METAL-INSULATOR TRANSITIONS

N.F. Mott

Cavendish Laboratory, Cambridge University, Cambridge, England

Abstract – Several kinds of metal-insulator transitions are discussed.
In non-crystalline systems a second order "Anderson" transition can be
described without considering electron-electron interaction. In cry-
stalline systems, we describe band-crossing transitions and "Mott" tran-
sitions, both of which should show a discontinuity in the number of
carriers when electron-electron interaction is included. The difficulty
of observing this directly is stressed. Sufficient disorder in the sys-
tem will remove the discontinuity. Fluid caesium, metal-ammonia and
expanded fluid mercury are contrasted with Si:P. Some crystalline com-
pounds of transition metals Ti_2O_3, VO_2, V_2O_3, NiS_2 and NiS are described,
each of which seems to involve its own different mechanism. A new sug-
gestion is made for Ti_2O_3. One other kind of transition is briefly
mentioned, the Verwey transitions in Fe_3O_4.

1. INTRODUCTION

In this talk I use the word "metal" to denote a material in which the resistivity tends to
a finite value (or zero) as the temperature decreases and "insulator" to denote one in which
it tends to infinity. According to this definition, a heavily doped semiconductor is a
"metal", a lightly doped one an "insulator". The transition from one state to another can
occur at low temperatures under pressure, stress, magnetic field or in alloys with change
of composition and in inversion layers with change of gate voltage. First order phase
changes between states with semiconducting and (apparently) metallic behaviour also occur
in many materials with increasing temperature, T, but in such cases, since our definition
of "metal" depends on the behaviour as T tends to zero, it is obviously not possible to
distinguish with certainty between a metal and a small-gap semiconductor. However, in
most of these materials the transition can be induced at low temperatures by pressure or
by alloying. In the latter case, since the alloys under consideration are disordered,
some discussion of the movement of electrons in a non-periodic field is necessary, and will
be given in the next section.

2. NON-CRYSTALLINE SYSTEMS

I will start therefore with a brief summary of the behaviour of electrons in a non-periodic
field. Our understanding depends on Anderson's (Ref. 1) paper of 1958, "Absence of diffu-
sion in certain random lattices", which uses the tight binding potential shown in Fig. 1,
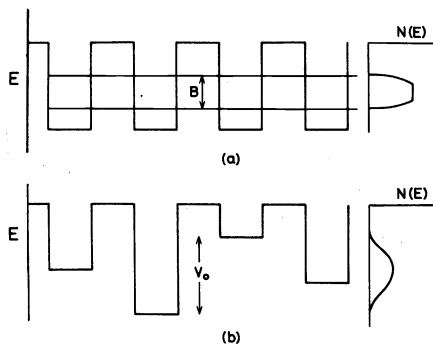

Fig. 1. Potential energy and density of states N(E) in the Anderson model.
(a) without and (b) with a random potential V_o.

in which atomic orbitals overlap with nearest neighbours only, leading to a narrow band of
energy levels of width B. Anderson then introduces a random potential V_0 at each well, and
finds that, if V_0/B is greater than a critical value, now thought to be of order 2 for
coordination number 6, all states are "localized"; that is, they become traps. If $V_0/B < 2$,
there exists a "mobility edge", which is an energy E_C, separating energies in the tail of
the band for which states are localized from those in which they are non-localized (extended).
The position of the mobility edge has been calculated for the Anderson model by Abou-Chacra
and Thouless (Ref. 2) and for other forms of disorder by Debney (Ref. 3) and by Davies (Ref.
4). For various systems in which an electron gas is degenerate at low T, the Fermi energy
E_F and/or the degree of disorder can be changed, so that E_F passes through E_C; a metal-
insulator transition takes place. Examples are, impurity-band conduction in doped silicon
where the degree of compensation can be changed, or conduction at a Si/SiO_2 interface in an
inversion layer in a MOSFET device, where the variable is the gate voltage (Refs. 5 and 6).
The "Anderson" transition that results is by now rather familiar, and is illustrated in
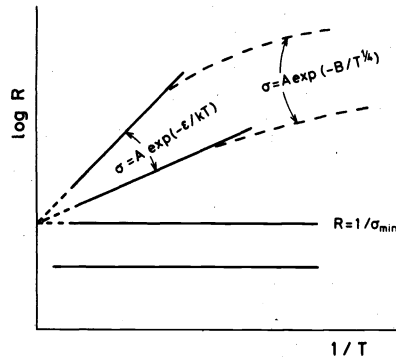Fig. 2, showing resistivity as a function of 1/T. For this lecture, the important points

Fig. 2. Resistivity R as a function of 1/T for a system showing an
Anderson transition.

are:

a) In the insulating range at high temperatures, the conductivity behaves like

$$\sigma = \sigma_{min} \exp(-\varepsilon/kT), \quad \varepsilon = E_C - E_F, \tag{1}$$

and ε goes continuously to zero as we approach the transition. At low temperatures conduc-
tion is by variable-range hopping, with

$$\sigma = A \exp(-B/T^{\frac{1}{4}}). \tag{2}$$

b) At the transition, the conductivity is σ_{min} (the minimum metallic conductivity),
given by

$$\sigma_{min} \simeq 0.05 \, e^2/\hbar a, \tag{3}$$

where a is the range of the fluctuations in the random potential. The theoretical value
of the numerical terms is uncertain by a factor of at least 2, depending on the co-ordina-
tion number etc. The existence of this quantity is somewhat controversial, and is not
confirmed by some theoretical work, (Ref. 7) but the experimental evidence for it is
strong (Ref. 8) and is confirmed over a wide range of a by the work on impurity bands. If
$a \sim 3 \, \text{Å}$, σ_{min} is $\sim 500 \, \Omega^{-1} \, cm^{-1}$. In two-dimensions, for instance in an inversion layer,
$\sigma_{min} = 0.1 \, e^2/\hbar$.

c) The theory leading to these results is that of non-interacting electrons.
The electrical behaviour in my opinion is not greatly affected by interaction, but the
magnetic behaviour might be. The effect of interaction between electrons is not fully
understood, and some other phenomena, such as the behaviour of the Hall coefficient, remain
quite unexplained (Ref. 9). The relevance of these findings to some crystalling systems
containing transition metal ions will appear later.

3. BAND CROSSING TRANSITIONS

Metal-insulator transitions in crystalline systems, the main subject of this lecture, are
of many types. Of these the simplest is the band-crossing transition, shown in Fig. 3.
Examples are ytterbium under pressure (Refs. 10 and 11) Bi-As alloys (Ref. 12) and Ti_2O_3
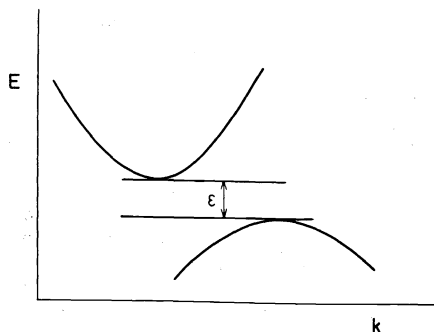with increasing temperatures (Ref. 13).

Fig. 3.   Energy as a function of κ for a band-crossing transition.

In a model without electron-electron interaction, the activation energy ε for conduction would go continuously to zero, but theory now suggests fairly conclusively that, when interaction is taken into account, there should be a discontinuous change both in $\underline{n}$ (the number of current carriers) and ε (the activation energy for conduction), $\underline{n}$ changing from zero to a value of the order

$$n^{1/3} \sim 0.1\ me^2/\kappa\ \hbar^2. \tag{4}$$

κ here is a background dielectric constant. In the metallic phase the material is a semi-metal, containing $\underline{n}$ electrons and $\underline{n}$ holes per unit volume. The argument (Ref. 14) in its simplest form is that the kinetic energy of such particles is const $(\hbar^2/m)n^{2/3}$, and their potential energy $-e^2n^{1/3}/\kappa$. The total energy has a minimum for a value of n given by (4), and the energy per pair of particles is thus $-E_0$, where $E_0 = $ const. $me^4/\hbar^2\kappa^2$, all constants depending on the band form. If the band gap ΔE is changed in some way, for instance by alloying, the discontinuous transition will occur when $E_0 = \Delta E$.

Since the change in the number of carriers is discontinuous, we would expect the plot of free energy against volume or composition in an alloy to show a kink as in fig. 4.
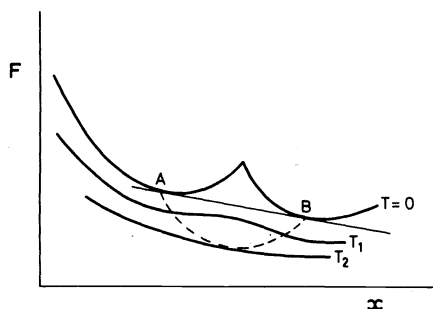

Fig. 4.   Free energy against x (=volume or composition) at a metal-insula-tor transition for T = 0 and increasing temperatures T_1, T_2. The dotted line shows a two-phase region.

Consequently at low temperatures under pressure there must be a discontinuous change in volume (from A to B) and in alloys a two-phase region. The only hope of seeing the discontinuity is to work with quenched alloys. Up till now I do not know of any direct observation of the discontinuity. The potential in a quenched alloy necessarily contains a random term, and this may suppress the discontinuity, the transition then being of Anderson type. I have given (Ref. 15) a discussion of the degree of disorder for this to be so.

The case of Ti_2O_3 is particularly interesting in this connection. The material (Refs. 13, 16 and 17) has a small gap between valence and conduction d-bands of a_{1g} and e_g (π) character, and at ∿ 400 K makes a rapid (but not first order) transition to a state with "metallic" conductivities of about 4000 Ω^{-1} cm^{-1}. At the same time there is a change in c/a in the corundum structure. If V_2O_3 is added, the same change in c/a occurs and for 10% of vanadium the material appears metallic. It was long thought that the change in c/a was the cause of the transition. But it is now clear that the vanadium goes in initially in the state $3d^3$, producing striking magnetic properties and is compensated by Ti^{4+} ions giving an acceptor impurity band. Metallic behaviour then occurs for quite low (∿1%) concentrations of V. But the sudden change of conductivity at ∿ 400 K persists (Ref. 7), showing that the gap has not shrunk, and this has recently been confirmed by optical measurements by Lucovsky (Ref. 18). Thus while holes in the valence of impurity band lead to a change in c/a, a change in c/a appears not to be the cause of the transition in Ti_2O_3

at 400 K. A possibility is, suggested here for the first time, that in Ti_2O_3 ΔE is near the critical value E_O defined above. The electrons (not the holes) in this material are known to have a large effective mass, so with increasing T the entropy (S) will increase and the free energy TS decrease. At 400 K, I postulate that F_O, the free energy of the electron-hole gas, becomes equal to the gap ΔE. If the electron gas is still degenerate at 400 K, the term TS will behave like $-T^2n^{1/3}$. Thus the negative term is

$$-(e^2/\kappa + const\ T^2)\ n^{1/3},$$

and increasing temperature leads to a free energy of the electron-hole gas of the form

$$E_o \sim (m/\hbar^2)\ (e^2/\kappa + const\ T)^2.$$

Further analysis is needed to show if a quantitative explanation can be obtained in this way.

4. MOTT TRANSITIONS

The next kind of transition that I shall consider is the "Mott Transition", which occurs when the distance between one-electron centres is varied. In 1949 I gave arguments (Ref.19) based on screening to show that such a transition must be continuous. I prefer now the formulation by Brinkman and Rice (Ref. 20); for large separations an antiferromagnetic sub-lattice is formed, which splits the conduction band into an upper and lower "Hubbard band", (Ref. 21) respectively empty and fully occupied by electrons. The transition is then similar to a band-crossing transition, occuring when the two bands cross. It should be from an antiferromagnetic insulator (AFI) to an antiferromagnetic metal (AFM), if the interaction between electrons and holes is neglected. If it is not, then again as for other band-crossing transitions a discontinuous change in n and ε is expected. The question then is, is an AFM region expected? I have estimated (Ref. 14, p.137) that the moments disappear when the number of carriers is about 1/10 of the number of atoms, and Rose, Shore and Sanders (Ref. 22) find a second order disappearance of the moments when r_s, the atomic radius, is about 4% below that for the transition. Further investigation is needed to see if the two results are comparable.

         If the change of n at the transition is greater than this, there should be no AFM region. The two types of behaviour are shown by NiS_2 (Krill et al) (Ref. 23) and V_2O_3 (for refs. see 12). For $NiS_{1.9}$ the disappearance of the moments under pressure is observed by Panissod et al. (Ref. 24).

         When the AFM makes its transition to a normal metal through the disappearance of moments, then according to the analysis of Brinkman and Rice (Ref. 25), the electron gas is "highly correlated" -- in the sense that only $\sim 10\%$ of the atomic sites carry pairs of electrons. This, as is intuitively obvious, greatly enhances the Pauli paramagnetism, the material behaving as if $N(E_F)$ were increased. The electronic specific heat is also increased, and the electrical resistance shows an abnormally large T^2 term of the type first discussed by Baber (Ref. 26). All these properties shown by V_2O_3 when the metallic phase is stabilized by pressure, but not by VO_2 and NiS; about NiS_2 the information does not exist.

         The description of a discontinuous transition given here is not the only one in the literature. We note particularly the model of Leroux Hugon and Ghazali (Ref. 27), developed for Si:P, in which the AFI lattice is shown to undergo a sudden instability due to the Clausius-Mossotti term as the density increases. The connection between the two models has yet to be investigated.

         When a transition from non-metal to metal occurs with increasing temperature, some of the entropy driving the transition is electronic, and some due to phonons. For V_2O_3, where in the metallic state the electron gas is highly correlated, the electronic entropy is the major term, as is also the case for Ti_2O_3, a band-crossing transition where the electrons have a large effective mass. For VO_2 and NiS, however, the softening of some phonon modes make the major contribution.

         For Mott transitions occurring with change of composition, as for band-crossing transitions, we have already mentioned the problem of whether for a quenched system, in view of the disordered nature of the alloy, there should be a discontinuity. The case of Si:P, much discussed in the literature, seems to have the characteristic of an Anderson transition; no discontinuity in ε is observed (Refs. 14 and 28). On the other hand some fluid systems, metal-ammonia solutions and expanded fluid caesium, show the solubility gap and critical points associated with the behaviour of Fig. 4. I believe that this is because the disorder is more like that in a liquid than that in a gas, and therefore smaller.

         Expanded fluid mercury has been extensively investigated by Professor Hensel and co-workers (Refs in 14). I argue that this is a band-crossing transition, and also that κ, the background dielectric constant, is larger because the electronic transition involves a charge transfer from one atom to another in the Mott transition. Thus any predicted critical point associated with the metal-insulator transition could lie at temperatures where the material is not fluid. It is interesting that the critical point for mercury does not lie near the metal-insulator transition, the conductivity there being about

$10^{-3} \ \Omega^{-1} \ cm^{-1}$ and much less than σ_{min}, while for Cs and metal-ammonia it is close to σ_{min} ($\sim 500 \ \Omega^{-1} \ cm^{-1}$).

5. SOME OXIDES AND SULPHIDES

I turn now to the materials VO_2, V_2O_3, NiS_2 and NiS, all of which show an insulator-metal transition (or semimetal-metal) with increasing temperature. In all cases the transition can be induced by alloying, in VO_2 with CrO_2, in V_2O_3 with Ti_2O_3, in NiS_2 with $NiTe_2$ and in NiS by excess S. In V_2O_3 and NiS it occurs under pressure.

V_2O_3 below the transition is an AFI and what has been called a "Mott Insulator". I would like to say there is no mystery about these materials and at T = 0 they can be described by formal band theory, if different wave functions are used for the two spin direction. This was shown by Slater (Ref. 29) in 1951, the AF lattice being supposed to split the zone into full and empty sub-bands. There was an old controversy about whether such materials should become metals above the Néel point; of course in reality usually they do not, although the AF lattice has disappeared. This problem has been extensively reviewed by Brandow (Ref. 30). I think the best way to look at this is to remember that, for an insulator, the Slater determinant using Bloch function is <u>identical</u> with that using localized Wannier functions, and one can of course have a gap when the latter, for a given spin direction, are not in a periodic array, just as one can have a gap in glasses.

The metal-insulator transition in V_2O_3 cannot of course be described without detailed discussions of the band structure, which is given in many publications and will not be reproduced here. I think the strongest evidence that the moments play an essential part and that the analysis as a Mott transition is relevant is the high specific heat γT and large T^2 term in the resistivity which the metallic phase shows, when stabilised by pressure.

I turn now to the case of VO_2. In one-electron systems the tendency to form pairs is always striking (metal-ammonia, expanded fluid caesium, Ti_4O_7 as described in §7), and in VO_2 by a band model. However, according to the treatment given by Zylbersztejn and Mott (Ref. 31), VO_2 would certainly be an (antiferromagnetic) insulator without pairing, the energy gap being of order of the Hubbard U; the pairing then has a minor effect on the energy and the gap. The metal-insulator transition is, these authors suggest, due to a shortening of one VO distance by distortion, and this raises the π^* oxygen states, which no longer mix with the $d_\parallel$ so either increasing U or decreasing the band width.

If this is correct, the transition in VO_2 is caused by a change in the degree of s-p hybridisation. We shall see that the same is true of NiS. We suggest now that in both these materials, in a rigid lattice, one may write the d-orbitals in the form

$$\psi_d + A \ \psi_p$$

where ψ_p is the overlapping oxygen orbital. As A decreases, the Hubbard U increases. The usual expression for the appearance of moments (Ref. 14, p.94)

$$a^3 U \ N \ (E_F) > 1$$

will lead to the formation of moments for a certain value of $U \ N \ (E_F)$. The plot of energy against A^2 should thus be as in Fig. 5, there being two minima. The lattice parameters will
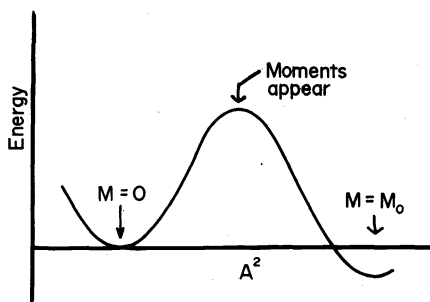

Fig. 5. Suggested curve of energy of a magnetic material as a function of A^2.

not of course be identical for the two minima. But the transition must be thought of primarily as one in which the system jumps from one electronic state to another. One is normally antiferromagnetic, but in VO_2 this property disappears because of pairing; the other is a normal metal, not a highly correlated one.

Turning now to NiS_2, the recent p-T diagram of Panissod et al is reproduced in Fig. 6, showing the region of the AFM. There is little change in the Néel temperature
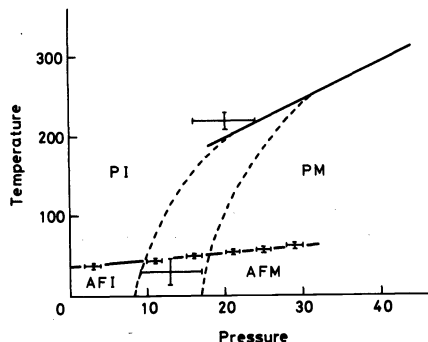
Fig. 6. Phase diagram (P, T in K bars and degrees C) for NiS$_{1.9}$. The solid line is deduced from measurements of resistivity under pressure (Moro et al, Ref. 35). The dashed lines show uncertainty in transition pressure (Panissod et al, Ref. 24).

between the AFI and the AFM. Measurement of electronic specific heat and resistivity where the antiferromagnetism has just disappeared would be of great interest, to see if the behaviour of a highly correlated gas is reproduced.

6. NICKEL SULPHIDE

NiS, recently investigated in some detail by Brusetti et al (Ref. 32), shows the resistivity-temperature curve shown in Fig. 7. The transition at low T also occurs as a function of
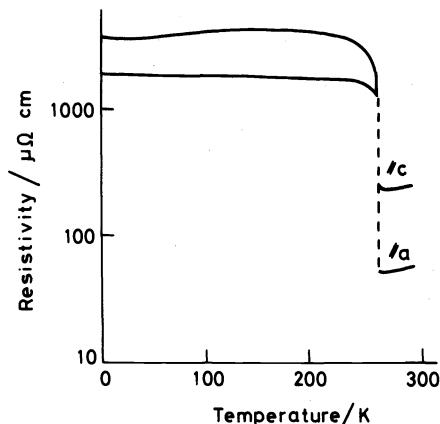

Fig. 7. Resistivity between 4.2 and 3000 K of stoechiometric NiS parallel to c and a axes (Brusetti et al, Ref. 32).

pressure, or composition in Ni$_x$S when x is about 0.95, or with the substitution of rhodium (Ref. 33) and other metals. The low temperature phase is antiferromagnetic, but the transition at 260 K, where the AF lattice disappears and the conductivity increases, is not a Néel point in the normal sense, the true Néel point lying at ∿ 1000 K. The low temperature phase appears to be a semi-metal, the conduction band being split by the AF lattice, as envisaged by Slater (Ref. 29), but the two bands overlapping. The absence of a variation of resistance with temperature could be characteristic of a "metallic" impurity band, but the evidence seems against this, particularly the very low values of the thermopower, except near the transition (Fig. 8). The phenomena are by no means fully understood but the increase with T suggests a semi-metal. If so, the very strong scattering implied by the constant resistivity, perhaps due to defects, needs theoretical investigation.

The metallic phase, however, is not magnetic and does not have the characteristics of a highly correlated gas. There is no structure change at the transition, but a slight contraction and softening of some phonons, which contribute more than half the entropy. Following White and Mott (Ref. 35) we think then that what has to be considered here is a discontinuous disappearance of the moments. As proposed in the last section for VO$_2$, it is suggested that overlap on to the d states by sulphur 3p orbitals can greatly change the Hubbard U, so that two electronic states are possible as in Fig. 6, one with moments and one without. The situation will be as in VO$_2$, except that the state with moments is an AFM.
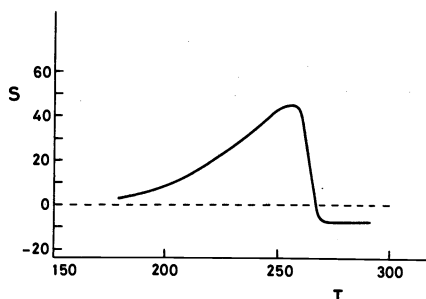
Fig. 8. Resistivity against 1/T for Fe_3O_4 and Ti_4O_7.

7. VERWEY TRANSITIONS

Finally there is a further very interesting type of transition, mainly I believe occurring in oxides. This is the so-called Verwey transition, of which Fe_3O_4 is typical. Here the "B-sites" are occupied by equal numbers of Fe^{3+} ions, and below 220 K they order giving a non-metallic state, while above that temperature there is no long-range order, and the conductivity approaches that of a metal. An unexplained property is that about 1% of impurity freezes-in the metallic behaviour down to the lowest temperatures. I think that this ordering of the carriers will occur only with a narrow d band, and this is more likely to be found in oxides than, for instance, sulphides. I have given a discussion of the experimental situation, and of a model for the high-temperature state in Ref. 36.

An even more intriguing behaviour (Refs. 37 and 38) is shown by Ti_4O_7, where equal numbers of Ti^{3+} and Ti^{4+} ions order at low temperatures. This shows two transitions In the intermediate phase the material is diamagnetic, and it is concluded that the carriers are mobile polarons, that is to say pairs of Ti^{3+} ions held together by covalent forces which decrease the distance between them, so that there is an activation energy for movement. This is another example of the tendency of one-electron centres to form pairs.

REFERENCES

1. P.W. Anderson, Phys. Rev. 109, 1492 (1958).
2. R. Abou-Chacra and D.J. Thouless, J. Phys. C, 7, 65 (1974).
3. B.T. Debney, J. Phys. C, 10, 4719 (1977).
4. J. Davies, Phil Mag (in press).
5. N.F. Mott, M. Pepper, S. Pollitt, R.H. Wallis, and C.J. Adkins, Proc. Roy. Soc. A 345, 169 (1975).
6. M. Pepper, Proc. Roy. Soc. A, 353, 225 (1977).
7. D. Weaive and B. Kramer, J. non-cryst. Solids 32, 131 (1979).
8. M. Pepper, J. non-cryst. Solids, 32, 161 (1979).
9. M. Pepper, Phil Mag 37, 83 (1978).
10. D.B. McWhan, T.M. Rice and Ph.D. Schmidt, Phys. Rev., 177, 1063 (1969).
11. R. Jullien and D. Jerome, J. Phys. Chem. Solids, 32, 257 (1971).
12. E.S. Itskevich and L.M. Fisher, Sov. Phys. JETP, 26, 66 (1968).
13. J. Dumas and C. Schlenker, J. de Phys. C., 37, 41 (1976).
14. N.F. Mott, Metal Insulator Transitions, Taylor & Francis, London (1974).
15. N.F. Mott, Phil Mag., 37, 377 (1978).
16. J. Dumas, Thesis, Grenoble (1978).
17. G.V. Chandrashekhar, Q. Won Chui, J. Moyo and J.M. Honig, Mat. Res. Bull., 5, 999 (1970).
18. G. Lucovsky, priv. comm.
19. N.F. Mott, Proc. Phys. Soc. A, 62, 416, (1949).
20. W.F. Brinkman and T.M. Rice, Phys. Rev. B, 1508, (1973).
21. J. Hubbard, Proc. Roy. Soc. A, 281, 401, (1964).
22. J.H. Rose, H.B. Shore and L.M. Sanders, Phys. Rev. (to be published).
23. G. Krill, M.F. Lapierre, F. Gautier, C. Robert, G. Czjezek, J. Fink and H. Schmidt, J. Phys. C, 9, 761 (1976).
24. P. Panissod, G. Krill, C. Vettier and R. Madar, Solid State Comm. 29, 67 (1979).
25. W.F. Brinkman and T.M. Rice, Phys. Rev. B 2, 4302 (1970).
26. W.G. Baber, Proc. Roy. Soc, A, 158, 303 (1937).
27. P. Leroux-Hugon and A. Ghazali, Phys. Rev. B, 14, 602 (1976); Phys. Rev. Ltt. 41, 1569 (1978).
28. N.F. Mott and E.A. Davis, Electronic Processes in non-crystalline materials, 2nd ed. (1979).
29. J.C. Slater, Phys. Rev., 82, 538 (1951).

30. B.H. Brandow, Adv. in Phys. 26, 657 (1977).
31. A. Zylbersztejn and N.F. Mott, Phys. Rev. B11, 4383 (1975).
32. R. Brusetti, J.M.D. Coey, J. Czjzek, J. Fink, F. Gompf, J. Phys. F. (in press)
33. T. Sawa and S. Anzai, J. appl. Phys. 49, 5612 (1978).
34. N. Moro, T. Mitzui and S. Yomo, Solid State Comm. 13, 1803 (1973).
35. R.M. White and N.F. Mott, Phil. Mag. 24, 845 (1971).
36. N.F. Mott, Festkorperprobleme (in press).
37. S. Lakkis, C. Schlenker, B.K. Chakraverty, R. Buder and M. Marezio, Phys. Rev. B14, 1429 (1976).
38. M. Gourmaler, C. Schlenker and J. Mercier, 9th Int. Conf. on Physics of Semiconductors, Edinburgh 1978, p. 469.