Psi = \Psi_{\text{IR}} \exp(-r/\xi). \quad (11)$$

ξ is called the localisation length, and in a theory of non-interacting electrons, near the edge

$$\xi \sim E_c a_E / (E - E_c)^\nu. \quad (12)$$

† Anderson (1958) found a much larger value. It should depend also on the coordination number.

Various calculations suggest $\nu = 1$ (Wegner 1979, 1982, Hikami 1980, 1982, Berezin *et al* 1980), though larger values have been found in numerical work.

For a degenerate electron gas, it is in principle possible to determine ν (of course, in the real model that includes electron–electron interactions) from observations of variable range hopping if carried out in a temperature sufficiently high for the Coulomb gap of Efros and Shklovskii (1975)[†] to be unimportant, since the conductivity should then vary as $\exp[-(T_0/T)^{1/4}]$ with

$$T_0 = \text{constant}/N(E_F)\xi^3 \quad (13)$$

(Pollitt 1976, Mott and Davis 1979 p 138, Shafarman *et al* 1986). We return to these investigations in § 8.

Calculations of the position of the mobility edge using the Anderson model have been made by Abou-Chacra and Thouless (1974). If $V_0/B \sim 1$, one can argue that the ‘tail’ shown in figure 1 will have a width $\sim V_0$. This will mean that for a fraction V_0/B of the wells the bottom is below the mobility edge. But in order to obtain localisation, the energy E_c must be so low that wells below it generate a band of width $\sim V_0/1.6$. The analysis appears to show that only a fraction $\sim V_0/B$ of the tail states become localised, so that any mobility edge lies well into the tail, at an energy of order $\frac{1}{2}V_0^2/B$, from the bottom of the band.

As regards equations (7) and (10) for the conductivity, we believe that they are valid when a_E (equation 3) is substituted for a .

4. Effects of multiple scattering

Equation (10) gives a value of the conductivity, $0.03e^2/\hbar a_E$, which the present author has called the ‘minimum metallic conductivity’, suggesting that it gave the value of the conductivity of a metal when E_F lies at E_c . The use of scaling theory by Abrahams *et al* (1979) first showed that, in a theory of non-interacting electrons, σ for a metal must in fact go continuously to zero as $E_F \rightarrow E_c$. What the effect of electron–electron interactions may be is still uncertain (see § 8), but there is now much evidence (Rosenbaum *et al* 1980) that σ , at low temperatures and in the absence of a magnetic field, does in fact go continuously to zero as $E \rightarrow E_c$.

In order to understand this, we look first at the effect of multiple scattering on the conductivity. According to Kawabata (1981, 1982), who used diagrammatic methods, multiple scattering introduces a correction to formula (9) for the conductivity, namely

$$\sigma = \sigma_B(1 - c/(k_F l)^2). \quad (14)$$

Here σ_B is the Boltzmann conductivity (6), k_F the wave-vector at the Fermi surface and c a constant of order unity. Bergmann (1983a, b, 1986), particularly in the two-dimensional problem, showed that the correcting term was the result of interference between waves scattered as shown in figure 3. If electrons are scattered from A to B by two equivalent paths in k -space ACDB, AC'D'B, the constructive interference will increase the scattering and reduce the conductivity. The form of (14) is easily understood as follows. The two paths will reinforce each other at B, and also within a circle round B of radius $\delta k = 1/l$ given by the uncertainty in k . Thus an area π/l^2 of the Fermi surface is effective, which is a fraction $\frac{1}{4}(kl)^{-2}$ of the whole area $4\pi k^2$. So the correcting term

[†] Near the transition, the temperatures at which the Coulomb gap is important must tend to zero as $n \rightarrow n_c$.

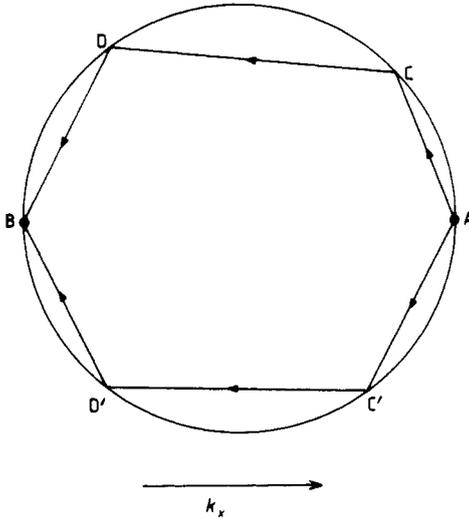


Figure 3. The Fermi surface of a metal. An electric field F is applied in the direction shown by the arrow, and $ACC'B$ and $ADD'B$ show two paths, the scattered waves interfering at B .

should be $c/(kl)^2$. The constant c has not been properly evaluated; it should be between 1 and 3 (see Mott and Kaveh 1985b).

One can also consider interference in real space. The correcting term then comes from the probability that an electron is scattering back to the point from which it started, the interference cutting down the intensity of the incident wave (see Kaveh and Mott 1987b).

If there is inelastic scattering, this interference will be diminished. Also, if the disorder is great enough to diminish the density of states, the Mott factor g must be introduced. Mott and Kaveh (1985a) write an amended form of (14) as

$$\sigma = \sigma_B g^2 \left[1 - \frac{c}{(k_F l)^2 g^2} \left(1 - \frac{l}{L_i} \right) \right]. \tag{15}$$

g represents the reduction in the density of states defined by (8), L_i is the inelastic diffusion length, which has already been defined (equation (1)). If electron–electron collisions are responsible, τ_i is proportional to $1/T^2$, so a correction to the conductivity linear in T is predicted†. This has been observed, both in amorphous metals (Howson 1984, Howson and Greig 1985) and in heavily doped semiconductors. However, electron–electron interaction also changes the density of states at the Fermi energy in such materials in a way which introduces another correction to σ proportional to $T^{1/2}$; this is also observed and is the most important correction at low T (§ 9).

† The literature sometimes refers to scattering by magnetic centres or ‘spin-flip’ collisions in metals as something different from electron–electron collisions. We believe this to be wrong at the zero of T . A moment in a metal, for instance due to a magnetic impurity, through Kondo-type interaction with the Fermi surface, produces in effect simply heavy pseudo-particles (see for instance Mott 1974, Grüner and Mott 1974). So ‘spin-flip’ collisions are only meaningful above the Kondo temperature $k_B T_k/\hbar = \Delta_F \exp(-E/\Delta)$ where Δ is the self energy and E is defined in the papers quoted. We know of no estimate for its magnitude in impurity bands, and it may be very small. Spin–orbit collisions are, however, something different, as discussed in § 8.

Kaveh and Mott (1985a, and earlier papers) propose that equation (15) can be extrapolated to the point where σ vanishes, and give reasons why this should be a good approximation (Kaveh and Mott 1987b). At the Ioffe–Regel limit one can write $k_F l = \pi$, for a half-filled band. (Near a band edge, if $l = a_E$, it will be smaller by $n^{2/9}$, where n is the number of states below E_c). In mid-gap, then, for $L_i = \infty$ (zero temperature), σ should vanish for $g = \frac{1}{3}$; near the band edge the critical values of g will be larger.

Whether E_F is in the mid band or not, the conductivity near the transition should behave like

$$\sigma_{\min} \left(\frac{2}{(k_F l)^2 g^2} \frac{d \ln g}{dR} (R - R_c) \right) \tag{16}$$

where R is the mean distance between atoms (or donors) and R_c the value at which σ vanishes. At the transition we put $(k_F l)^2 g^2 = 1$. For an impurity band, $N(E_F)$ without broadening by disorder should vary as $\exp(-\alpha R)$, where α is the rate at which the donor wavefunctions fall off; with broadening, this is decreased by a term depending on $e^2/\kappa R$, varying more slowly, since αR is normally of order four at the transition. So near the transition (16) may be written

$$\begin{aligned} \sigma &= \sigma_{\min} \left(2(\alpha R) \frac{R - R_c}{R_c} \right) \\ &= \frac{8}{3} \sigma_{\min} (n - n_c)/n_c \end{aligned} \tag{17}$$

where n is the concentration of donors and n_c the critical value.

We may compare this with the results of Hertel *et al* (1983) shown in figure 4 for the metal–insulator transition in a-Si : Nb. It will be seen that $n_c = 0.115N$, N being the total number of atoms cm^{-3} and that σ reaches the value σ_{\min} when $(n - n_c)/n_c = 0.39$. We predict from equation (17) the value $\frac{8}{3} = 0.375$, in excellent agreement. However, it must be realised that, for a metal, electron–electron interactions may make a large difference to this value (cf § 9), and for a heavy metal like Nb so may spin–orbit scattering.

At the transition, at any finite temperature, we must have $L_i \ll \xi$, so that (15) becomes

$$\sigma(E_c) = 0.03 e^2 / \hbar L_i. \tag{18}$$

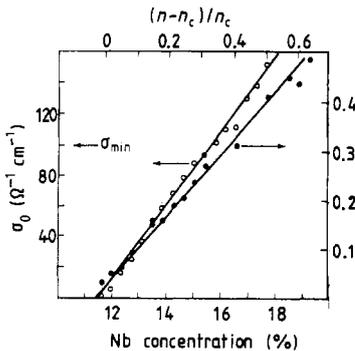


Figure 4. The low-temperature conductivity of amorphous Si–Nb as a function of niobium concentration (Hertel *et al* 1983). The density of states is also shown.

This equation must be relevant to the conduction band in non-crystalline semiconductors (L_i being the result of the interaction with phonons); it has been suggested (Mott 1985a) that the conductivity should be

$$\sigma = \sigma(E_c) \exp[-(E_c - E_F)/k_B T]. \tag{19}$$

Following work by P Thomas and co-workers (Dersch and Thomas 1985, Müller and Thomas 1984, Fenz *et al* 1985), this may be incorrect, as we shall show in § 6.

In a magnetic field H , at zero temperature, or if $L_H \ll L_i$, (15) should take the form

$$\sigma = \sigma_B g^2 \left[1 - \frac{c}{(k_F l g)^2} \left(1 - \frac{l}{L_H} \right) \right] \tag{20}$$

where L_H is the cyclotron radius $(e\hbar/cH)^{1/2}$. Physically it is clear that a magnetic field will destroy the phase correlation illustrated in figure 3, so that the field leads to a *negative* magnetoresistance, varying as

$$\Delta\rho \propto - \text{constant } H^{1/2}.$$

This will normally be the major term in the magnetoresistance for small fields. On the other hand, it has been realised for many years that in doped semiconductors a magnetic field will decrease the radius of the orbits and thus decrease the band width B in the absence of disorder, g in (20) therefore decreasing. This leads to a linear term

$$\Delta\rho \propto \text{constant } H.$$

Thus with increasing field σ should first increase and then decrease. Then, as Shapiro (1984) has pointed out, the mobility edge as a function of energy and field H should behave as in figure 5; along the dotted line an increasing field should first give a transition to a metallic state and then again to an insulating state. These transitions will not show a minimum metallic conductivity (Mott and Kaveh 1985b p 857), σ tending continuously

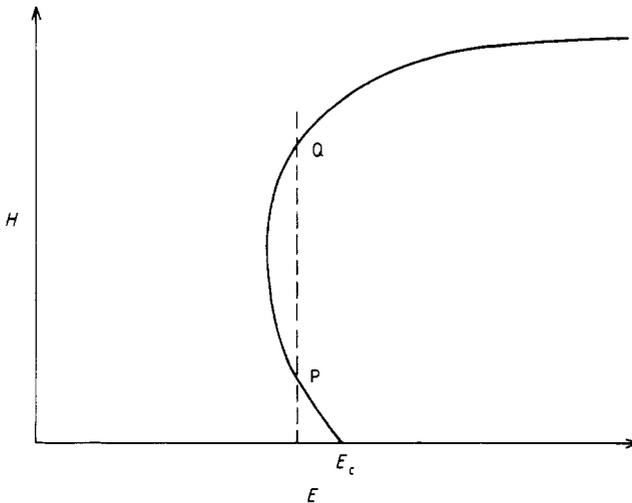


Figure 5. Phase diagram separating extended from localised states as a function of energy E and magnetic field H . If the field is increased along the dotted line, transition between localised and extended behaviour can occur at P and Q.

to zero. Such an effect neglects electron–electron interactions and may not be present if the result of these is large.

On the other hand, when L_H is less or comparable with l (equal to a at the Ioffe–Regel limit), the correcting term in (20) seems to disappear and a minimum metallic conductivity σ_{\min} given by figure 6 appears. This is not fully understood theoretically, but seems to follow from the experimental work of Biskupski *et al* (1981, 1984) and Long and Pepper (1984a). In doped and compensated InP and InSb, behaviour of the type shown in figure 6 is observed, experiments having been carried out down to 40 mK. A

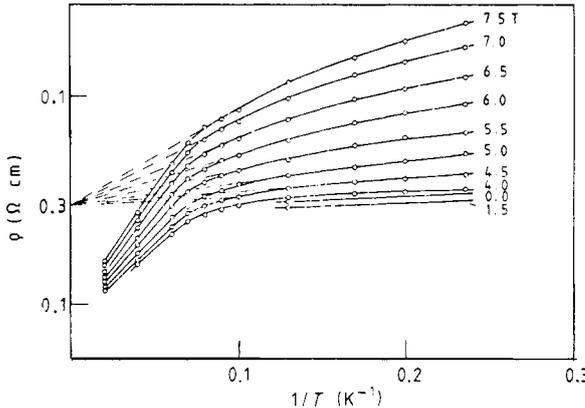


Figure 6. Ln (resistance) of doped InP plotted against $1/T$ for increasing values of the field, the Tesla (Biskupski 1982).

minimum metallic conductivity is clearly observed and the same quantity appears—as it should—in the pre-exponential factor for activated charge transport. Moreover σ_{\min} can be plotted against a (a^{-3} is the concentration of electrons), giving

$$\sigma_{\min} \sim 0.037 e^2 / \hbar a$$

as shown in figure 7. The factor 0.037 is very near to our prediction (0.031). A minimum metallic conductivity of this order, for a transition induced by a strong field, is also observed in Si:Sb by Long and Pepper (1984b).

This effect appears only in materials in which the term due to interactions ($\Delta\sigma \propto T^{1/2}$) is small (see § 8). This is discussed by Mott and Kaveh (1985b) and Kaveh and Mott (1987b), who propose that σ goes rapidly but continuously to zero with increasing H .

In § 7 we describe another case (liquid semiconductors) in which the correcting term certainly disappears, because in this case $l \approx L_i$.

In equation (15), in the absence of a field and at the absolute zero of temperature, L will be the size of the specimen. In an alternating field, we may write

$$L_\omega = (D/\omega)^{1/2}$$

where ω is the frequency. Another length, relating to spin–orbit scattering, is discussed in § 8.

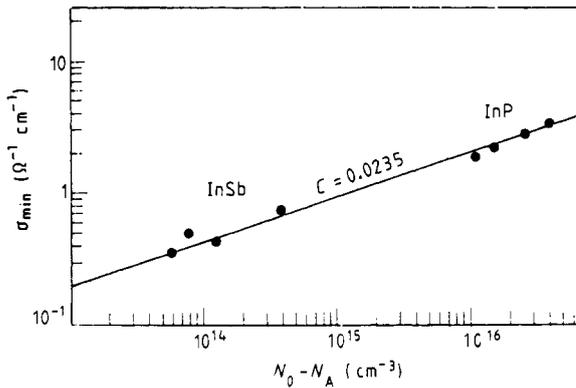


Figure 7. Plot of $\sigma_{\min}(\text{obs})$ against $N_0 - N_A$ (Biskupski 1982). C is the constant in $\sigma = Ce^2/\hbar a$.

5. A deduction of $\sigma(E)$ from the Kubo–Greenwood formula†

In this section we start with the Kubo–Greenwood formula (Mott and Davis 1979 p 11). This formula gives for the conductivity in the metallic regime

$$\sigma = \frac{2\pi e^2 \hbar^3 L^3}{m^2} (M^2)_{\text{av}} [N(E)]^2 \quad (21)$$

where

$$M = \int \Psi_1^* \frac{\partial}{\partial x} \Psi_2 d^3x$$

Ψ_1, Ψ_2 , being wavefunctions near the Fermi energy. This formula is simply the limit, as $\omega \rightarrow 0$, of the optical absorption coefficient resulting from transitions from states just below to just above the Fermi energy of a metal at zero temperature. When in 1979 Abrahams and co-workers used the scaling theory to show that $\sigma(E)$ should (in the absence of interactions) tend linearly to zero as $E \rightarrow E_c$, the present author felt the need to show how (21), involving only an average of squared terms, could vanish. This could not be ascribed to the density of states $N(E)$; theory (Thouless 1970) showed that there should be no discontinuity in $N(E)$ at the transition‡ (still less a zero value); and this was confirmed by measurements of Kobayashi *et al* (1977) of the (electronic) specific heat of Si:P in the metallic regime.

Mott (1984) and Mott and Kaveh (1985a, b) discussed how this could occur. We now give a somewhat revised account of their argument.

† Kubo (1956), Greenwood (1958).

‡ The calculation of Thouless was for non-interacting electrons. When interaction is taken into account (§ 9), one must distinguish between $\tilde{N}(E)$, the value when the free-energy is minimised, which does not vanish, and $N(E)$, the value representing the excitation of a single electron, which does. $N(E)$ is appropriate for discussions of photoexcitation and hopping conduction (Efros and Shklovskii, 1975), $\tilde{N}(E)$ for metallic conductivity following the equation

$$\sigma = e\tilde{N}(E)\mu$$

where μ is the mobility, and also for the specific heat.

We first consider states with energies lying between E_c and some lower value E such that

$$\Delta E = E_c - E$$

is small. Wavefunctions will normally be non-degenerate and therefore, in the absence of a magnetic field, real. We suppose them to be of the form (cf equation (11)).

$$\text{Re}(\Psi_{\text{ext}}) \exp(-r/\xi)$$

where Ψ_{ext} may be the wavefunction Ψ_{IR} defined by equation (11), but the constants c_n in (5) remain unknown. As we have seen, when ΔE is small

$$\xi \sim aE_c/\Delta E.$$

Consider two localised wavefunctions Ψ_1, Ψ_2 , at a distance R from each other, and suppose also that they are not orthogonal. To orthogonalise them we write

$$a\Psi_1 + b\Psi_2$$

$$a\Psi_2 - b\Psi_1$$

and argue that the overlap integral must be less than ΔE , if they are both to lie in the range ΔE . The transfer integral will then be of the form

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

so the *minimum* distance between them is

$$R = \xi \ln(H_0/\Delta E) \approx \ln(\xi/H_0 E_c). \tag{22}$$

H_0 and E_c should be of the same order, so we write

$$R \sim \xi \ln(\xi/a). \tag{23}$$

This gives us a number of states per unit volume in the range ΔE equal to $1/R^3 \approx [(\Delta E)/aE_c]^3 (a/R)^3$, while the true number must be $N(E)\Delta E$, which is much greater.

This is discussed by Mott and Kaveh (1985a). They point out that the statements above should be limited to states with energies lying within a range $\hbar\omega_c$ of each other, where

$$\hbar\omega_c = 1/N(E)\xi^3. \tag{24}$$

The correct density of states is then obtained.

We discuss next the form of the wavefunctions of states with energies in a range ΔE above the mobility edge. We argue that within a length R each wavefunction Ψ at energy E above E_c will be similar to *one* of the localised states at the same energy to below it. We make the assumption (which we have not been above to prove) that, within the energy range $\hbar\omega_c$ and the length R , all wavefunctions are similar. We illustrate our concept of the envelopes of the functions in figure 8. In the maxima, the wavefunctions over an energy range $\hbar\omega_c$ are highly degenerate (but may no longer be orthogonal). We argue, however, that within one 'bump' we can superimpose the functions so that all are real.

If we use the Kubo–Greenwood formula to calculate σ , we found (Mott 1972) that at the mobility edge $\sigma = 0.03e^2/\hbar a$, assuming that phases of Ψ in adjacent atoms are

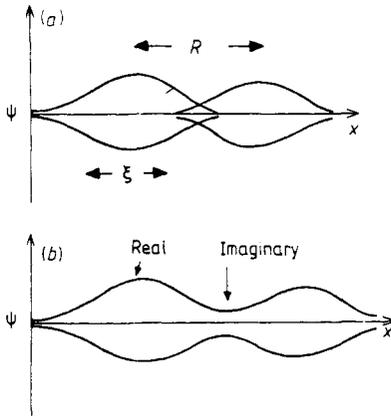


Figure 8. Suggested form of the envelopes of the wavefunctions for energies (a), just below; and (b), just above E_c .

uncorrelated. We now argue that within one bump (figure 8) two functions with neighbouring energies are both real and essentially identical. Thus

$$\int \Psi_1 \frac{\partial}{\partial x} \Psi_2 d^3x = \frac{1}{2} \int dy dz [\Psi]_x^{x+R} \tag{25}$$

where the integral is over a volume R^3 and $x, x + R$ are at the two minima. The term $(\xi/a)^{1/2}$ which arose in the author's treatment (1972) from the assumption that the phases in adjacent atoms are uncorrelated now disappears, so the integral is reduced by $(a/\xi)^{1/2}$. The normalisation of the functions introduces a factor $[\exp(R/\xi)]^2$ in the conductivity, which just cancels the effect of the limits $x, x + R$ in (25).

This very approximate and intuitive argument gives

$$\begin{aligned} \sigma(E) &= \sigma_{\min} a/\xi \\ &= 0.03e^2/\hbar\xi. \end{aligned} \tag{26}$$

This form with similar value of the constant was given first by Abrahams *et al* (1979) and Wegner (1979).

If we compare (21) with equation (6) we obtain an expression for ξ

$$a/\xi = \frac{2}{3}(n - n_c)/n_c$$

where n is the number of donors and n_c the value of n for which σ vanishes. If we write in the neighbourhood of E_c

$$(n - n_c)/n_c = (E - E_c)/E_0$$

we find

$$a/\xi = \frac{2}{3}(E - E_c)/E_0. \tag{27}$$

E_0 is equal to E_c if $N(E)$ is linear between E_c and the band edge; if not, as shown in figure 9, it is somewhat less than E_c . Such a relationship, however, depends on our definition of ξ in the non-localised range.

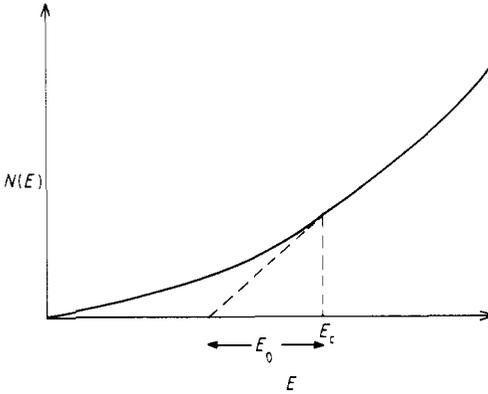


Figure 9. Density of states near E_c to illustrate the quantity E_0 .

As already stated, however, since ξ will tend to infinity at the mobility edge, sufficiently near the edge we must always put

$$L_i \ll \xi.$$

We shall next derive equation (18), namely

$$\sigma = c(E)e^2/\hbar L_i \quad c(E) \sim 0.03$$

under these conditions†. We use the Kubo–Greenwood equation (20) and set

$$M = (m\delta E/\hbar^2)x_{12}$$

where

$$x_{12} = \int \Psi_1 \alpha \Psi_2 \, d^3x \tag{28}$$

the integral being over the volume L_i^3 , and

$$\delta E = 1/L_i^3 N(E_F).$$

This gives

$$\sigma = \frac{2\pi e^2}{\hbar} \frac{|x_{12}|_{\text{av}}^2}{L_i^3}. \tag{29}$$

It is, in our view, correct to evaluate x_{12} over the volume L_i^3 . There is no one-electron wavefunction extending from one such block to another. So the conductivity of the whole block, of size L^3 , is the same as that of each block of size L_i^3 . In (28), Ψ_1 and Ψ_2 must be orthogonal. We write (with $L = L_i$)

$$x_{12} = \iiint dy \, dz \int_{-L}^{+L} x \Psi_1 \Psi_2 \, dx \tag{30}$$

and, to ensure orthogonality, we approximate by writing

$$\Psi_1 \Psi_2 = 2x/L^2$$

† The result obtained differs slightly from that given by Kaveh and Mott (1987b).

the 2 being inserted to ensure that Ψ_1^2 and Ψ_2^2 are of order $1/L$ at the extremities of the range of x . Thus (30) becomes $L/6$ and the conductivity is

$$\frac{e^2}{\hbar L_i} \frac{2\pi}{3} \left(\frac{1}{6}\right)^2. \quad (31)$$

The factor $\frac{1}{3}$ arises because in (31) one could write y or z instead of x , in which case the integral vanishes.

The constant in (31), 0.058, is near enough to our value 0.03 obtained in the last section, taking into account the approximations involved.

In our view, contrary to that expressed by Kaveh and Mott (1987b), a temperature variation of the constant $c(E)$ can only occur if L_i approaches a , in which case, of course, the mobility edge becomes broad and the extrapolation of (15) to the transition doubtful.

6. The pre-exponential factor in the conductivity of a-Si-H

This was discussed by the present author† (1985a) on the assumption that one could write

$$\sigma = \sigma_0 \exp[-(E_c - E_F)/k_B T] \quad (32)$$

with $\sigma_0 = \sigma(E_c)$ and that E_F would vary with temperature, but not $\sigma(E_c)$ or E_c . Taking $N(E_c) = 5 \times 10^{20} \text{ cm}^{-3} \text{ eV}^{-1}$ (Spear 1983) and the band width $B = 10 \text{ eV}$ (Kramer *et al* 1983), we found $a_E = 5.8 \text{ \AA}$ and $\sigma_{\min} = 126 \text{ } \Omega^{-1} \text{ cm}^{-1}$. We supposed that $\sigma(E_c) = 0.03e^2/\hbar L_i$ and gave an argument in favour of the factor 0.03, even when E_c lies near the bottom of the band. This argument is as follows.

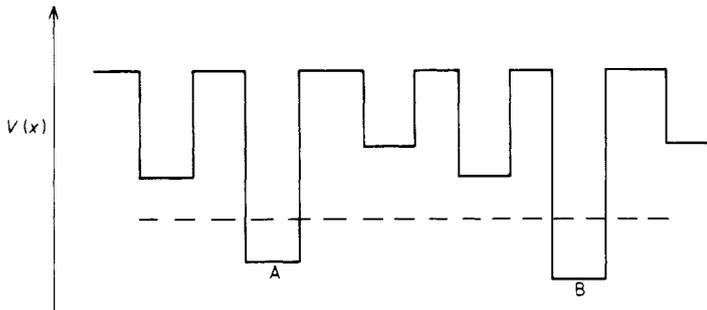


Figure 10. The potential energy introduced by Anderson, illustrating the behaviour at the bottom of a band.

If in the model proposed by Anderson (1958) illustrated in figure 2, one takes in figure 10 an energy E near the bottom of the wells and assumes, say, a Gaussian distribution of well depths, then only those wells with depths below E contribute to the density of states, $N(E)$, namely those marked A and B in figure 10. A narrow band formed from these will result. We argue that the treatment used by Anderson can be applied to this band, and that localisation will occur when $g = \frac{1}{3}$.

† For a correction to this paper see Mott (1986a).

We also gave an estimate of the uncertainty in the position of the mobility edge, ΔE , introduced by inelastic scattering; this is as follows. We write

$$\Delta E = \hbar/\tau_i$$

where τ_i is the time between inelastic collisions. We wrote

$$1/\tau_i = D/L_i^2$$

so

$$\Delta E = \hbar D/L_i^2.$$

Also

$$\sigma = e^2(E)D$$

so

$$\Delta E = \sigma \hbar / e^2 N(E) L_i^2.$$

Also

$$N(E) \sim 1/a_E^3 E_0$$

and

$$\sigma = 0.03e^2/\hbar L_i \quad (33)$$

giving

$$\Delta E/E_0 = 0.03e^2/\hbar L_i. \quad (34)$$

This we now think may be an underestimate, since the coefficient in (33) is too small (see below).

We estimated L_i from the data of Vardeny *et al* (1981) on the rate of loss of energy of photoelectrons in a-Si-H and found $L_i \sim 30 \text{ \AA}$, from which we obtained $\sigma(E_0) = 25 \text{ \Omega}^{-1} \text{ cm}^{-1}$.

Comparison with experiment rested mainly on measurements of the drift mobility μ_D by the Dundee group, using the relationships

$$\mu_D = \mu_0 \exp(-\Delta E/k_B T) \quad (35)$$

and

$$\mu_0 = \sigma_0/eN(E_A)k_B T \quad (36)$$

and Spear's estimate for $N(E_A)$, the effective density of states at the bottom of the band.

More recently Overhof (1986) has examined the probable temperature-dependence of E_F , and come to the conclusion that, to account for the experimental observations on the conductivity, σ_0 should be much greater, $\sim 1000 \text{ \Omega}^{-1} \text{ cm}^{-1}$. Thomas and co-workers[†], pointing out that σ_0 should be given by

$$\sigma_0 = - \int \frac{\partial f}{\partial E} \sigma(E) dE \quad (37)$$

where f is the Fermi-Dirac function

$$1/\{1 + \exp[(E - E_F)/k_B T]\}$$

[†] Müller and Thomas (1984), Fenz *et al* (1985), Thomas (1987a, b).

find that even at zero temperature this is not equal to

$$\sigma(E_c) \exp[-(E_c - E_F)/k_B T].$$

There is no discontinuity in $\sigma(E_c)$ at E_c (since L_i is finite even at $T = 0$) and the hopping conductivity for the non-degenerate gas, first described by Grant and Davis (1974), is considerable. By using a mode-coupling analysis due to Götze (1978, 1979, 1981) and Belitz and Götze (1983) these authors calculate $\sigma(E)$ near E_c and find that the maximum current does not normally lie at E_c .

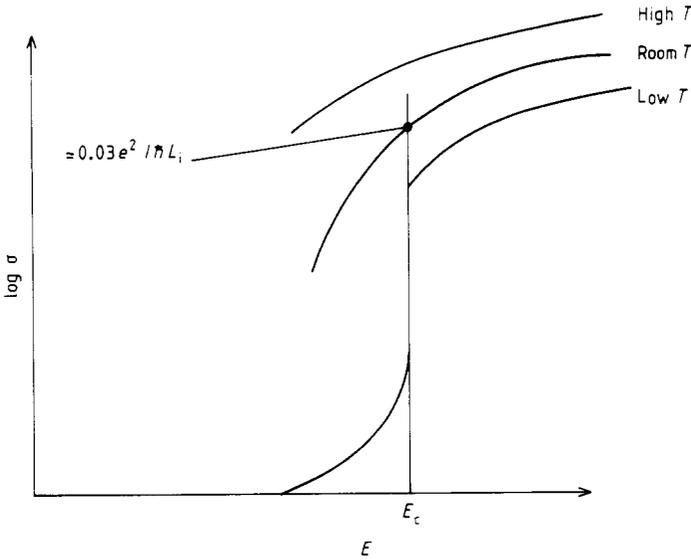


Figure 11. Plot of $\sigma(E)$ as a function of E near the mobility edge E_c in conduction band of a semiconductor (schematic).

Figure 11 shows schematically the behaviour of $\sigma(E)$ near E_c at high and low temperatures. At low T it lies near the bottom of the band, or even in the exponential tail, and conduction is by hopping; then another channel appears at a value of E above E_c , where the plot of $\ln \sigma$ versus $1/T$ is flat and $\sigma \sim \sigma_{\min}$. Thus, as the temperature rises further and L_i decreases and the value of $\ln \sigma(E)$ drops less rapidly with decreasing T , the current path moves to lower energies below E_c . Thus the conductivity is of the form

$$\sigma(E_M) \exp[-(E_M - E_F)/k_B T]$$

where E_M is the energy of the current path, and approximately

$$E_M - E_F = \varepsilon - \gamma T \tag{38}$$

where

$$\varepsilon \geq E_F - E_c \quad \text{at } T = 0.$$

The term $\exp(\gamma/k_B)$ leads to an increased value of the observed pre-exponential factor. However, we do not believe that this factor can be greater than ~ 10 . It follows then that

equation (34) somewhat underestimates the width of the mobility edge. Instead of (33) we should take

$$\sigma_0 = c_0 e^2 / \hbar L_i$$

where c_0 is up to 10 times greater than 0.03, so

$$\Delta E / E_0 \leq (a_E / L_i)^3.$$

With $a_E \sim 6 \text{ \AA}$ and L_i at room temperature $\sim 30 \text{ \AA}$, we find $\Delta E / E_0 \sim 0.013$.

Mott (1985a) found that his value of $\sigma(E_c)$ agreed closely with data obtained from the drift mobility, assuming the Dundee value of $N(E_A)$. An examination of the Dundee data as well as that of Tiedje *et al* (1981) by Davis *et al* (1985) and Michiel *et al* (1987) leads to the conclusion that, while Spear's value $N(E_c) \approx 5 \times 10^{20} \text{ cm}^{-3} \text{ eV}^{-1}$ is reasonable, a larger value of $\mu_{\text{ext}} \sim 70$ (instead of 10) gives slightly better agreement with experiment.

However, following discussions with Professor Spear and co-workers, we are convinced that the effects predicted by Thomas are not shown in the drift mobility measurement, and must be small. The formulae we use are

$$\begin{aligned} \mu_D &= \mu_0 \exp(-\Delta E / k_B T) \\ \mu_0 &= \mu_{\text{ext}} N(E_c) / N(E_A) \\ \mu_{\text{ext}} &= \sigma_0 / e N(E_c) k_B T \end{aligned} \quad (39)$$

from which we find

$$\mu_0 = \sigma_0 / e N(E_A) k_B T.$$

Measurement at high T , at which (39) breaks down, enable us, following Spear, to obtain μ_{ext} and $N(E_c) / N(E_A)$ separately. If σ_0 is given by Mott's value $0.03 e^2 / \hbar L_i$, and $N(E_c) / N(E_A) \approx 6$, the experimental results are reproduced; if σ_0 contains a factor $\exp(\gamma / \hbar)$ of order 10, $N(E_A)$ and $N(E_c)$ become comparable, which seems unlikely†.

We now believe (Spear 1987) that the argument given by Spear *et al* (1980) postulating a movement of E_c with temperature are incorrect; the kink in the $\sigma - 1/T$ shown in figure 7 of that paper at 420 K is given an alternative explanation by Kakalios and Street (1986). The large values of the pre-exponential factor, that is a large term $\exp(\gamma/k)$, are a result of the shrinkage of the band gap with increasing temperature, resulting from phonon interactions, observed in optical absorption and described theoretically by Fan (1951) for crystals and for non-polar amorphous materials by Griffith (1977). A shift of $E_c - E_v$ with T of about $6.5 \times 10^{-4} \text{ eV } k_B^{-1}$ is observed, which gives $\exp(\gamma/k)$ about 1000. The point is here that E_c and E_A move together, relative to the vacuum and to E_F , so the drift mobility is not affected, only the pre-exponential in σ . This value is confirmed by both photoconductivity experiments (Spear and LeComber 1985) combined with absorption from E_v to E_F (Djamdi and LeComber 1987).

The pre-exponential is a maximum for uncompensated material, when E_F is pinned by the dangling bands. For the compensated material $N(E_F)$ increased as E_F approaches either band, so the statistical shift of E_F gives a term in σ_0 of the type $\exp(-\gamma'/k)$. A decrease in σ_0 with compensation is therefore observed (Heintze and Spear 1986).

† This, however, is the conclusion obtained most recently by Thomas and Overhof (1987).

The large values of σ_0 demanded by Silver *et al* (1986 and earlier work) are now thought by those authors to be limited to experiments with a large concentration of electrons and holes which can screen out the random field, so such large values do not seem acceptable for normal specimens on any hypothesis.

Spear's model used in these calculations proposes the following values for the density of states below E_c ;

$$N(E) = N_0(\Delta - E)/\Delta \quad E_c < E < 0.12 \text{ eV} \quad (40)$$

with E_c as the zero of energy, $N_0 = 5 \times 10^{20} \text{ cm}^{-3} \text{ eV}^{-1}$ and $\Delta = 1.3 \text{ eV}$, and

$$N(E) = N_a \exp[(0.12 - E)/k_B T_c] \quad 0.12 < E < 0.30$$

with $T_c = 310 \text{ K}$, and $N_a = 4 \times 10^{19} \text{ cm}^{-3} \text{ eV}^{-1}$. The model thus assumes a linear drop in $N(E)$ followed by an exponential decay. Marshall *et al* (1985) point out that better agreement with experiment, particularly the absence of dispersive transport observed by the Dundee group, is obtained if it is assumed that the cross section for trapping into the states in the exponential tail is small. We therefore give some speculations about the origin of the exponential tail. Davies (1980) made calculations of the value of E_c in a-Si-H, ascribing the localised states to the random dihedral angles in the material. We suggest that the linear tail of localised states may have this origin, giving a capture cross section of order 10^{-15} cm^2 as in the model just described. Overhof and Beyer (1983, 1984) have maintained that in addition there should be long-range fluctuations of potential resulting from point charges; this is necessary to explain the observed differences between the activation energy for conduction and that for the thermopower. We suggest that these are the origins of the exponential tail. The transitions into such states might have low probability because they do not overlap those in the linear region.

To summarise, we now think that, to account for the observations, we have to make the following hypotheses.

(i) In undoped a-Si-H, the lower part of the conduction band, that is *both* E_A and E_c , move together with increasing T with reference to the vacuum (as in crystalline Si) and to E_F . The large pre-exponential in the conductivity must be ascribed both to this effect *and* the shift of the current path as described by Thomas and co-workers.

(ii) The latter, however, can only be a very minor part of the effect, since the constants in the drift mobility are not greatly affected up to room temperature.

(iii) The kink in the curve plotting $\log \sigma$ versus $1/T$ at 420 K has its origin in a phase change involving hydrogen (Kakalios and Street 1986).

(iv) The pre-exponential should be given by

$$\sigma_0 \approx (0.03e^2/\hbar L_i) \exp(\gamma/k_B)$$

where for undoped specimens γ is the result of band-gap shrinkage and for doped specimens also of the statistical shift of E_F .

7. Liquid semiconductors

The present author (Mott 1985b) has suggested that in liquids all collisions are inelastic, and therefore in equation (15) $L_i = l$ so that the correcting term is absent. The strongest experimental evidence is perhaps that of Cutler (1977) on liquid $\text{Te}_{1-x}\text{Tl}_{2+x}$ alloys;

the Pauli susceptibility (proportional to g) for various compositions and temperatures is plotted in figure 12 against $\sigma^{1/2}$ with σ in the range 100 to 2500 (less than the Ioffe-Regel value). A straight line is obtained.

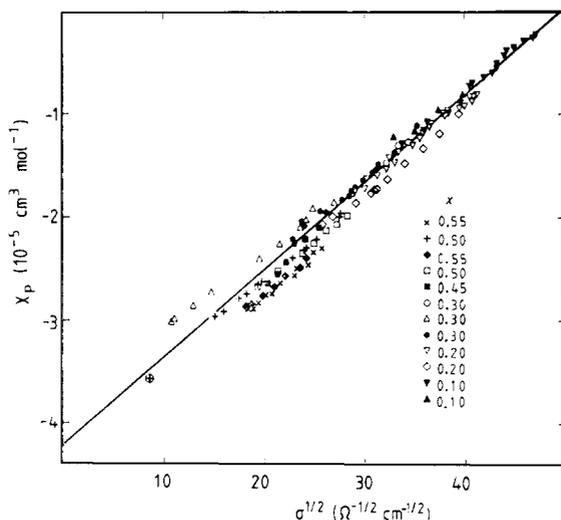


Figure 12. Pauli susceptibility of liquid $\text{Te}_{1-x}\text{Pb}_{2+x}$ alloys plotted against $\sigma^{1/2}$ for various temperatures and compositions (Cutler 1977).

A minimum metallic conductivity $\sigma = 0.03e^2/\hbar a_E$ should exist and give the pre-factor in σ . Other evidence is described by Mott (1985b).

Recently Barnes *et al* (1987) have described the conductivity and thermopower of liquid Mn–Te alloys. In these materials conduction is solely by holes in a manganese 3d band broadened by hybridisation with 5p electrons from the tellurium. In Mn–Te each Mn atom has given two electrons to form Te^{2-} , leaving Mn^{2+} with 5 electrons in the d band forming an antiferromagnetic insulator. Addition of Te produces holes in the band described above. Figure 13 shows the conductivity as a function of concentration and figure 14 the thermopower. We believe that for specimens for which $\sigma < \sigma_{\min}$ ($\sim 350 \Omega^{-1} \text{cm}^{-1}$) the states are localised. The high values of the thermopower are difficult to explain otherwise, but are consistent with the equation

$$S = (k_B/|e|)(\Delta E/k_B T + 1)$$

with $\Delta E = 0.2 \text{ eV}$. Also the measurements of σ show a small positive value of $d\sigma/dT$, which is consistent with the same activation energy, 0.2 eV.

We believe, then, that for liquids $\sigma(E)$ is near to the value σ_{\min} given by (6) at the mobility edge and for some way above it.

There appear, then, to be two cases in which the correcting term in the Kawabata formula is absent; these are liquids, and under a magnetic field when $(\hbar v_F/eH)^{1/2}$ is comparable or less than l , though explanations are quite different (Kaveh and Mott 1987a).

Enderby and Barnes (1987) have used the value $\sigma_{\min} \sim 1000 \Omega^{-1} \text{cm}^{-1}$ in a discussion of the electrical conductivity of molten sulphide and solid mixtures, maintaining that, when σ drops below this value, conduction may be mainly ionic.

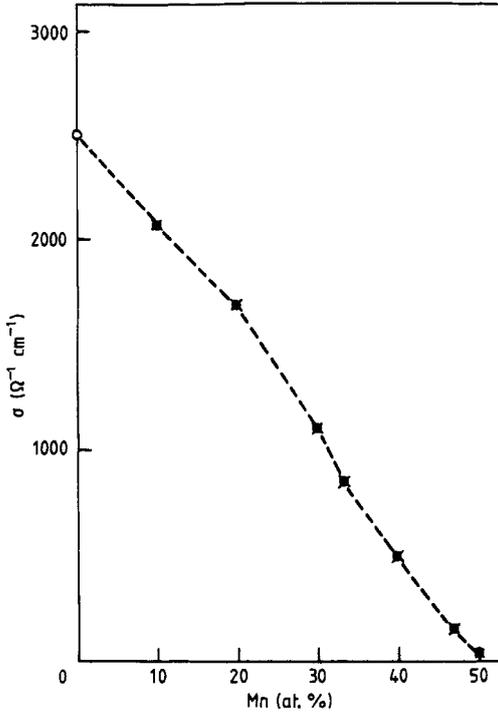


Figure 13. The conductivity of liquid Mn_xTe_{1-x} alloys as a function of x extrapolated to 1500 K. The temperature dependence is slight (Barnes *et al* 1987, figure 4).

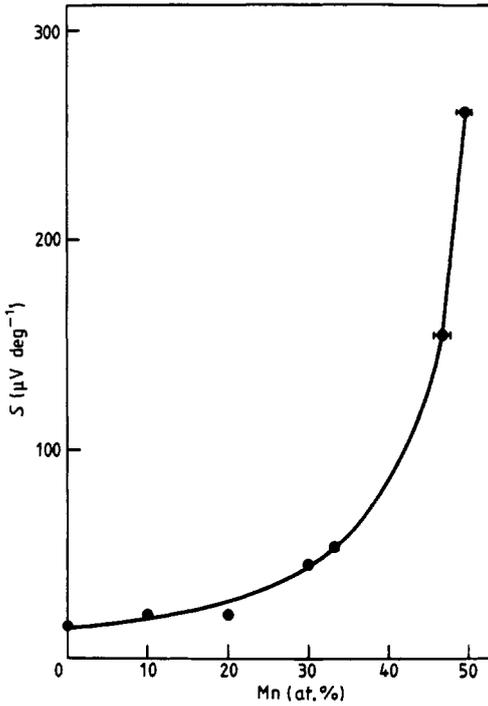


Figure 14. Thermopower at 1500 K of alloys shown in figure 13 (Barnes *et al* 1987, figure 5).

8. Spin-orbit scattering

This is a process in which the electron changes its spin on scattering, the angular momentum being taken up by the specimen as a whole. The relevant cross section is believed to be independent† of T , and to increase rapidly with atomic number Z of the scatterer. Thus Hahn and Enderby (1967) investigated the spin-lattice relaxation in both solid and liquid lithium through the effect of adding Mg, Zn and Ag to the ESR line width; the effect of Ag atoms dissolved in broadening the ESR line was about 500 for Mg; a variation of $(Z - Z')^4$ is often assumed, which would give a ratio 200. Here Z' is the atomic number of the matrix. The most recent experimental and theoretical work is that of Monod and Schultz (1982).

The effect of spin-orbit scattering on the resistivity of non-crystalline materials seems first to have been discussed by Hickami *et al* (1980). A time τ_{so} between such collisions can be defined, leading to an inelastic diffusion length L_{so} , given by

$$L_{so} = (D\tau_{so})^{1/2}. \quad (41)$$

This is independent of T , and may be greater or less than L_i . Both Hahn and Enderby (1967) and Maekawa and Fukuyama (1981) give formulae for τ_{so} , but they have not been worked out in any particular case.

Experimental work on thin metallic non-crystalline films, treated as two dimensional, were carried out by Bergmann (1982) and Kawaguchi and Fujimori (1982). Discussions are given by Bergmann (1983a, b, 1986). The magnetoresistance is normally negative, because the constructive interference illustrated in figure 3 is destroyed by the field. Bergmann shows, however, that after a spin-orbit collision this does not occur, so the resistance increases due to the contraction of the orbits. For weak fields, at any rate, where $L_i \gg L_H$, the magnetoresistance is positive.

In the next section we discuss the effect on the resistivity of Si:Sb.

9. Some effects of electron-electron interactions

While these are probably negligible for electrons in the conduction band of a semiconductor, they are of major importance for a degenerate electron gas when the mean free path is short, and *a fortiore* when the Fermi energy E_F lies near the mobility edge E_c . Altshuler and Aronoff (1979) first showed that these interactions lead to a change in the density of states near the Fermi energy given by

$$\delta N(E) = (2\pi\hbar D l)^{-1} \left(\frac{1}{3} - 2F \right) \{ -1 + l|E - E_F|/\hbar D \}^{1/2}. \quad (42)$$

It will be seen that the second term in (42), being proportional to $D^{-3/2}$, becomes important only when D is small, and thus for 'dirty' systems such as amorphous metals or impurity bands. It is deduced that $\sigma = \sigma(0) + \delta\sigma(T)$ where

$$\delta\sigma(T)/\sigma(0) = 0.03 \left(\frac{\hbar}{mD} \right)^{1/2} \left(\frac{1}{3} - 2F \right). \quad (43)$$

† Paalanen *et al* (1986) deduce from ESR measurements in Si:P a temperature-variation of the spin-flip scattering rate below 1 K, which they ascribe to spin-orbit scattering. We think this may be caused by a temperature-dependence of $N(E_F)$ when averaged over the range $k_B T$, which leads to the $T^{1/2}$ correction to σ , as described in the final section.

In the last term, the $\frac{1}{3}$ comes from exchange interactions between electrons and $2F$ from the Hartree interaction. F is given by

$$F = x^{-1} \ln(1 - x) \quad x = 2k_F \lambda$$

where λ is the screening length. $\delta N(E)$ will decrease with T (Altshuler and Aronov 1983, Fukuyama 1985).

The term $mT^{1/2}$ has been observed in amorphous metals and in doped semiconductors, and also for deformed metallic bismuth where the mean free path is short and can have either sign. Kaveh and Mott (1982, 1987a) argued that the inclusion of electron–electron interaction modifies the Kawabata formula to

$$\sigma = g^2 \sigma_B \left[1 - \frac{1}{g^2 (k_F l)^2} \left(1 - \frac{l}{L_i} \right) - \frac{c_1}{g^2 (k_F l)^2} \left(1 - \frac{l}{L_T} \right) \right] \quad (44)$$

where $L_T = (\hbar D / k_B T)^{1/2}$ and (Kaveh and Mott 1987a)

$$c_1 = \frac{4}{3} \nu_0 \bar{F}$$

with (Finkelstein 1983, Altshuler and Aronov 1983, Kaveh and Mott 1987a)

$$\bar{F} = 32[(1 + \frac{1}{2}F)^{3/2} - (1 + \frac{2}{3})]/3F.$$

The last term in (42) gives the $T^{1/2}$ dependence of $\delta\sigma$. Though it can have either sign, Kaveh and Mott (1987a) argue that only for many-valley uncompensated semiconductors can the Hartree term predominate. A change of sign in the $T^{1/2}$ term observed in Si:P and GeSb near the transition (Paalanen *et al* 1986, 1987), cannot be attributed to a reduction of the Hartree term—but rather to the dominance of L_i^{-1} compared with L_T^{-1} . Near the transition (Kaveh and Mott 1987a)

$$\sigma = AT^{1/3} \quad A = \frac{e^2}{\hbar} \left(\frac{c_L}{3\pi^2} \right)^{2/3} (k_B N(E_F))^{1/3}. \quad (45)$$

If we write

$$\sigma = \sigma_0 [(n - n_c) / n_c]^\nu \quad (46)$$

there seem to be two classes of material; those for which $\nu = 1$ as predicted by non-interacting theory and those for which $\nu = \frac{1}{2}$, which include Si:P (Thomas *et al* 1982) and other uncompensated many-valley semiconductors. References are given by Kaveh (1985), according to whom these are all materials for which the Hartree term dominates, with m in (42) positive (except near the transition for reasons given above). Kaveh's (1985) derivation, assuming that (41) can be extrapolated to the transition, is as follows. Near the transition we can write

$$D = c_1 / \xi + c_2 / D \xi \quad (47)$$

where the first term comes from localisation and the second from electron–electron interactions. Since $\delta\sigma_T / \sigma_B = \frac{2}{3} \delta N / N \approx 1 / D \xi$, ξ is substituted for l in equation (41) for reasons given by Kaveh (1985, p 15). Equation (47) is a quadratic equation for D , with solutions

$$D = -\frac{1}{2} \left[-\frac{c_1}{\xi} + \left(\frac{c_1}{\xi^2} + \frac{4c_2}{\xi} \right)^{1/2} \right].$$

For large ξ this becomes

$$D \approx \frac{c_1}{2\xi} + \left(\frac{c_2}{\xi}\right)^{1/2}$$

and so the square root dominates as $\xi \rightarrow \infty$. If, as we believe, $\xi \sim 1/(n - n_c)$, whether or not interactions are included, as seem to be verified by the work of Castner's school on variable-range hopping (Shafarman *et al* 1985), we find here the explanation of the result that $\nu = \frac{1}{2}$ for certain semiconductors.

If c_2 has the opposite sign, ν will be unity as in the theory without interactions.

In this model, the localisation length behaves like $(E - E_c)^\nu$ with $\nu = 1$, not $\frac{1}{2}$. The present author has given reasons why values of ν less than $\frac{2}{3}$ are unlikely (Mott 1976, 1981b), a conclusion confirmed by the work of Chayes *et al* (1986).

Since we believe that the transition takes place in the impurity band, we must suppose that many-valley behaviour persists there. How this should occur is discussed below, when we consider the effect of the Hubbard U for uncompensated material.

The theory with interactions explained in outline here depends on the extrapolation of equation (41) to the transition, and we have no proof that this can be done without qualitative error. Certainly there remain phenomena, involving the transition, which have not been explained satisfactorily. One is the behaviour of the conductivity of Si:Sb at low temperatures, observed by Long and Pepper (1984b), shown in figure 15. Some speculations about this behaviour are given by Kaveh and Mott (1985b, p 388); we return to this problem at the end of this paper.

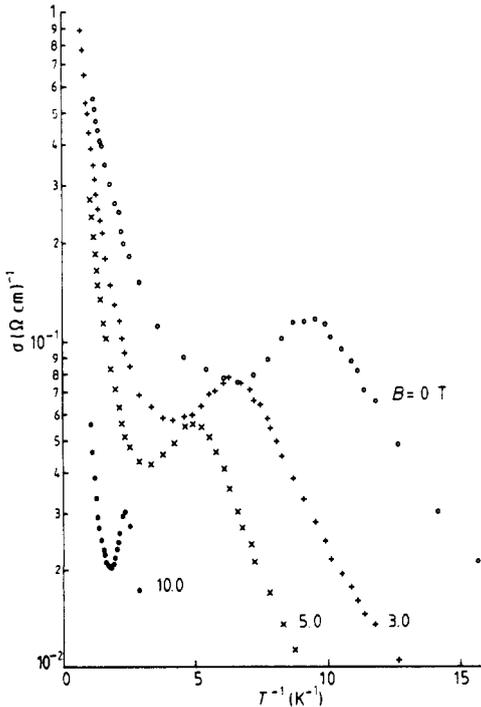


Figure 15. Conductivity σ plotted against $1/T$ for Si-Sb with dopant concentration $2.7 \times 10^{18} \text{ cm}^{-3}$ for varying magnetic field in Tesla (Long and Pepper 1984b).

For *uncompensated* semiconductors near the transition the intrasite interaction, given by the Hubbard U defined by

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

must play a role, though the evidence shows that the metal–insulator transition is of an Anderson type, driven by disorder in an impurity band, and not the discontinuous Mott transition that one would expect for an array of donors located on a lattice. The present author in various publications (e.g., Mott and Kaveh 1985b) has described this in the following way. An impurity band is split by the Hubbard U into the upper and lower bands (cf also Fritzsche 1978 and earlier papers) of which the centres are separated by an energy U and with widths B_1, B_2 . B_1 and B_2 increase as the number n of donors is increased, and in the absence of disorder a Mott transition would take place when

$$\frac{1}{2}(B_1 + B_2) = U$$

and this equation is expected to determine approximately the value of n for which metallic conduction occurs. However, in the real system, states at the edge of both bands are localised in the Anderson sense, so slightly more overlap is required to bring the Fermi energy E_F to E_c for the overlapping bands, giving a transition of an Anderson type.

Since a factor $\exp(-R/a_H)$ enters into the band widths of B_1 or B_2 and this is the major factor in determining the concentration for the transition for both compensated and uncompensated specimens, the concentrations at which they occur are not expected to differ greatly, and take place when $R/a_H \sim 4$.

At the transition, then, the wavefunctions at the Fermi energy are made up of orbitals from the upper and lower Hubbard bands. Kaveh's explanation of the value $\frac{1}{2}$ for the critical index in σ depends on some 'many-valley' properties remaining. Takemori and Kamimura (1983) have shown that this should be so for the upper band (see also Mott 1986b).

The descriptions given here, particularly of the behaviour near the transition, differs from that of Castellani *et al* (1986 and earlier papers) based particularly on the work of Finkelstein (1983). In their work they treat a system of dimension $2 + \varepsilon$, where ε is small, and extrapolate to $\varepsilon = 1$. In the paper quoted they come to the conclusion that in the absence of spin-flip and of spin–orbit scattering, a minimum metallic conductivity may exist, though they do not give its value and consider that in a later paper this is doubtful. As this—for zero temperature and in the absence of a magnetic field—has not been observed, they ascribe the observed index ($\nu = 1$) for the conductivity to the presence of spin–orbit scattering, likely in their view to be significant because of the presence of the heavy element Nb. They state (without proof) that, when spin–orbit scattering is predominant, σ at $T = 0$ goes linearly to zero.

They also discuss the value $\nu = \frac{1}{2}$ for Si:P in terms of a small spin–orbit scattering—but consider their conclusion tentative. We prefer the explanation of Kaveh (1985) described here. According to this, *compensated* Si:P should give $\nu = 1$, while the conjecture of Castellani *et al* (1987) suggests that $\frac{1}{2}$ should be retained. Experiments

† The discontinuous nature of the transition depends on the Coulomb charge on the carriers. Professor P Nozières has pointed out that the concept of a Mott transition may apply to ^3He , and here no discontinuity is expected.

at mK temperatures have not been made for compensated material, though the plot of σ against concentration for Ge-Sb (compensated) appears linear down to $\sigma \sim \sigma_{\min}$ (Thomas *et al* 1982, Thomas 1986, p 183), in contrast to that for uncompensated Si:P, which is parabolic.

Experiments on Si:P, compensated with boron, down to the transition, would be of interest. If Castellani and co-workers are correct, we might expect $\nu = \frac{1}{2}$; according to Kaveh (1985) we should find $\nu = 1$.

We now return to the problem of Si:Nb for which some results of Long and Pepper are illustrated in figure 15. Nb being a heavy element, we suggest that the diffusion length for spin-orbit scattering may be such that L_{so} is less than L_i at low T , specifically below 0.2 K. From the work of Bergmann (1983a, b) we suppose that, below this temperature, destructive interference of the scattered waves enhances the conductivity, while from Castellani and co-workers we take it that σ tends to zero linearly as n decreases. If this is so, the mobility edges with and without spin-orbit scattering should differ, as in figure 16. The upper curve gives $\sigma(T=0)$ with spin-orbit scattering, the

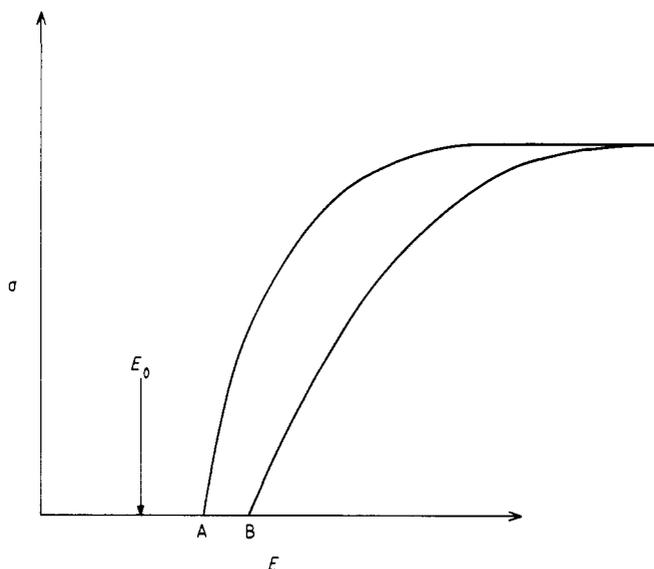


Figure 16. Zero-temperature conductivity plotted against electron concentrations in InSb, according to the hypothesis of this paper. Lower curve, without spin-orbit; upper curve, with spin-orbit scattering.

lower one gives the linear form expected without interaction. We suppose that the concentration $2.7 \times 10^{18} \text{ cm}^{-3}$ is just below A, so that below 0.1 K states are weakly localised. Above 0.1 K our hypothesis is that $L_i \ll L_{so}$, so that $\sigma(0)$ curve should have the form marked (1), so that localisation becomes stronger and the conductivity drops. The effect of a magnetic field is discussed elsewhere (Kaveh and Mott 1987c).

The same effects should be observed in Si:As—but at considerably lower temperatures, below those for which experiments have been carried out.

Acknowledgments

The author is indebted for discussions and correspondence with E A Davis, J Enderby, M Kaveh, M Pepper, W E Spear and P Thomas.

References

- Abou-Chacra R and Thouless D J 1974 *J. Phys. C: Solid State Phys.* **7** 65
- Abrahams E, Anderson P W, Licciardello D C and Ramakrishnan T W 1979 *Phys. Rev. Lett.* **42** 693
- Altshuler B L and Aronov A G 1979 *Solid State Commun.* **36** 115
- 1983 *Solid State Commun.* **46** 429
- Anderson P W 1958 *Phys. Rev.* **109** 1492
- Banyai L 1964 in *Physique de Semiconducteurs* ed. M Hulin (Paris: Dunod) p 417
- Barnes A C, Laundy B and Enderby J E 1987 *Phil. Mag.* **B 55** 497
- Belitz D and Götze W 1983 *Phys. Rev. B* **28** 5445
- Berezin E, Hikami S and Zinn-Justin J 1980 *Nucl. Phys. B* **165** 258
- Bergmann G 1982 *Phys. Rev. B* **25** 2937
- 1983a *Phys. Rev. B* **28** 2914
- 1983b *Phys. Rep.* **107** 1
- 1986 *Localization and Interaction Scottish Univ. Summer Schools in Physics* ed. D M Finlayson (Edinburgh: Edinburgh University Press) p 149
- Biskupski G 1982 *PhD 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 Remenyi G 1984 *J. Phys. C: Solid State Phys.* **17** L411
- Castellani C, Katlear G and Lee P A 1987 *Preprint*
- Chayes J T, Chayes L, Fisher D S and Spencer T 1986 *Phys. Rev. Lett.* **57** 2999
- Cohen M H, Fritzsche H and Ovshinsky S R 1969 *Phys. Rev. Lett.* **22** 1065
- Culter M 1977 *Liquid Semiconductors* (New York: Academic)
- Davies J H 1980 *Phil. Mag.* **B 41** 373
- Davis E A, Michiel H and Adriaenssens G J 1985 *Phil. Mag.* **B 52** 261
- Dersch V and Thomas P 1985 *J. Phys. C: Solid State Phys.* **18** 5015
- Djandji F and LeComber P G 1987 *Phil. Mag.* **B** at press
- Efros A L and Shklovskii B I 1975 *J. Phys. C: Solid State Phys.* **8** L44
- Elyutin T V, Hickey B, Morgan G J and Weir G F 1984 *Phys. Status Solidi b* **124** 279
- Enderby J E and Barnes A C 1987 *J. Electrochem.* in press
- Fan H Y 1951 *Phys. Rev.* **82** 900
- Fenz P, Müller H, Overhof H and Thomas P 1985 *J. Phys. C: Solid State Phys.* **18** 391
- Finkelstein A M 1983 *Sov. Phys.-JETP* **57** 97
- Fritzsche H 1978 in *The Metal Non-Metal Transition in Disordered Systems* ed. L R Friedman and D P Tunstall p 193
- Fröhlich H 1947 *Proc. R. Soc. A* **188** 521
- Fukuyama H 1985 in *Electron-Electron Interactions in Disordered Systems* ed. A L Efros and M Pollak (Amsterdam: North Holland)
- Ghazali A and Serre J 1982 *Phys. Rev. Lett.* **48** 883
- Götze W 1978 *Solid State Commun.* **27** 1393
- 1979 *J. Phys. C: Solid State Phys.* **12** 1279
- 1981 *Phil. Mag.* **B 43** 219
- Grant A J and Davis E A 1974 *Solid State Commun.* **15** 563
- Greenwood D A 1958 *Proc. Phys. Soc.* **71** 585
- Griffith R W 1977 *J. Non-Cryst. Solids* **24** 413
- Gruner G and Mott N F 1974 *J. Phys. F: Met. Phys.* **4** L16
- Hahn C E W and Enderby J E 1967 *Proc. Phys. Soc. (London)* **92** 418
- Heintze M and Spear W E 1986 *J. Non-Cryst. Solids* **78/79** 495
- Hertel G, Bishop D J, Spencer E C, Rowell J M and Dynes R C 1983 *Phys. Rev. Lett.* **50** 743

- Hikami S 1980 *Prog. Theor. Phys.* **64** 1466
 — 1982 *Anderson Localization* (Berlin: Springer) p 15
- Hikami S, Larkin A I and Nagaoka Y 1980 *Prog. Theor. Phys.* **63** 707
- Howson M A 1984 *J. Phys. F: Met. Phys.* **14** L25
- Howson M A and Greig D 1985 *Phys. Rev.* **30** 4805
- Ioffe A F and Regel A R 1960 *Prog. Semicond.* **4** 237
- Jerome D, Ryter C, Schutz H J and Friedel J 1985 *Phil. Mag.* **B 52** 403
- Jones H, Mott N F and Skinner H W B 1934 *Phys. Rev.* **35** 470
- Kakalios J and Street R A 1986 *Phys. Rev. B* **34** 6014
- Kaveh M 1985 *Phil. Mag.* **52** L1
- Kaveh M and Mott N F 1982 *J. Phys. C: Solid State Phys.* **15** L697
 — 1987a *Phil. Mag.* **B 55** 1
 — 1987b *Phil. Mag.* **B 55** 9
 — 1987c *Phil. Mag.* **B** at press
- Kawabata A 1981 *Solid State Commun.* **38** 823
 — 1982 *Anderson Localization* (Berlin: Springer) p 122
- Kawaguchi T and Fujimori Y 1982 *J. Phys. Soc. Japan* **51** 3
- Kobayashi N, Ikehata S, Koboyashi S and Sasaki S 1977 *Solid State Commun.* **24** 67
- Kramer B and King H 1983 *Z. Phys. B* **53** 1
- Kubo R 1956 *Can. J. Phys.* **34** 1274
- Landau L D 1957 *Sov. Phys.-JETP* **3** 900
- Long A P I and Pepper M 1984a *J. Phys. C: Solid State Phys.* **17** 3391
 — 1984b *J. Phys. C: Solid State Phys.* **17** L425
- MacKinnon A 1983 *J. Non-Cryst. Solids* **59/60** 73
- Marshall J M, Le Comber P G and Spear W E 1985 *Solid State Commun.* **54** 11
- Maekawa S and Fukuyama J 1981 *J. Phys. Soc. Japan* **50** 2516
- Meyer J R, Bartoli F J and Mott N F 1985 *Phil. Mag.* **B 52** L57
- Michiel H, Andriaenssens G J and Davis E A 1987 *Phys. Rev. B* at press
- Monod P and Schultz S 1982 *J. Physique* **43** 393
- Mott N F 1967 *Adv. Phys.* **16** 49
 — 1972 *Phil. Mag.* **26** 1015
 — 1974 *J. Phys. F: Met. Phys.* **4** L46
 — 1976 *Commun. Phys.* **1** 200
 — 1981a *Phil. Mag.* **43** 941
 — 1981b *Phil. Mag.* **B 44** 265
 — 1984 *Phil. Mag.* **B 49** L75
 — 1985a *Phil. Mag.* **B 51** 19
 — 1985b *Phil. Mag.* **B 52** 169 (Erratum 1986a *Phil. Mag.* **B 53** 91)
 — 1986b *Localization and Interaction* ed. D M Finlayson (Edinburgh: Edinburgh University Press) p 29
- Mott N F and Davis E A 1979 *Electronic Processes in Non-Crystalline Materials* 2nd edn (Oxford: OUP)
- Mott N F and Kaveh 1985a *Phil. Mag.* **B 52** 177
 — 1985b *Adv. Phys.* **34** 329
 — 1987a *Phil. Mag.* **B 55** 1
 — 1987b *Phil. Mag.* **B 55** 9
- Müller H and Thomas P 1984 *J. Phys. C: Solid State Phys.* **17** 5337
- Paalanen M A, Sachdev S and Bhatt R N 1987 *Proc. Int. Conf. Physics of Semiconductors (Stockholm) 1987* in press
- Paalanen M A, Sachdev S, Bhatt R N and Ruckenstein A E 1986 *Phys. Rev. Lett.* **57** 2061
- Pollitt S 1976 *Commun. Phys.* **1** 207
- Rosenbaum T F, Andres K, Thomas G A and Bhatt R N 1980 *Phys. Rev. Lett.* **45** 1723
- Overhof H 1986 *Preprint*
- Overhof H and Beyer W 1983 *Phil. Mag.* **B 7** 377
 — 1984 *Phil. Mag.* **B 49** L9
- Shafarman W M, Castner T G, Brooks J S, Martin K P and Naughton M J 1986 *Phys. Rev. Lett.* **56** 980
- Shapiro B 1984 *Phil. Mag.* **B 50** 241
- Silver M, Adler D, Shaw M P and Cannella V D 1986 *Phil. Mag.* **B 53** L89
- Spear W E 1983 *J. Non-Cryst. Solids* **59/60** 1
 — 1987 private communication

- Spear W E, Allan D, LeComber P G and Gaith A 1980 *Phil. Mag.* B **41** 419
- Spear W E and LeComber P E 1985 *Phil. Mag.* B **52** 247
- Takemori T and Kamimura H 1983 *Adv. Phys.* **32** 715
- Thomas G A 1986 in *Localization and Interaction* ed. D M Finalyson (Edinburgh: Edinburgh University Press) p 183
- Thomas G A, Ootuka Y, Katsumoto S, Kobayashi S and Sasaki W 1982 *Phys. Rev.* B **25** 4288
- Thomas G A, Paalanen M A and Rosenbaum T F 1983 *Phys. Rev.* B **27** 3897
- Thomas P 1987a unpublished
- 1987b *J. Non-Cryst. Solids* in press
- Thomas P and Overhof H 1987 private communication
- Thouless D 1970 *J. Phys. C: Solid State Phys.* **3** 1559
- Tiedje T, Cebulka J, Morel D and Abeles B 1981 *Phys. Rev. Lett.* **46** 1425
- Vardeny Z, Tauc Z and Fang C J 1981 *J. Physique* **42** C4 539
- Wegner F 1979 *Z. Phys.* B **35** 207
- 1982 *Anderson Localization, Springer Series in Solid State Science* vol 39, ed. Y Nagaoka and H Fukuyama (Berlin: Springer) p 8