PHYSICAL REVIEW B VOLUME 20, NUMBER 4 15 AUGUST 1979

Analysis of the pretransition range of the metal-insulator transition in doped semiconductors

G. F. Neumark
Philips Laboratories, Briarcliff Manor, New York 10510

(Received 30 June 1977)

Various theories of the metal-insulator transition have been considered in regard to their likely importance in the doping range prior to this transition, and screening appears to dominate. It is then shown that screening—through its effect on impurity energies and radii—can explain the various activation energies in this range. This is consistent with the recent conclusion that the screening model can explain the temperature dependence of Hall data.

I. INTRODUCTION

The metal-insulator (MI) transition in doped semiconductors, first noted by Hung and Gliessman[1] in 1950, has since been extensively studied, with recent comprehensive reviews by Mott *et al.*[2] One observes that at a particular ("critical") doping density (N_D^c) the low-temperature conductivity changes from an activated behavior to a temperature-independent conductivity. By definition, a material whose conductivity is activated is termed "insulating", whereas one with no temperature dependence is metallic, and hence one has an MI transition. Also of interest are some of the details of the activated behavior as one approaches the transition. It was shown by Fritzsche[3] that, for doping densities less than N_D^c, the conductivity (σ) can be approximated by,

$$\sigma \approx \sigma_1 e^{-\epsilon_1/kT} + \sigma_2 e^{-\epsilon_2/kT} + \sigma_3 e^{-\epsilon_3/kT} , \tag{1}$$

where $\sigma_1 >> \sigma_2$, σ_3, and $\epsilon_1 > \epsilon_2 > \epsilon_3$. It is generally assumed[2] that the energy ϵ_2 is the one relevant to the above MI transition, which takes place in the limit of $\epsilon_2 \rightarrow 0$. The process responsible for ϵ_2 is likely some form of interimpurity transport. The detailed meaning of ϵ_2 is a main topic of the present paper. The energy ϵ_1 is identified[2] as that from an impurity into the conduction band, and ϵ_3 is the energy required for hopping[2] from a neutral to an ionized majority impurity.

There are many prior theoretical analyses of the MI transition[2]—each with use of different approximations for the potential term in the Hamiltonian. One of the earliest treatments of this problem was by Mott[4] in 1949. He considered the screened Coulomb interaction among charges in an assembly of positive ions and electrons. This problem has lately been extensively analyzed for various screening potentials and various complications due to the detailed band structure.[5-11] In all these cases,[4-11] with strong screening, which occurs for a high density of ions and electrons, there are no bound states in such a system. Thus, metallic behavior ensues. Some later approaches are based on the interaction of two electrons on the same site,[2,3] but neglect Mott's (original) long-range screened Coulomb interaction. These later analyses use a Hamiltonian introduced by Hubbard,[12] and will thus be referred to as the Hubbard model. Interestingly, it was shown by Johansson[13] that the two treatments (Mott and Hubbard) are mathematically equivalent at the transition point. Nevertheless, there is no apparent physical equivalence between the short-range correlation of two electrons on one site and the (long-range) screening of a sea of electrons. In a recent paper[14] we have briefly argued that the two approaches lead to different results in the doping range just below the transition, with the screening model giving improved insight into, and agreement with, experimental Hall results. Another interaction which has been considered in the literature[15-17] is that induced by the bound-electron response (polarization) to the potential of a center, i.e., a change in dielectric function. However, here again screening appears to dominate in the *pretransition* range. This is shown in Fig. 1, which compares the energy reductions predicted by several non-Hubbard theories. The energy (E_A) normalized to the energy at infinite dilution (E_0) is shown as a function of net impurity concentration (N) normalized by the impurity Bohr radius at infinite dilution (a_0). The polarization theory of Hugon and Ghazali[17] is compared with screening results given by Krieger and Nightingale[5] for two different screening functions (Thomas-Fermi and Lindhard). Over most of the range of finite E_A/E_0 both screening functions give an appreciably greater reduction in energy than the polarization effect. Lastly, an important approach[2] to the MI transition is the one which considers the disorder due to random location of the impurities and consequently of the random nature of the potential. This has two consequences.[2] First, it affects the density of states. Second, it can lead to an Anderson[18] transition, i.e., to mobility edges[2] as well as band edges. Here we

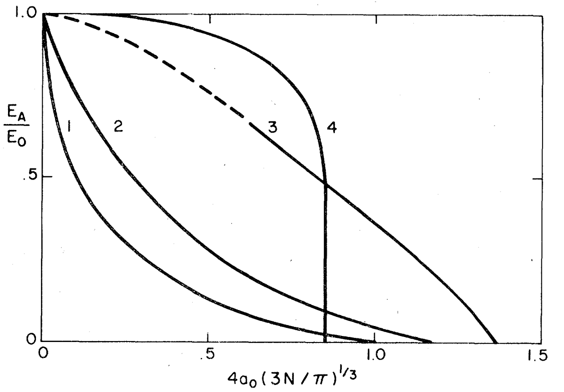

FIG. 1. Energies (E_A) predicted by various theories vs net doping concentration (N). Appropriate normalized quantities are used, namely, E_A/E_0 and $a_0N^{1/3}$, where E_0 is the energy and a_0 the Bohr radius at infinite dilution. Curves 1 and 2 are for the effect of screening, curve 1 for Thomas-Fermi and curve 2 for dielectric (from Ref. 5). Curve 3 is for disorder (from Ref. 20); since no results are given in Ref. 20 for low doping, this part is shown dashed. Curve 4 is for the effect of polarization (from Ref. 17).

shall neglect the second effect, which essentially means that we assume the relevant mobility edge to be very close to the (semiconductor) band edge. The reason for this assumption is that, as discussed earlier,[14] a mobility edge would give carrier activation, and such carrier activation appears incompatible with Hall results. Calculations on the density of states include those of Matsubara and Toyozawa[19] for unscreened Coulomb potentials, with an improved calculation of this type by Chao.[20] In Fig. 1 we also show the results of Chao[20] (using the peak of his distribution as E_A/E_0) as representing disorder alone, without screening. It can be seen that, again, screening gives the stronger effect. It thus appears, both from Fig. 1 and from our earlier work,[14] that screening is the dominant effect in the pretransition range. Nevertheless, it remains to be shown that the activation energies (and some other parameters) of the screening model give reasonable agreement with experimental values, such as those reported by Davis and Compton[21] on Ge(Sb). It is the aim of the present paper to investigate this point.

It is well known that screening reduces the binding energy of an effective-mass-type (hydrogenic) impurity and in fact it was shown[22] that this process accounts for the reduction of the ϵ_1 activation energy in the "intermediate" doping range (see Ref. 22 for the definition of "intermediate"). As the doping increases beyond this range towards the MI transition, some additional effects must also be considered: (i) The impurity distribution broadens. (ii) There can be a transfer of an electron from one donor to another, giving a state with two electrons on one impurity (or holes, in p material), the so-called D^- level.[3,23] (iii) There will be an increase in the effective radius of the impurity.[6] As a consequence of the latter two effects, carrier transfer among the impurities becomes more and more likely with increasing doping. In the present analysis one key aspect is that this transfer gives increased screening. A second important aspect is that the carrier transfer can be a hopping process. Thus, the energy ϵ_2 can be attributed to an activated *mobility*.[24] As shown earlier,[14] an activated mobility can explain the data reported by Davis and Compton[21] on Hall coefficients of Ge(Sb), whereas these data cannot be explained with the usual[2] Hubbard or Hubbard–Anderson approach.

The model outlined above is presented in Sec. II. Activation energies and conduction processes are reviewed in Sec. II A, and the detailed role of screening, with a resultant separation into two regions, is discussed in Sec. II B. It is then shown (Sec. III) that this model leads to a good description of the Hall data[21] in the pretransition range. As a speculation, we also point out (Sec. III) that the model may explain persistent photoconduction. Some further discussion and a summary are presented in Sec. IV.

II. MODEL

A. Energies and conduction processes

In any semiconductor system, there is the usual impurity to conduction-band activation energy E_A, with a corresponding "normal" conduction-band conductivity. Moreover, there can be a distribution of such activation energies around an average value E_A, with a half-width ΔE.

An additional factor which must be considered is the possibility of forming donor "molecules", with a resultant ionized state. Specifically, if one considers two donors, an electron can be transferred from one to the other, leaving a hole behind

$$2D^0 \rightleftarrows D^- + D^+ . \tag{2}$$

This is the so-called D^- state[3,23] (the hole part is usually neglected). The excitation energy E_u to this state is[2]

$$E_u \approx E_A - Q , \tag{3}$$

where Q is the electron affinity of the D^- state. The energies E_A and E_u are shown, schematically, in Fig. 2(a) for the case of low doping. The situation at higher doping, assuming $\Delta E >> E_u$, is shown in Fig. 2(b). Regarding conduction, this is possible once carriers are excited to these states. Such conduction can be either band- or hopping-type. Band-type is through extended states, and has often been postulated[2] for the upper Hubbard band or for excitation

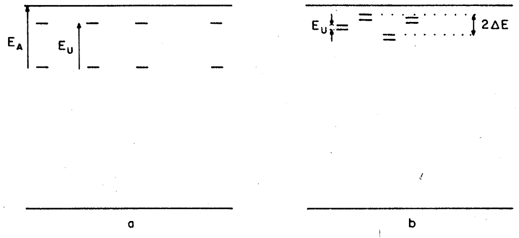

FIG. 2. Schematic of the relevant energies. (a) shows the energy to the conduction band (E_A) and that to a state with two electrons per donor, the D^- state (E_u). It is drawn for low doping and assumes that the energy spread (ΔE) among the different impurity levels is too small to show up on the scale of this figure. (b) shows the situation at higher doping, after E_A had decreased, and for the case $\Delta E > E_u$.

above a mobility edge in the Anderson model.[2] Hopping-type is via localized states. Such a process has been considered earlier by Davis and Mott,[25] who suggested that the conductivity in this case is of the form

$$\sigma \sim \exp[-(E_u + E_h)/kT] \ , \tag{4}$$

where E_h is the "hopping energy" and is related to the bandwidth (see below).

To simplify the subsequent discussion (Secs. II B and III), we shall here assume that E_u and E_h are in general not comparable, i.e., one or the other dominates. In this case the experimentally observed energy will be the larger of these two energies (E_u or E_h). Moreover, since the hopping energy E_h can be attributed to an activated *mobility*,[24] for $E_h >> E_u$, the overall conduction is of an activated-mobility-type. A further point to note is that the hopping expected for the case of Eq. (2) proceeds by transfer between D^0 and D^- states, i.e., by formation of states with two electrons. (We neglect "hole" transport via the D^+ states; the probability of the D^- hopping is expected to be larger than that of this D^+-type due to the larger radii of the D^- states). This differs from the ϵ_3-type hopping[2] which consists of transfer of an electron from a neutral (D^0) donor to an initially ionized (D^+) donor.

It remains to estimate, quantitatively, the various energies (E_A, E_u, ΔE, and E_h) introduced above. This can be done reasonably well at lower doping, i.e., in the "intermediate" doping range defined in Ref. 22. We first present primarily the results for this case, and subsequently will discuss the difficulties at higher doping. Following the approach of Ref. 22, we assume that, approximately, the energy is given by a screened hydrogenic Hamiltonian

$$H = (p^2/2m^*) - (e^2/\kappa r)e^{-qr} \ , \tag{5}$$

with the solution for E_A of[22,26]

$$E_A/E_0 \approx 1 - 1.81 a_0 q + 0.81 (a_0 q)^2 \ , \tag{6}$$

where q is the inverse screening length, κ is the dielectric constant, E_0 is the energy and a_0 (where $a_0 = \kappa \hbar^2/m^* e^2$) the Bohr radius for $q = 0$, and the other symbols are standard (and are defined in Ref. 22). The inverse screening length q, neglecting impurity band conduction, is given for n material by[27,28]

$$q^2 \approx \frac{4\pi e^2}{\kappa k}\left[\frac{n}{T} + \frac{N_A + n}{T}\left(1 - \frac{N_A + n}{N_D}\right)\right] \ , \tag{7}$$

where n is the carrier concentration, T is the temperature, k is the Boltzmann constant, and N_A and N_D are the acceptor and donor concentrations, respectively. For conduction in the impurity band, the carrier concentration does not depend on the occupation statistics, and consequently the screening is of the Debye–Huckel form,

$$q^2 = \frac{4\pi e^2}{\kappa k T} n_{\text{imp}} \approx \frac{4\pi e^2}{\kappa k T}(N_D - N_A)_{\text{imp}} \ . \tag{8}$$

Regarding E_u, the effect of screening on the electron affinity Q will be relatively minor,[29] so that E_u will decrease approximately as much as E_A [following Eq. (3)]. As to the value of Q, this has been determined as 1.6 meV by independent[30] measurements of the photoconductivity of the D^- level. The next quantity to be considered is the broadening in the density of states. That in E_A has been analyzed by Morgan[31] (see also Stern[27]), who obtains a Gaussian distribution[32] $\rho(E)$

$$\rho(E) = \frac{N_D}{(2\pi)^{1/2}\Delta E}\exp\left[-\left(\frac{E - E_A}{(2)^{1/2}\Delta E}\right)^2\right] \ , \tag{9}$$

where

$$\Delta E = \left(\frac{e^2}{\kappa}\right)\left(\frac{2\pi(N_D^+ + N_A^-)}{q}\right)^{1/2} \ . \tag{10}$$

Regarding the broadening for the D^- states, we are not aware of any work on this subject. However, this problem can be circumvented. Thus, for $E_u << \Delta E$, there will be an effective broadening of the joint density of states of the D^0 and D^- states. For this case, it thus seems reasonable to assume that Eq. (10) holds, at least approximately, for this joint density of states. And, as regards $E_u >> \Delta E$, in the situation where this case is of interest in the present problem—see discussion in Sec. II B—one does not require ΔE.

It remains to discuss the relation between E_h and ΔE. This problem has been analyzed (for both low and high doping) by Hill[33] for a density distribution which is uniform over a range of energies $2E_m$. We shall henceforth assume[32] that E_m is roughly equivalent to the ΔE of the Gaussian distribution, i.e., $E_m \approx \Delta E$. Hill's conclusion[33] is that the relation between E_h and ΔE depends on two parameters, z

and x, where

$$z = \frac{\beta}{2}\left[\frac{3}{4\pi N_D}\right]^{1/3} , \tag{11}$$

$$x = (\Delta E/kT) , \tag{12}$$

and β is the inverse Bohr radius. For small x ($x \lesssim 10$), there is[33] a relatively fast transition with decreasing z from the case of nearest-neighbor hopping with

$$E_h \approx \Delta E , \tag{13}$$

to that of constant range hopping with

$$E_h \approx kT . \tag{14}$$

(For larger x the transition region becomes broader and also includes the variable range hopping regime.) As an example,[33] for $x = 2$, the relation $E_h = \Delta E$ holds for $z \gtrsim 2$ whereas that of $E_h = kT$ applies for $z \lesssim 0.8$. Note that z decreases with increasing doping.

We next consider the situation for E_A, ΔE, and E_u at higher doping. For this range quantitative estimates for these energies are unfortunately not very reliable. There are several difficulties: (i) Use of the one-electron, one-impurity Hamiltonian of Eq. (5) can no longer be justified.[22] (ii) It is not certain which form of screening[5,10-11] (Lindhard, Hubbard–Sham, etc.) is best. (iii) Band-structure effects[5,7-11] should be included. (iv) Corrections for dynamic screening[34] also may be non-negligible. (v) Dielectric polarization effects become appreciable.[15-17] Probably the best available calculation for E_A in this high-doping range is that by Aldrich,[11] who uses Lindhard screening and who includes the band structure effects. [But note that points (i), (iv), and (v) are not included.] We shall use his calculation for evaluating E_A in this high-doping range. Regarding ΔE, we are not aware of any calculation comparable to that of Aldrich for E_A. We therefore analyze only the lowest-doped sample of Davis and Compton[21] which shows the ϵ_2 activation energy, and assume that the value of ΔE in this case does not deviate too far from that given by Eq. (10). As to E_u, since we analyze only the lowest-doped sample of Davis and Compton,[21] we again assume that the value as given in the low-doping case [Eq. (3), with $Q \approx 1.6$ meV] applies approximately. (Moreover we require the value of E_u primarily to show that $E_u << \Delta E$, see Sec. III. Thus, a rough estimate is satisfactory.)

B. Screening, and resultant separation into two regions

As just discussed, screening lowers the activation energy E_A, and also the excitation energy E_u. Moreover, screening will increase the effective impurity radius.[6] As a consequence of the decrease in E_u and/or the increase in radius, there will be a relatively rapid increase in interimpurity, i.e., D^-, conduction. To see this, note, first, that a lower E_u leads to increased excitation [Eq. (4)]. Second, it is known[35] that the hopping probability depends not only on energy, as given in Eq. (4), but also on the impurity separation (R) and the inverse Bohr radius (β) as

$$\sigma_h \sim \exp(-A\beta R) , \tag{15}$$

where $A \approx 2.8$. A necessary next step is to realize that mobile[36] carriers are required for good screening (with immobile carriers the potential in the Poisson equation cannot adjust—also see Ref. 22 for further discussion), so that higher conduction gives improved screening. This results in more conduction, etc., i.e., there is a *feedback* effect.

Because of the random distribution of the impurities, the feedback effect has several relevant consequences. For a better understanding, we must remember that the impurities are distributed at random spatially, and that regions with a (random) denser concentration will have better screening even in the absence of the above feedback effect, and thus contain the shallower levels. Introducing the feedback effect, this will now operate preferentially on clusters of closer, shallower impurities. This has as further consequence that because of the now even better screening, the energy of these impurities becomes lower than a hypothetical value which would exist in the absence of feedback. The deeper impurities, on the other hand, are not affected. With this selective lowering of the energy of the shallower impurities, the overall energy distribution is no longer uniform, but becomes split, as indicated in Fig. 3. Moreover, since the feedback and lowering of impurity energies takes place in clusters, there is also *spatial* separation between (interspersed) regions of shallower and deeper impurities. The impurity conduction thus takes place only along certain paths along shallower impurities, i.e., conducting channels are formed. As a further consequence, the number of

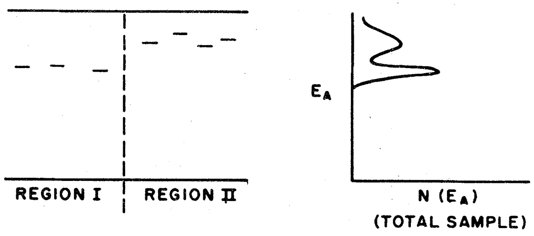

FIG. 3. Energy levels as postulated for the region with the deeper impurities (region I) and that with the shallower ones (region II) are shown on the left. The overall density of states for the sample (both regions) is shown on the right.

carriers available for this conduction is now equal to the number of impurities in the conducting region, and not to the overall impurity density in the sample. It must also be noted that in order to give dc conduction, the conducting channels must extend from one end of the sample to the other. Problems of this type are amenable to a percolation analysis, and it has been shown by Webman, Jortner, and Cohen[37] that for a random mixture this leads to the requirement that the conducting region occupy 15% of the total volume.

The observed activation energies of Eq. (1) must still be associated with the various energies of the model. For this, we define the region of deeper impurities as region I, and correspondingly region II for the shallower impurities.

We first consider region I, and note that the largest activation energy expected from the present model is E_A in this region, $\equiv E_A^{\mathrm{I}}$. We now identify this with the largest observed energy, ϵ_1. Moreover, in region I we expect $Q << E_A$ and therefore $E_u \sim E_A$ [Eq. (3)]. Under this condition, conduction in the conduction band will dominate over that in the impurity band, since the former has both a higher density of states and a higher mobility. Thus, neither E_u nor E_h will affect the conduction. Also, note that it is only in region I where, by hypothesis, there is the possibility of $E_u >> \Delta E$ (see below). Thus, since E_h (and thus ΔE) will not affect the conduction here, we do not require the value of ΔE when $E_u >> \Delta E$ (as asserted in Sec. II A). With no further[38] energies required for region I, the energy ϵ_2 is assigned to region II. As mentioned earlier in the present paper, we now assume that since the data of Davis and Compton[21] can be fitted[14] by treating ϵ_2 as an activated *mobility*, the dominant conduction in region II is hopping under the condition $E_u << E_h$ [see Eq. (4) and associated discussion]. We thus identify[39] ϵ_2 with E_h. (We omit superscripts on E_h, E_u, and ΔE, since these are of no interest in region I—see above.)

III. COMPARISON TO EXPERIMENT

As shown earlier[14] the present theory gives a good fit to the Hall data on Ge(Sb) reported by Davis and Compton,[21] using the activation energies obtained by these authors. It remains to be shown that these activation energies are consistent with the screening model.

The present analysis will be restricted to sample $L3$ of Davis and Compton.[21] Of those of their[21] samples which are of interest here (i.e., those showing the ϵ_2 activation energy over a reasonable temperature range), sample $L3$ is the lowest doped, and thus comes closest to satisfying the criteria of "intermediate" doping. For the calculation of E_A via Eq. (6), the parameter values $E_0 = 10.2$ meV (Ref. 40) with a corresponding $a_0 = 39$ Å [from the defining equation with $m^* = 0.22$ (Ref. 41)] are used, with q calculated via Eq. (7) or (8). The value of κ was taken as 16 (Ref. 42).

In region I, $N_D = 4.5 \times 10^{16}$, $N_A/N_D = 0.02$ (Refs. 14 and 21). Use of n as evaluated earlier[14] showed that "intermediate" doping[22] was satisfied for $T \gtrsim 40$ °K. Resultant evaluation at this temperature via Eqs. (6) and (7) ($n_{40}^{\mathrm{I}} = 1.8 \times 10^{16}$ cm^3, as obtained from the earlier[14] analysis of the Hall data[21]) gave $E_A^{\mathrm{I}} \approx 4.4$ meV. Similar evaluation at 100 °K ($n_{100}^{\mathrm{I}} = 3.5 \times 10^{16}$ cm^{-3}) gave $E_A^{\mathrm{I}} \approx 5.6$ meV. This range of E_A^{I} compares well indeed with the observed value of $\epsilon_1 \approx 5.5$ meV. Our identification of E_A^{I} with ϵ_1 is thus seen to be fully consistent.

For region II, we have identified E_h with ϵ_2 (Sec. II B), which requires that the inequality $E_u << E_h$ be satisfied (Sec. II A). We thus wish to show that $E_u << \epsilon_2$ or, from Eq. (3), that $(E_A^{\mathrm{II}} - Q) << \epsilon_2$. To obtain E_A^{II}, we first require q and consequently n_{imp} [Eq. (8)] in region II. The problem in this evaluation is that for inhomogeneous samples, Hall measurements give, at best, an approximate value of carrier concentration as averaged over the entire sample.[43] In view of the increase of impurity radius with screening,[6] the area of region II is not well known; consequently knowledge of the average concentration gives only limited information regarding the required concentration in region II. However, in order to show the adequacy of the model, it seems sufficient that a plausible assumption regarding this carrier density gives results consistent with our identification of the experimental values. Moreover, this approach is expected to be in line with the approximations already inherent in the present analysis of the *non*-"intermediate" situation. First, note that because of the impurity conduction $n_{\mathrm{imp}} \approx N_D - N_A$. Then, one must remember that the impurities in this region II are somewhat more closely spaced than those in the overall sample, since it is such a closer spacing which leads these impurities to have the shallower energies (Sec. II B). Assuming for specificity, perhaps a 50% increase in density then leads to $(N_D - N_A) \approx 6.6 \times 10^{16}$ cm^{-3}. For $T \approx 8$ °K (which is in the range showing ϵ_2), this leads to a value of $q = 3.27 \times 10^6$ cm^{-1} [Eq. (8)]. As discussed (Sec. II A), we use the results of Aldrich[11] to obtain E_A/E_0. From Fig. 1 of Aldrich,[11] $E_A/E_0 \approx 0.05$, giving $E_A^{\mathrm{II}} \approx 0.05$ meV. This value is too low, since by definition $E_A^{\mathrm{II}} > Q$, and since it is known from photoconductivity measurements[30] that $Q \approx 1.6$ meV. Nevertheless, this is not surprising in view of the approximations (see Sec. II A). Moreover, this result indicates that $E_A^{\mathrm{II}} - Q$ must be close to zero, so that the inequality $E_u << E_h$ is indeed satisfied. It now still remains to be shown that the value of ϵ_2 is reasonable. To check this, we now evaluate ΔE [Eq.

(10)] and then E_h. For the impurity band conduction

$$N_D^+ + N_A^- = N_D + N_A = 6.9 \times 10^{16}\ \text{cm}^{-3} ,$$

and consequently $\Delta E \approx 3.3$ meV. Again considering $T \approx 8$ °K, this leads to $x \approx 4.8$ [Eq. (12)]. For evaluation of z [Eq. (11)], the inverse Bohr radius β is also required. Assuming, for sample $L3$, an increase[44] of ~50% in the radius over that at infinite dilution, i.e., $\beta \approx (\frac{1}{60}\ \text{Å})$, gives $z \approx 1.3$. From Hill's[33] Fig. 4, for these values $E_h \approx \frac{1}{2}\Delta E$, leading[45] to $E_h \sim 1.6$ meV. Since $\epsilon_2 \approx 2$ meV (Refs. 14 and 21), the identification of ϵ_2 with E_h is consistent. It can also be noted that with a 50% increase in radius in region II, and since the impurities in this region amount to about 5% of the total (Ref. 14), the volume fraction of region II is about 17%. This satisfies the percolation criterion of 15% for conductivity (Sec. II B).

Before concluding this section, we also still wish to mention the possible application of the screening model to the phenomenon of "persistent" or "storage" photoconduction. It has been observed[46] on many semiconductors that at low-temperatures photoconductivity, once initiated, often does not decay. This conduction can then be quenched only on heating. Moreover, Kulp *et al.*[47] have shown that the transport properties of those of their CdS crystals which showed this effect "can be explained on the basis of impurity conduction". We thus suggest that this type of photoconduction results because of the extra carriers, and thus better screening, introduced by the illumination. The resultant feedback then "switches" the original nonscreened insulating state to a screened, conducting system. However, details on this respect have not yet been worked on.

IV. DISCUSSION

A disconcerting feature in understanding the "ϵ_2" metal-insulator transition in doped semiconductors is that there are several theories which can explain the gross features of such a transition. Past attempts to decide between such theories by more detailed analyses have not met with much success. As two recent examples: Mott, after generally favoring the Hubbard view, has recently[48] concluded that the detailed behavior of the Hall coefficient as a function of concentration is better understood by the Anderson model. Arguments against this latter interpretation have, however, already been advanced by Deshmuk.[49] Similarly, on the basis of localization calculations it has been concluded by Antoniou and Economou[50] that, "at least for uncompensated specimens, the electron-electron correlation and not the randomness is responsible...". Whether these arguments are conclusive has also already been questioned, by Weaire and Srivastava.[51]

The present work is a continuation of such attempts to differentiate among the various theories. As we have shown previously,[14] the temperature dependence of the Hall coefficient, in the pretransition range, is incompatible with excitation either to an upper Hubbard band or to an Anderson mobility edge. A screening model, coupled with a separation into two regions, however *can* explain the data. In the present paper, we have presented this model in more detail. We have also estimated the appropriate activation energies for the sample of Davis and Compton which is most amenable to theoretical predictions. The calculations, without adjustable parameters, gave $\epsilon_1 = 4.4$ to 5.6 meV in the temperature range of interest, as against an observed value of 5.5 meV. For ϵ_2 some estimates are required (for instance, the increase in impurity density in region II); however, with reasonable values, the prediction is 1.6 meV, as against an observed value of 2 meV. We feel these results strongly support the validity of the model in the pretransition range. As regards the MI transition itself, there is no question that screening can explain such a transition, since sufficient screening always leads to zero activation energy.[4–11] However, as can be seen from Fig. 1, polarization and disorder effects become increasingly important close to the transition. We thus still regard this aspect as an unanswered question.

ACKNOWLEDGMENTS

The author is indebted to Dr. E. Arnold, Dr. R. Bhargava, Dr. J. Krieger, and Dr. F. Stern for several helpful discussions, as well as to Dr. R. Bhargava and Dr. G. Blom for a critical reading of the manuscript.

---

[1] C. S. Hung and J. R. Gliessman, Phys. Rev. 79, 726 (1950).

[2] N. F. Mott, *Metal-Insulator Transitions* (Taylor and Francis, London, 1974); N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials* (Oxford University, London, 1971); N. F. Mott and Z. Zinamon, Rep. Prog. Phys. 33, 881 (1970).

[3] H. Fritzsche, J. Phys. Chem. Solids 6, 69 (1958).

[4] N. F. Mott, Proc. Phys. Soc. London 62, 416 (1949).

[5] J. B. Krieger and M. Nightingale, Phys. Rev. B 4, 1266 (1971).

[6] C. S. Lam and Y. P. Varshni, Phys. Rev. A 4, 1875 (1971).

[7] F. Martino, G. Lindell, and K. F. Berggren, Phys. Rev. B 8, 6030 (1973).

[8] O. P. Sinha and O. P. Puri, Phys. Rev. B 12, 1395 (1975).

[9] P. Csavinszky, Phys. Rev. B 14, 1649 (1976).

[10] R. L. Greene, C. Aldrich, and K. K. Bajaj, Phys. Rev. B 15, 2217 (1977).

[11] C. Aldrich, Phys. Rev. B 16, 2723 (1977).

[12] J. Hubbard, Proc. R. Soc. London Sect. A 281, 401 (1964).

[13] B. Johansson, J. Phys. C 6, L71 (1973).

[14] G. F. Neumark, in *Physics of Semiconductors,* edited by F. G. Fumi (Marves, Rome, 1976), p. 301.

[15] G. W. Castellan and F. Seitz, *Semiconducting Materials* (Butterworths, London, 1951), p. 8.

[16] T. G. Castner, N. K. Lee, G. S. Cieloszyk, and G. L. Salinger, Phys. Rev. Lett. 34, 1627 (1975).

[17] P. Leroux Hugon and A. Ghazali, Phys. Rev. B 14, 602 (1976).

[18] P. W. Anderson, Phys. Rev. 109, 1492 (1958).

[19] T. Matsubara and Y. Toyozawa, Prog. Theor. Phys. 26, 739 (1961).

[20] K. A. Chao, F. A. Oliveira, and N. Majlis, Solid State Commun. 21, 845 (1977).

[21] E. A. Davis and W. D. Compton, Phys. Rev. 140, A2183 (1965).

[22] G. F. Neumark, Phys. Rev. B 5, 408 (1972).

[23] See, for example, H. Nishimura, Phys. Rev. 138, A815 (1965).

[24] See, for example, H. Scher and M. Lax, Phys. Rev. B 7, 4502 (1973).

[25] E. A. Davis and N. F. Mott, Philos. Mag. 22, 903 (1970).

[26] This approach neglects the details of the band structure. However, multivalley aspects do not appear to be too important in the low-doping range—see Ref. 9.

[27] F. Stern, Phys. Rev. B 9, 4597 (1974).

[28] The screening length as given in Ref. 22 should have included terms due to preferential neutralization of deeper donors, as introduced for ionized impurity scattering by Brooks and Herring [H. Brooks, in *Advances in Electronics and Electron Physics,* edited by L. Marton (Academic, New York, 1955), Vol. VII, p. 85]. These terms are mostly more important than the ion-mobility terms considered in Ref. 22, and neglect of these ion-mobility terms then leads to the result of Eq. (7). It must also be noted that Eq. (7) breaks down at very low temperatures (this aspect is discussed in Ref. 27), but is expected to be approximately valid for the present applications.

[29] The change of Q with screening has been examined more carefully by J. B. Krieger (private communication). He also concludes that this energy is not greatly influenced by screening—appreciably less than E_A.

[30] M. Taniguchi, M. Hirano, and S. Narita, Phys. Rev. Lett. 35, 1095 (1975).

[31] T. N. Morgan, Phys. Rev. 139, A343 (1965).

[32] For purposes of the present paper, the exact form of the distribution is not of importance, the primary requirement being that such a distribution results in a hopping process and thus to an activated mobility. However, the Gaussian form is convenient for estimating ΔE from the doping density. Also, it is unlikely that the same ΔE applies to both the E_A and the E_u states. In practice, we here consider only the case $E_u << \Delta E$, and assume that for this case the overall broadening is approximately ΔE.

[33] R. M. Hill, Phys. Status Solidi A 34, 601 (1976). Also, note that since in Hill's treatment E_h (Hill's ΔE) is a function of temperature, an improved definition of E_h over that of Eq. (4) is $E_h = -d \ln\sigma / d(1/kT)$, and is indeed the definition used by Hill.

[34] G. Mahler and J. L. Birman, Phys. Rev. B 16, 1552 (1977); and G. Mahler (private communication).

[35] See, for example, C. H. Seager and G. E. Pike, Phys. Rev. B 10, 1435 (1974).

[36] The ratio of the hopping mobility in region II to that in region I (i.e., μ_{II}/μ_I) can be obtained from Eqs. (4) and (15). With the assumptions $E_u^{I} >> E_h^{I}$ and of $E_u^{II} << E_h^{II}$, these give

$$\frac{\mu_{II}}{\mu_I} \approx \exp - \left[\frac{E_h^{II} - E_u^{I}}{kT} + 2.8[(\beta R)^{II} - (\beta R)^{I}] \right] .$$

For region I, the appropriate values are $E_u \approx 4$ meV, $(1/\beta) \approx a_0 = 39$ Å, $R = 174$ Å. For region II, $E_h = 2$ meV, and assuming a 50% increase in impurity density and a 50% increase in radius (i.e., reasonable values) gives $R = 153$ Å and $(1/\beta) = 26$ Å, respectively. At $T = 8$ °K, the result is $\mu_{II}/\mu_I \approx 3 \times 10^3$, which shows that μ_I is very appreciably smaller than μ_{II}.

[37] I. Webman, J. Jortner, and M. H. Cohen, Phys. Rev. B 11, 2885 (1975).

[38] We suggest, as a conjecture, that the hopping responsible for ϵ_3 may be active mainly in region I. In region II, one would expect an activation energy for ϵ_3 comparable to that for ϵ_2 (since both, essentially, depend on ΔE).

[39] It is to be emphasized that the identification of ϵ_2 with E_h and the neglect of E_u applies to samples where the data can be fitted with an activated mobility. This does not exclude the possibility that the activation to the D^- level, E_u, may provide the limiting factor in other samples, nor, even, that it may cause some perturbation in the present samples. Another energy of region II which is neglected here is the impurity to band one, E_A^{II}. Although in principle E_A^{II} should contribute a band conduction, in practice it was not required for fitting the data (Ref. 14); since its inclusion would provide one more parameter available for fitting, such inclusion could make the fit only easier, without changing the basic principle of the present approach.

[40] J. H. Reuszer and P. Fisher, Phys. Rev. 135, A1125 (1964).

[41] See, for example, J. S. Blakemore, *Semiconductor Statistics* (Pergamon, New York, 1962).

[42] The value of $\kappa = 16$ which we have used is the one for low doping, which however may increase with doping. For a discussion of this point (and data for Si) see T. G. Castner and N. K. Lee, Solid State Commun. 19, 323 (1976) and references given there. Early data on Ge(Sb) by F. A. D'Altroy and H. Y. Fan, Phys. Rev. 103, 1671 (1956) had indicated that κ for sample $L3$ would not be appreciably increased. A more recent estimate by Castner (private communication, based on results to be published) is that $\kappa \approx 25$ for this sample. This would increase our estimates for E_A and decrease that of ΔE somewhat, but would not affect the main arguments of this paper.

[43] See, for example, J. Volger, Phys. Rev. 79, 1023 (1950).

[44] The value of radius obtained from percolation analyses of the MI transition [see Ref. 35 and D. F. Holcomb and J. J. Rehr, Jr., Phys. Rev. 183, 773 (1969)] is close to twice that of the Bohr radius at infinite dilution. However, most of the samples determining the percolation value are more heavily doped than $L3$. Consequently, an increase

of 50% appears reasonable for this sample.

[45] It is to be noted that Hill's analysis does predict some variation of E_h with T for the present values of x and z, but a relatively minor one.

[46] See, for example, M. K. Sheinkman and A. Ya. Shik, Sov. Phys. Semiconduct. 10, 128 (1976).

[47] B. A. Kulp, K. A. Gale, and R. G. Schulze, Phys. Rev. 140, A252 (1965).

[48] N. F. Mott, J. Phys. (Paris) 37, C4-301 (1976); Philos. Mag. 35, 111 (1977).

[49] V. G. I. Deshmukh, Philos. Mag. B 37, 649 (1978).

[50] P. D. Antoniou and E. N. Economou, Phys. Rev. B 16, 3768 (1977).

[51] D. Weaire and V. Srivastava, Solid State Commun. 23, 863 (1977).