Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Infrared Physics & Technology xxx (2006) xxx–xxx
www.elsevier.com/locate/infrared

Energy gap–refractive index relations in semiconductors – An overview

N.M. Ravindra *, Preethi Ganapathy, Jinsoo Choi

Department of Physics, New Jersey Institute of Technology, Newark, NJ 07102, USA

Received 23 September 2005

Abstract

An overview of the understanding of correlations between energy gap and refractive index of semiconductors is presented here. The atomic approach of Moss, the nearly free electron model of Penn, the oscillator concept of Wemple and the optical polarizability approach of Finkenrath are considered in this study. The Ravindra relation is discussed in the context of alternate approaches that have been presented in the literature. Case studies of applications of these relations to infrared materials and wide band gap semiconductors are presented.

1. Introduction

The refractive index and energy gap of semiconductors represent two fundamental physical aspects that characterize their optical and electronic properties. The applications of semiconductors as electronic, optical and optoelectronic devices are very much determined by the nature and magnitude of these two elementary material properties. These properties also aid in the performance assessment of band gap engineered structures for continuous and optimal absorption of broad band spectral sources. In addition, devices such as photonic crystals, wave guides, solar cells and detectors require a pre-knowledge of the refractive index and energy gap. Application specific coating technologies (ASPECT™) [1] including antireflection coatings and optical filters [2] rely on the spectral properties of materials.

The energy gap determines the threshold for absorption of photons in semiconductors. The refractive index in the semiconductor is a measure of its transparency to incident spectral radiation. A correlation between these two fundamental properties has significant bearing on the band structure of semiconductors. In 1950, Moss [3] proposed a basic relationship between these two properties using the general theory of photoconductivity which was based on the photo effect studies of Mott and Gurney [4], Smekal [5], Zwicky [6], Gudden and Pohl [7] and Pearson and Bardeen [8]. According to this theory, the absorption of an optical quantum will raise an electron in alkali halides to an excited state rather than freeing it from the center. Thermal energy then moves this electron to the conduction band from the lattice. Such a photo effect takes place in imperfections at certain lattice points, and thus, the electron behaves similar to an electron in an isolated atom with a dielectric constant of the bulk material. As a result of this effective dielectric constant, $\varepsilon_{\mathrm{Eff}}$, the energy levels of the electron are scaled down by a factor of $1/\varepsilon_{\mathrm{Eff}}^2$ which approximately corresponds to the square of the refractive index, n. This factor, thus, should be proportional to the energy required to raise an electron in the lattice to an excited state as given by the Bohr formula for the ionization energy, E, of the hydrogen atom, $E = 2\pi^2 m^* e^4/\varepsilon^2 h^2$, where, m^* is the electron effective mass, e is the electronic charge, ε is the relative permittivity and h is the Plank constant. This minimum energy determines the threshold wavelength, λ_e, which then varies as the fourth power of the refractive index. Experimental data on different photoconductive compounds show that the values of n^4/λ_e were close to 77 throughout a range of refractive indices. The

* Corresponding author.
E-mail address: ravindra@njit.edu (N.M. Ravindra).

1350-4495/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.infrared.2006.04.001

similarities in the quotient show that the photoelectrons stem from the same type of lattice imperfections, or alternatively, the binding energies to the different types of hydrogen-like centers are similar. Thus the Moss relation was formulated as [3]:

$$\frac{n^4}{\lambda_e} = 77/\mu\text{m} \tag{1}$$

where n is the refractive index and λ_e is the wavelength corresponding to the absorption edge. In terms of energy gap, this is [9]:

$$n^4 E_g = 95\ \text{eV} \tag{2}$$

According to this relation, the refractive index of a semiconductor can be determined with a known energy gap, E_g. This relation, again, was based on the general assumption that all energy levels in a solid are scaled down by a factor of $1/\varepsilon_{\text{eff}}^2$.

2. Penn model

In 1962, Penn [10] proposed a simple model for an isotropic semiconductor with electrons in a sphere of momentum space and are characterized by an isotropic energy gap. In his investigation of a nearly free electron model for a semiconductor, Penn showed that two common assumptions were flawed in Callaway's approximation of the static dielectric constant [11]. Previous semiconductor models did not allow for the formation of standing waves in the Brillouin zone. The Umklapp process was neglected. The energy, E, and wave function, ψ, with respect to the state k for this model are given by,

$$E_{\mathbf{k}}^{\pm} = \tfrac{1}{2}\{E_{\mathbf{k}}^0 + E_{\mathbf{k}'}^0 \pm [(E_{\mathbf{k}}^0 - E_{\mathbf{k}'}^0)^2 + E_g^2]^{1/2}\}$$
$$\psi_{\mathbf{k}} = (e^{i\mathbf{k}\cdot r} + \alpha_{\mathbf{k}}^{\pm} e^{i\mathbf{k}'\cdot r})/[1 + (\alpha_{\mathbf{k}}^{\pm})^2]^{1/2}$$

where

$$\alpha_{\mathbf{k}}^{\pm} \equiv \tfrac{1}{2}E_g/(E_{\mathbf{k}}^{\pm} - E_{\mathbf{k}'}^0)$$
$$E_{\mathbf{k}}^0 \equiv (\hbar^2/2m)k^2, \quad \mathbf{k}' \equiv \mathbf{k} - 2K_F\hat{k} \tag{3}$$

Here, E is the electron energy, $\mathbf{k}$ is the wave vector, $\mathbf{K}_F$ is the Fermi wave vector and α represents an averaged Jones boundary [12]. The electron energy E as a function of wave vector k for an isotropic three-dimensional free electron model is plotted in Fig. 1. This model of placing the energy gap above the Fermi surface was first suggested by Callaway [11] in his investigation of the correlation energy of electrons.

Fig. 2 shows the density of states vs. the energy and describes the treatment of the Penn model. The dashed line represents Callaway's model and the solid line represents the Penn model. According to the Penn model, the states from the energy gap are removed and piled up at the top of the valence band and the bottom of the conduction band [10]. The resulting graph is thus asymptotic.

Based on this model, Penn describes the dielectric function as:

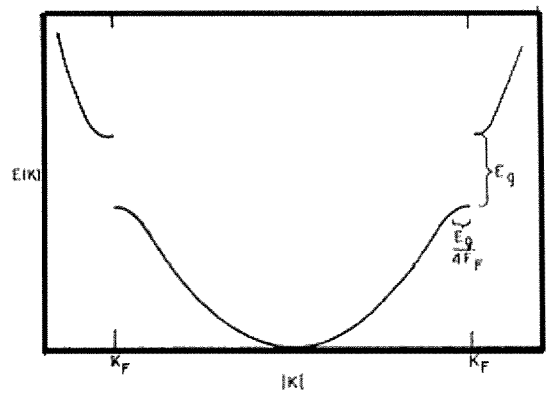

Fig. 1. Electron energy as a function of k for isotropic three-dimensional nearly free electron model [10].

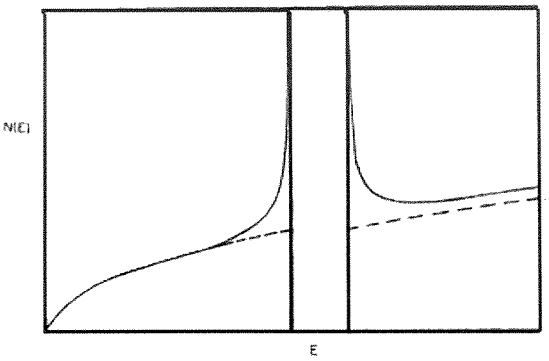

Fig. 2. Density of states vs. energy. The solid curve represents the three-dimensional nearly free electron model while the dashed curve represents Callaway's model [10].

$$\epsilon(\mathbf{q}) = 1 + \frac{8\pi e^2}{q^2}\left[\sum_{\mathbf{k}} N_{\mathbf{k}}(1 - N_{\mathbf{k}+\mathbf{q}})\frac{|\langle \mathbf{k}|e^{-i\mathbf{q}\cdot\mathbf{r}}|\mathbf{k}+\mathbf{q}\rangle|^2}{E_{\mathbf{k}+\mathbf{q}}^{+} - E_{\mathbf{k}}^{-}} + \sum_{\mathbf{k}} N_k(1 - N_{k'})\frac{|\langle \mathbf{k}|e^{-i\mathbf{q}\cdot\mathbf{r}}|\mathbf{k}'\rangle|^2}{E_{\mathbf{k}'}^{+} - E_{\mathbf{k}}^{-}}\right] \tag{4}$$

In this equation, $\mathbf{k}' = \mathbf{k} + \mathbf{q} - (2K_F)(\mathbf{k}+\mathbf{q})/|\mathbf{k}+\mathbf{q}|$ where $\mathbf{K}$ is the reciprocal lattice vector and $\mathbf{q}$ is a wave number. N is the occupation number for the states $\mathbf{k}$ and $\mathbf{k}+\mathbf{q}$. e is the electronic charge. If the Brillouin zone is divided into sections, then the dielectric function can be written as:

$$\epsilon(\mathbf{q}) = 1 + \frac{8\pi e^2}{q^2}\sum_{\mathbf{k}i}\delta_{\mathbf{k}}^i\frac{|M_i|^2}{E_i} \tag{5}$$

where $\delta_k^i = 1$ when k is in the i region and 0 otherwise. M_i is the matrix element. For the specific case when $q = 0$,

$$\varepsilon(0) \approx 1 + (\hbar\omega_p/E_0)^2 + [1 - (E_g/E_F/4) + ((E_g/E_F)^2 + 48)] \tag{6}$$

where ω_p is the plasma frequency and E_F is the Fermi energy. The expression in the brackets is nearly equal to 1 for materials with band gaps in the commonly occurring range where $(E_g/E_F) = 0.3$ [12]. Since the most significant varia-

tion occurs in the expression before the square brackets, (5) can be written as,

$$\varepsilon(0) \approx 1 + (\hbar\omega_p/E_0)^2 \cdot S_0 \tag{7}$$

where S_0 represents the quantity in brackets in Eq. (6). Penn approximates S_0 to be nearly 1, thereby neglecting the smaller quantities of E_g/E_F. The dielectric function is thus:

$$\varepsilon(0) \approx 1 + (\hbar\omega_p/E_0)^2 \tag{8}$$

Grimes and Cowley [12] found a more accurate value of S_0. In their investigation, Grimes and Cowley found that the value of S_0 is only weakly dependent on the band gap and that a value of 0.62 is a fairly good representation of S_0. Thus, with this slightly more accurate value for S_0, the energy gap can be determined by using appropriate values of the dielectric constant.

In the Penn model, the dielectric constant was found to have reasonable values for small values of wave number ($\mathbf{q}$). This model is quite simplistic in design, though it accounts for the formation of standing waves near the Brillouin zone. It does not consider the degeneracy that may occur in the Brillouin zone. In spite of the simplicity of the Penn model, the mere isotropy of the system allows the model to be applied to a liquid or amorphous semiconductor.

3. Ravindra relation

The Ravindra relation, which was initially proposed empirically [13] was shown to be an approximation [14] of the Penn model [10]. One feature common to all the semiconductor band structures is that the valence and conduction bands are more or less parallel to each other at least along the symmetry directions. This formed the basis for Gupta and Ravindra [14] to define:

$$E_p = E_g + K \tag{9}$$

Using Eq. (9) in the Penn model in Eq. (8) and considering the situation for which $E_g/K < 1$, Gupta and Ravindra arrived at an expression for the refractive index:

$$n = K_1 - K_2E_g + K_3E_g^2 - K_4E_g^3 \tag{10}$$

where the values of K are calculated by Gupta and Ravindra [14] and the higher values are neglected since they are too small to be considered. Eq. (10) then reduces to a form:

$$n = 4.16 - 1.12E_g + 0.31E_g^2 - 0.08E_g^3 \tag{11}$$

The oscillator model from Wemple [15] defines the dielectric constant as: $\varepsilon = 1 + E_d/E_0$ where ε is the real part of the complex dielectric constant, E_d is the oscillator strength and E_0 is the average energy gap which is approximately equal to the Penn energy gap, E_p.

Performing a similar treatment with the dielectric function in the optical region, they [14] show a similar equation for the refractive index with the constants evaluated,

$$n = 4.16 - 0.85E_g \tag{12}$$

Other forms of the Ravindra relation include a variation of the relation with respect to temperature and pressure [16]:

$$\frac{dn}{dT} = -0.62\frac{dE_g}{dT} \tag{13}$$

$$\frac{dn}{dp} = -0.62\frac{dE_g}{dp} \tag{14}$$

3.1. Moss relation and its implications

It is perhaps noteworthy to mention that other optical properties of semiconductors are merely integral relations to the energy gap and refractive index. Moss [9] pointed out that the absorption edge can be calculated using the refractive index. Moss [9] showed that the refractive index and absorption edge are related by:

$$n - 1 = \frac{1}{2\pi^2} \times \int \left(K \frac{d\lambda}{1 - \frac{\lambda^2}{\lambda_e^2}} \right) \tag{15}$$

which for long wavelengths at zero frequency becomes,

$$n - 1 = \frac{1}{2\pi^2} \times \int K\, d\lambda \tag{16}$$

Here, K is the absorption coefficient and λ_e is the wavelength corresponding to the absorption edge. This function is integrated over the wavelength rather than energy. This expression shows that the long wavelength refractive index is determined simply by the total area under the curve of absorption coefficient vs. wavelength – it is independent of the absorption spectrum [18]. As a result, if a certain level of absorption persists over a given energy interval – as it occurs where absorption is due to transitions between two allowed energy bands – then, the smaller the width of the forbidden zone, the greater will be the spread of absorption in wavelength, resulting in a greater value of the integral and larger n [9].

Moss [9] provided an illustration of this relation in his paper. This analysis shows that much of the refractive index originates from narrow bands of intense absorptions. The absorption occurs at wavelengths below 1 μm which corresponds to the very short wavelength compared to the absorption edge.

With these observations, Moss concluded that the narrow absorption band of semiconductors in the UV region is the factor which determines the refractive index. Two conclusions followed from this assumption: (1) the relation between the energy gap and UV oscillator frequency are factors in determining the relationship between the refractive index and the long wavelength absorption edge; (2) any modifications of the absorption edge that do not vary at the same time are not expected to produce a change in the refractive index.

Moss compared his relation with the Ravindra relation [9]. The Ravindra relation predicts no results of indices

beyond 4.1. However, very reliable infrared materials, such as PbSe and PbSnTe, do exist at refractive indices of 4.7 and 7.0 respectively. The Moss formula, however, closely estimates these values with the formula: $n^4E_g = 108$ eV where the constant has been revised. This formula yields refractive indices of 4.6 for PbSe and 6.03 for PbSnTe. For low refractive indices, the Ravindra relation predicts impractical results. It predicts that the refractive index will become zero when the energy gap is 6.6 eV. The Moss formula on the other hand requires that the energy gap be infinity for the same result. For the index to be unity, the Ravindra relation predicts the energy gap to be 5 eV. However, there exist materials which exceed this value. The Moss relation requires that the energy gap should be greater than 100 eV to obtain a refractive index of 1. This is much larger than the refractive index of any material.

In a theoretical derivation of optical polarizability, Finkenrath [17] described a treatment in the transitions from the valence band to the conduction band. These transitions are represented by differential oscillator densities that occur at distributed resonant frequencies. The combined polarizabilities are summed and replaced by an integral over the relevant frequencies. These differential oscillators are identified through the wave number and occupation probability based on spherical energy surfaces. The result of this study showed that $(\varepsilon_{\text{eff}} - 1)^2E_g$ is a complex expression that is basically constant. From this expression, one can see that neglecting unity compared to the dielectric constant gives the Moss relation. According to Finkenrath's estimate, the average value of the constant is nearly 95 eV. These observations are made for direct interband transitions. It must be understood that, if the transitions are not direct, then deviations could arise from ignoring the transitions from other bands and the varying densities of states and the matrix elements.

Moss [9] thus showed the interrelations between the refractive index and the absorption. He showed that the important factor that determines the index is the frequency of the UV absorption peak. Reasonable values for the refractive index can be found using the model of a single classical oscillator at this frequency. The relation between the refractive index and the absorption edge band energy arises from the comparative nature of the band edge and the resonance energy. Assuming a constant energy difference, a relation similar to the linear Ravindra relation can be obtained. However, the Ravindra relation is shown to have a few shortcomings at high and low index values. The Moss relation out performs the Ravindra relation in these areas and is further supported by formal theory from Finkenrath. From the different semiconductors known, one can understand that a relationship between refractive index and energy gap, or for that matter, other optical constants that are integrally related, can only be general and approximate at best. From Moss [9], it is seen that where the Ravindra relation fails, the Moss relation holds approximately true.

4. Alternate approaches and interpretations

Further relations were developed as a modification or addition to the Moss and Ravindra relations. While the Moss formula is limited by the structure of the material, the Ravindra relation is restricted by the refractive index. From the Ravindra relation, the refractive index cannot be greater than a value of 4.1, which corresponds to an energy gap of 6.587 eV. In an effort to broaden the application of these two concepts, several authors have presented variations of the Moss and Ravindra relations.

In 1992, a relation similar to the Moss relation was proposed by Reddy and Anjaneyulu [19]. According to their formula, the relation is:

$$E_g e^n = 36.3 \tag{17}$$

This relation holds true for energy gaps greater than 0 eV.

Reddy and Ahammed [20] proposed an empirical relation that was a modification of the Moss formula [18]

$$n^4(E_g - 0.365) = 154 \tag{18}$$

This relation was proposed to overcome some of the drawbacks of the Moss relation. However, it is not valid when the energy gap is less than 0.36 eV nor does it hold for infrared materials such as lead salts and InSb.

Further derivations based on the Moss and Ravindra relations include approaches by Herve and Vandamme [21]. They separated the semiconductors into a covalently bonded group and ionically bonded group. For a covalently bonded crystal, the dielectric function is:

$$\varepsilon(\omega) = 1 + 4\pi Nq^2/[m(\omega_0^2 - \omega^2)] \tag{19}$$

where N is the density of valence electrons, m the rest mass and ω_0 the UV resonance frequency. Eq. (19) describes the dielectric constant for elements like Si and Ge. For the case of binary compounds such as NaCl or GaAs, the dielectric function needs to account for the ionic bonding. Thus the dielectric function in an ionic diatomic lattice is,

$$\varepsilon'(\omega) = \varepsilon'(\omega_{\text{op}}) + \Omega_p^2/(\omega_T^2 - \omega^2) \tag{20}$$

where Ω_p is similar to the plasma frequency for ions and ω_T is the infrared resonance frequency. To account for materials in the optical range, $\omega_T < \omega_{\text{op}} < \omega_0$, Herve and Vandamme proposed,

$$n = [\varepsilon'(\omega_{\text{op}})]^{1/2} = n = \sqrt{1 + \left(\frac{A}{E_g + B}\right)^2} \tag{21}$$

where A is the hydrogen ionization energy $\approx$13.6 eV and B is 3.47 eV. Herve and Vandamme based this equation on oscillatory theory, assuming the UV resonance energy has a constant difference with band gap energy,

$$h\omega_0 = E_g + B \tag{22}$$

Herve and Vandamme claim that their model provides the lowest deviation for III–V, I–VII and chalcopyrites. This

model is accurate for most materials used in optoelectronic device structures and high band gap materials. Yet it does not explain the behavior of the IV–VI group.

In the Moss, Ravindra and Herve and Vandamme relations, the variation of the energy gap with temperature leads to three relations. By differentiating their relation with respect to temperature, we arrive at:

$$\frac{dn}{dT} = -\frac{13.6^2}{n(E_g+B)^3} \times \left(\frac{dE_g}{dT} + B_1\right) \tag{23}$$

where $B_1 = \frac{dB}{dT}$ Allowing, $E_g + B = \sqrt{\frac{13.6^2}{n^2} - 1}$,

$$\frac{1}{n}\frac{dn}{dT} = -\frac{(n^2-1)^{3/2}}{13.6n^2(E_g+B)^3} \times \left(\frac{dE_g}{dT} + B_1\right) \tag{24}$$

Similarly, the Moss and Ravindra relations were differentiated to arrive at temperature coefficient for the energy gap.

A few values of (dE_g/dT) and (dn/dT) calculated using this relation for some semiconductors are given in Table 1.

In conducting this study, Herve and Vandamme [22] found two particular temperature dependences. Generally, the energy gap and temperature are inversely proportional for most semiconductors, that is, as T increases, the energy gap decreases, and, as a result, the refractive index increases. This occurs in diamond and similar structures. On the other hand, infrared detector materials including the lead sulphide, lead selenide and lead telluride exhibit a positive temperature coefficient of energy gap and negative refractive index temperature dependence.

Herve and Vandamme [22] found that the Moss relation showed the strongest deviation at lower energy gaps (<1.43 eV), as shown in Fig. 3. Above this value, the Moss relation provides results closest to the experimental values.

Calculations made with the Ravindra relation show more accurate estimations for energy band gap values less than 1.43 eV. These results deviate at energy gaps greater than 1.43 eV. Herve and Vandamme found that their model presents the best results below the specified energy gap and is quite close to the experimental results above 1.43 eV.

Table 1
Energy gap, its temperature coefficient and refractive index temperature dependence for some semiconductors as calculated by the Herve Vandamme equation [22]

Material	E_g (eV)	dE_g/dT (eV K^{-1})	$(dn/dT)/n$ (K^{-1})
InSb	0.18	-2.8×10^{-4}	6.9×10^{-5}
PbSe	0.278	$+5.1\times10^{-4}$	-2.1×10^{-4}
Ge	0.67	-3.7×10^{-4}	6.9×10^{-5}
GaSb	0.75	-3.7×10^{-4}	8.2×10^{-5}
Si	1.1	-2.8×10^{-4}	4.0×10^{-5}
InP	1.35	2.9×10^{-4}	2.7×10^{-5}
GaAs	1.43	-3.9×10^{-4}	4.5×10^{-5}
AlAs	2.15	-4.0×10^{-4}	4.6×10^{-5}
AlP	2.41	-3.7×10^{-4}	3.6×10^{-5}
SiC	2.86	3.3×10^{-4}	2.9×10^{-5}
GaN	3.5	-4.8×10^{-4}	2.6×10^{-5}
C	5.48	-5.0×10^{-5}	4.0×10^{-6}

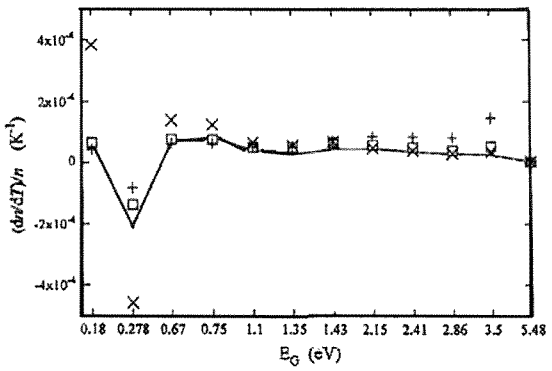

Fig. 3. Temperature coefficient of refractive index vs. energy band gap. The solid line represents the experimental results. The cross (×) represent the differentiated Moss formula, the (+) represent the differentiated Ravindra formula and the open squares (□) represent the Herve Vandamme model [22].

5. Case studies of applications to condensed matter physics

5.1. IR detector materials

One application of the energy gap–refractive index relation helps to determine the strontium composition in PbSe and lead-alkaline-earth-chalcogenide materials ($Pb_{1-x}Sr_xSe$ [23]). The compounds are notably used in mid infrared (mid-IR) lasers and mid-IR/ultraviolet (UV) detectors. The energy gap and refractive indices of these materials depend on the strontium composition. Majumdar et al. [23] sought to determine the electronic and optical properties of the ternary compound for Sr compositions. PbSe has a direct band gap at 0.3 eV, at the L point of the Brillouin zone [23]. Contrastingly, SrSe has a wide indirect band gap between the X–Γ bands of the Brillouin zone. As the amount of Sr in $Pb_{1-x}Sr_xSe$ increases, the band gap changes from narrow and direct to wide and indirect. Majumdar et al. have determined the composition at which this change takes place.

In their experiment [23], the transmission spectra measured from the epitaxial layers of $Pb_{1-x}Sr_xSe$ were used to calculate the refractive index and the absorption coefficient. The refractive index was approximated from odd interference peaks in the long wavelength region, which is below the energy gap. This approximated value was then fitted with the experimental curve from which the true refractive index and absorption coefficient were determined. The refractive index of $Pb_{1-x}Sr_xSe$ was calculated and was found to be in the range of 4.8–2.04 for room temperature and for a strontium composition of 0–1 at a wavelength of 4 μm. Similarly, SrSe has a refractive index of 2.04 at room temperature and at a wavelength of 4 μm. This is shown in Fig. 4.

From Fig. 4, the refractive index of PbSe shows a negative temperature coefficient while SrSe shows a positive coefficient. When the composition is nearly $Pb_{0.82}Sr_{0.18}Se$, the refractive index at the specified wavelength is indepen-

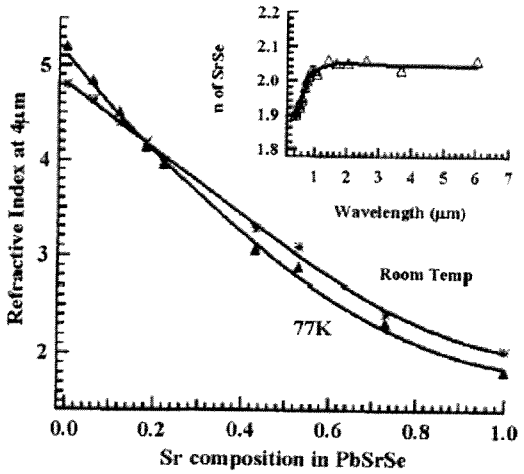

Fig. 4. Refractive index of $Pb_{1-x}Sr_xSe$ for different strontium compositions at room temperature and liquid nitrogen temperature (77 K) at a wavelength of 4 μm. Inset shows refractive index of SrSe at 77 K [23].

dent of temperature in the measured range. The slope of the curvature of the absorption coefficient changes dramatically with the increase of Sr in the compound. Such a change in curvature of the absorption coefficient represents a change in band gap.

The absorption coefficients for direct and indirect transitions are calculated as follows:

$$\alpha_{\text{direct}} = \frac{A}{(h\upsilon)^{1/2}}(h\upsilon - E_g)^{1/2}$$
$$\alpha_{\text{indirect}} = \frac{B}{h\upsilon}(h\upsilon - E_g \pm E_p)^2 \quad (25)$$

Here, A and B are constants, $h\upsilon$ is the photon energy and E_p is the phonon energy. The phonon energy is much lower than the energy band gap and so is neglected. These equations only hold true in the region where the photon energy is greater than the energy band gap. Taking the second derivative of the absorption coefficient (α) (25) with respect to the photon energy yields,

$$\frac{d^2\alpha_{\text{direct}}}{d(h\upsilon)^2} = \frac{-AE_g}{4}\{4(h\upsilon)^3 - 3(h\upsilon)^2E_g\}\{(h\upsilon)^4 - (h\upsilon)^3E_g\}^{-3/2}, \quad (26)$$

which is negative for $h\upsilon > E_g$ and,

$$\frac{d^2\alpha_{\text{indirect}}}{d(h\upsilon)^2} = 2\frac{BE_g^2}{(h\upsilon)^3} \quad (27)$$

which is positive for $h\upsilon > E_g$. As the absorption coefficient changes from negative to positive, the band gap changes from direct to indirect due to the increase in strontium composition [23]. Fig. 5 shows the direct and indirect band gap plotted against the strontium composition at 77 K for $Pb_{1-x}Sr_xSe$.

The following equations were found to describe the best fit for the curve of the calculated direct and indirect band gap energies:

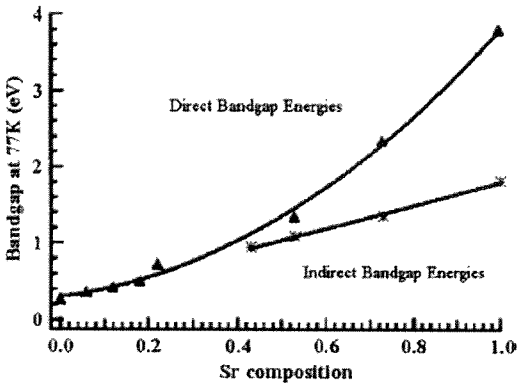

Fig. 5. Direct and indirect band gap energies for $Pb_{1-x}Sr_xSe$ at different compositions of strontium at 77 K [23].

$$E_{g_{\text{direct}}} = 0.278 + 1.356x + 1.040x^2 + 1.144x^3 \text{ (eV)}$$
$$E_{g_{\text{indirect}}} = 0.29 + 1.527x \text{ (eV)} \quad (28)$$

As can be seen, the direct band gap shows a dependence on a polynomial function of the strontium composition (x) while the indirect band gap is linear in nature with respect to the strontium composition. In this experiment, Majumdar et al. found the indirect band gap of SrSe at 77 K to be 1.82 eV and the direct band gap to be 3.81 eV. For PbSe, the direct band gap occurred at 0.278 eV at 77 K. Majumdar et al. assumed that the lowest indirect band gap energy for all strontium compositions (0.43–1) occurs at similar bands which separate linearly as the amount of strontium increases. This occurs as the indirect band gap of $Pb_{1-x}Sr_xSe$ changes linearly with the strontium composition. When the indirect band gap energy curve of lower Sr composition is extrapolated, the direct energy band gap curve crosses at $x \sim 0.2$. This suggests that at $x \sim 0.2 - 1$, the lowest energy band gap of $Pb_{1-x}Sr_xSe$ is direct.

Majumdar et al. have thus, determined the refractive indices and direct and indirect band gaps of $Pb_{1-x}Sr_xSe$ for strontium compositions ranging from 0 to 1. Using the data obtained in this experiment, it would be possible to aid further development of optoelectronic devices based on $Pb_{1-x}Sr_xSe$.

5.2. HgZnTe (MZT)

Ali Omar and El-Akkad [24] performed a study on the optical parameters of mercury zinc telluride (MZT) in solid solution. MZT is a direct band gap material with an energy gap between −0.15 and 2.20 eV depending on the composition ratio. An equation describing the energy gap is,

$$E_g = a_0 + a_1x^{1/2} + a_2x + a_3x^2 + a_4x^3 \text{ (eV)} \quad (29)$$

The best fit was found to be,

$$E_g = -0.1016 + 1.978x + 0.3144x^2 \quad (30)$$

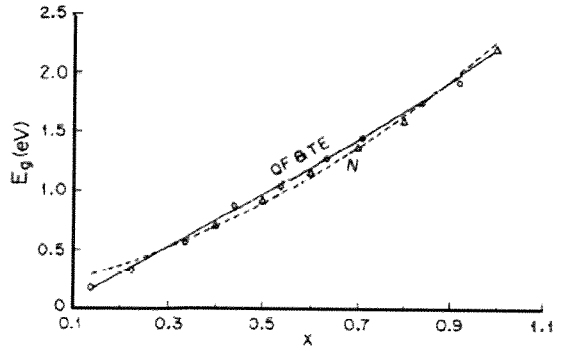

Fig. 6. E_g vs. percent composition. TE refers to the energy gap equation (29) and QF refers to the best fit equation (30). N refers to the ternary solution equation (32) [24].

where x represents the percent composition. According to Nag [25], an empirical formula illustrating the dependency of energy gap on percent composition is given by,

$$E_g = (a - b\overline{Z})^2 \tag{31}$$

where a and b depend on the semiconductor group. When this formula is applied to the ternary solution, $E_g(x)$ becomes:

$$E_g(x) = 0.15 + 0.8637x + 1.245x^2 \text{ (eV)} \tag{32}$$

Fig. 6 describes the results of Eqs. (29)–(32).

Relating the energy gap to the refractive index, Ali Omar and El-Akkad refer to the Ravindra relation. With reference to the best fit, this equation is:

$$n = 4.036 - 1.127x - 0.262x^2 \tag{33}$$

In the range of $0 \leqslant x \leqslant 1$ the quadratic approximation is

$$n_{CM} = 3.84 - 1.56x + 0.428x^2 \tag{34}$$

From graphical illustrations of the refractive index vs. percent composition, as in Fig. 7, the refractive index is seen to decrease with x.

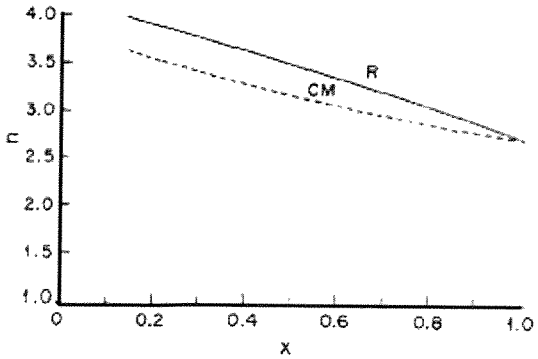

Fig. 7. Refractive index vs. composition ratio. R refers to refractive index with respect to best fit, CM refers to the quadratic approximation [24].

Studying the relation between these parameters, composition and E_g for MZT allows for an understanding and use of the material for IR detectors and solar cells.

5.3. GaN

Bourissa [26] performed a study on the pressure dependence of optoelectronic properties of GaN in the zinc-blende structure. The refractive index was calculated at different pressures with the Moss, Ravindra and Herve Vandamme relations. On a graph of the band lineup vs. pressure, the conduction band minimum is seen to increase with the increase in pressure at Γ yet it is seen to decrease at X. The valence band maximum decreases at Γ (Fig. 8).

The lowest direct energy gap increases with increasing pressure while the indirect gap is seen to have a smaller dependence on pressure. As a result, the direct band gap of zinc-blende GaN cannot be expected to crossover to indirect behavior in the pressure range studied (Fig. 9).

The dependency of band gap on pressure can be represented by,

$$E = a^* + b^*\left(\frac{\Delta a}{a_0}\right) + c^*\left(\frac{\Delta a}{a_0}\right)^2 \tag{35}$$

where $\Delta a = a_p + a_0$ and a_p and a_0 are lattice constants at pressure p and zero pressure. Applying this to the equation describing least square fit, namely,

$$E(p) = E(0) = \frac{dE(p)}{dp}p \tag{36}$$

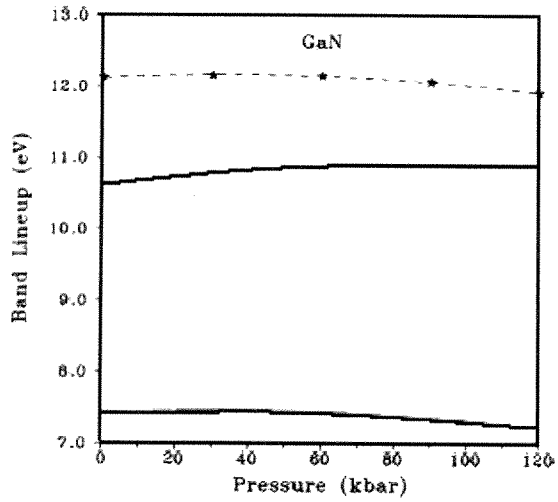

Fig. 8. Band lineup in eV vs. pressure in (kbar). The valence band maximum Γ is indicated by (—) while the conduction band minimum Γ is indicated (···). (-★-) indicates the conduction band X bottom [26].

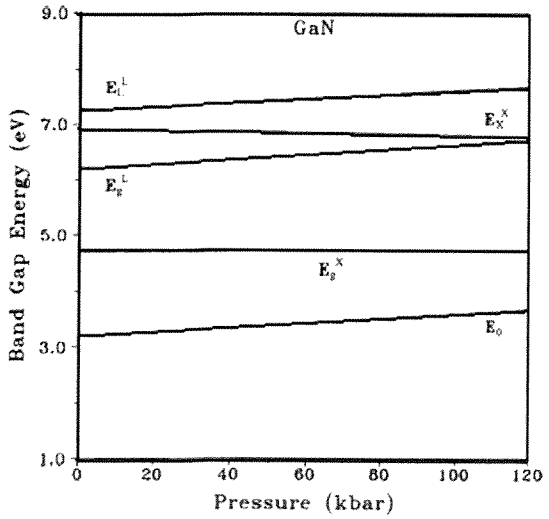

Fig. 9. Direct and indirect band gaps of GaN as a function of pressure in (kbar) [26].

is then,

$$E_0 = 3.2 - 23.83\left(\frac{\Delta a}{a_0}\right) + 195.62\left(\frac{\Delta a}{a_0}\right)^2$$

$$E_X^X = 6.91 + 3.24\left(\frac{\Delta a}{a_0}\right) - 240.65\left(\frac{\Delta a}{a_0}\right)^2$$

$$E_L^L = 7.23 - 23.60\left(\frac{\Delta a}{a_0}\right) + 106.10\left(\frac{\Delta a}{a_0}\right)^2 \qquad (37)$$

$$E_g^X = 4.70 - 0.16\left(\frac{\Delta a}{a_0}\right) + 1.22\left(\frac{\Delta a}{a_0}\right)^2$$

$$E_g^L = 6.20 - 25.48\left(\frac{\Delta a}{a_0}\right) + 166.76\left(\frac{\Delta a}{a_0}\right)^2$$

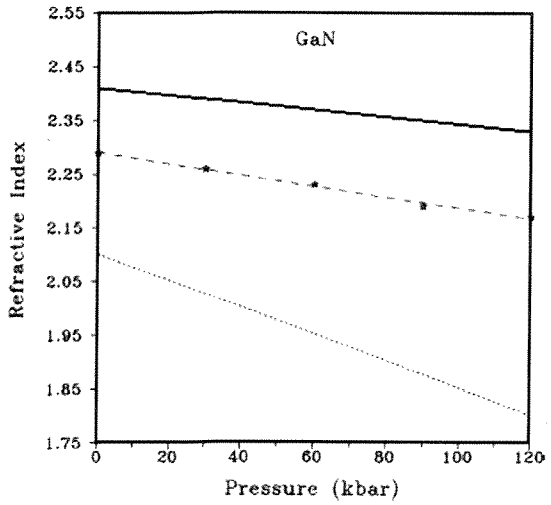

Fig. 10. Refractive index as a function of pressure for GaN. (—) indicates Moss relation, (···) represents Ravindra relation and (-★-) is the Herve Vandamme relation [26].

Table 2
Refractive indices at zero pressure and the first pressure derivative for GaN [26]

Model	n	$\frac{dn}{dp}$ (10^{-3} kbar^{-1})
Moss relation	2.41	−0.67
Ravindra et al. relation	2.10	−2.40
Herve and Vandamme relation	−2.29	−1.13

These dependencies show a non-linear behavior. From Fig. 9, the valence band width is shown to increase with pressure. This can be attributed perhaps to decreasing ionicity under hydrostatic compression which is typical of semiconductors.

The refractive index obtained by the three models is also plotted as a function of pressure in Fig. 10.

In Fig. 10, the refractive index is seen to decrease linearly with respect to increasing pressure, as opposed to the energy gap which shows the opposite behavior. It can be concluded that for GaN and other III–V compounds, the smaller the energy band gap, the larger the refractive index.

In Table 2 are the refractive indices at zero pressure and the first pressure derivative. The refractive index obtained by the Moss relation shows a large difference when compared with the Ravindra and Herve Vandamme relations. Since no known experimental data has been published on the refractive index of GaN, Bourissa calculated the index using $\varepsilon_\infty = n^2$. The results of the refractive index for the three relations were 5.81, 4.41 and 5.24 respectively.

Compared to the experimental value of 5.15 for the dielectric constant obtained by Yu et al. [27], the value 5.24 as obtained by the Herve Vandamme relation is a better result. However, the value of the dielectric function of 5.8, theoretically calculated by Zheng et al. [28] is almost in agreement with the Moss relation.

The pressure derivative of the refractive index is another significant parameter which explains pressure induced distortion. The first pressure derivative calculated through the Ravindra relation indicates a greater effect of applied hydrostatic pressure to the refractive index. Results of calculations show that the pressure derivative is negative and this appears to be the trend among the III–V group. However, a study by Johanssen et al. [29] shows a direct linear relation between refractive index and increase in pressure. Thus, it is believed that ionicity of the material plays a significant role in determining the magnitude and sign of the pressure coefficient of refractive index.

6. Conclusions

The various energy gap–refractive index relations and their applications to semiconductors have been summarized in the above study. The analysis examines the contributions of Moss, Penn, Finkenrath, Wemple and Ravindra. Applications of these relations to IR detector materials such as $Pb_{1-x}Sr_xSe$ and $Hg_{1-x}Zn_xTe$ and UV detector material such as GaN have been discussed.

Acknowledgements

N.M. Ravindra is indebted to Dr. T.S. Moss for his enormous appreciation of the fundamental physics of semiconductors and beyond. He appreciates very much his interactions with Dr. Moss for over 15 years which continued without any impediment from physical or geographical boundaries. Dr. Moss will be dearly missed but his contributions to fundamental semiconductor physics will go a long way. The work reported in this paper is dedicated to Dr. T.S. Moss.

References

[1] N.M. Ravindra, NJIT Internal Memorandum, Applications Specific Coating Technology, July 2005.
[2] Feiming M. Tong, Nuggehalli M. Ravindra, Walter F. Kosonocky, Opt. Eng. 36 (1997) 549.
[3] T.S. Moss, Photoconductivity in the Elements, Academic Press Inc., New York, 1952.
[4] N.F. Mott, R.W. Gurney, Electronic Processes and Ionic Crystals, Oxford University Press, London, 1940.
[5] A. Smekel, Z. Phys. 55 (1929) 350.
[6] F. Zwicky, Proc. Natl. Acad. Sci. USA 15 (1929) 816.
[7] B. Gudden, R.W. Pohl, Z. Phys. 37 (1926) 881.
[8] G.L. Pearson, J. Bardeen, Phys. Rev. 75 (1949) 865.
[9] T.S. Moss, Phys. Stat. Sol. (B) 131 (1985) 415.
[10] D. Penn, Phys. Rev. 128 (1962) 2093.
[11] J. Callaway, Phys. Rev. 116 (1959) 1368.
[12] R.D. Grimes, E.R. Cowley, Phys. Rev. 53 (1975) 2549.
[13] N.M. Ravindra, Sushil Auluck, V.K. Srivastava, Phys. Stat. Sol. (B) 93 (1979) K155.
[14] V.P. Gupta, N.M. Ravindra, Phys. Stat. Sol. (B) 100 (1980) 715.
[15] S.H. Wemple, J. Chem. Phys. 67 (1977) 2151.
[16] N.M. Ravindra, Infrared Phys. 21 (1981) 283.
[17] H. Finkenrath, Infrared Phys. 28 (1988) 327.
[18] T.S. Moss, Optical Properties of Semi-Conductors, Butterworths Scientific Pub., London, 1959.
[19] R.R. Reddy, S. Anjaneyulu, Phys. Stat. Sol. (B) 174 (1992) K91.
[20] R.R. Reddy, Nazeer Y. Ahammed, Infrared Phys. Technol. 36 (1995) 825.
[21] P.J.L. Herve, L.K.J. Vandamme, Infrared Phys. 35 (1994) 609.
[22] P.J.L. Herve, L.K.J. Vandamme, J. Appl. Phys. 77 (1995).
[23] A. Majumdar et al., J. Appl. Phys. 95 (2004) 939.
[24] M. Ali Omar, R. El-Akkad, Infrared Phys. Technol. 39 (1998) 115.
[25] B.R. Nag, Infrared Phys. Technol. 36 (1995) 831.
[26] N. Bourissa, Mater. Chem. Phys. 73 (2002) 51.
[27] G. Yu, H. Ishikawa, M. Umeno, T. Egawa, J. Watanabe, T. Soga, T. Jimbo, Appl. Phys. Lett. 73 (1998) 1472.
[28] R. Zheng, T. Taguchi, M. Matssura, J. Appl. Phys. 87 (2000) 2526.
[29] P.G. Johannsen, G. Reiss, U. Bohle, J. Magiera, R. Muller, H. Spiekermann, W.B. Holzapfel, Phys. Rev. B 55 (1997) 6865.

CHAPTER 3

A Review of the Metal–Insulator Transition in Doped Semiconductors

R.F. MILLIGAN

Muhlenberg College
Allentown, PA 18104, USA

T.F. ROSENBAUM

The James Franck Institute, University of Chicago
Chicago, IL 60637, USA

R.N. BHATT and G.A. THOMAS

Bell Laboratories
Murray Hill, NJ 07974, USA

Electron–Electron Interactions
in Disordered Systems
Edited by
A.L. Efros and M. Pollak

Contents

List of symbols

Defined symbols

e	= charge on electron	ω	= frequency
m^*	= electron effective mass	ϵ	= dielectric constant
$\hbar$	= Planck's constant	ϵ_{Si}	= dielectric constant for silicon
c	= speed of light	ϵ_X	= dielectric constant for impurity
μ_B	= Bohr magneton	χ	= dielectric susceptibility
σ	= conductivity	χ_s	= spin susceptibility
ρ	= resistivity	D	= diffusion constant
$\sigma(0)$	= conductivity at zero temperature	I	= nuclear spin
σ_M	= Mott's minimum metallic conductivity	J	= exchange interaction energy
σ_{IR}	= Ioffe–Regal value of conductivity	T_1	= spin lattice relaxation time
T	= temperature	K	= Knight shift
T_F	= Fermi temperature	Ω	= atomic volume
k_F	= Fermi wave vector	$N(E)$	= electron density of states
l	= electron mean free path	E_F	= Fermi energy
H	= magnetic field strength	E_c	= mobility edge energy
E	= electric field strength	g	= spectroscopic splitting factor
k	= Boltzmann's constant	ξ	= localization length
n	= electron density	N_D	= donor density
n_c	= electron density at MI transition	N_A	= acceptor density
d	= average spacing between impurities	n_{Si}	= refractive index of silicon
d_c	= average spacing between impurities at n_c	$n(0)$	= zero frequency refractive index
a^*	= effective Bohr radius	$\alpha(\omega)$	= IR absorption coefficient

Symbols less commonly encountered or defined in terms of previously defined symbols.

D^+	= lower (hole) Hubbard band	$\alpha(W_N)$	= absorption coefficient of cluster
D^-	= upper (electron) Hubbard band	$R \equiv n^{1/3}\alpha(h\nu)$	= $\alpha_0(h\nu) + \alpha_1 \exp(\epsilon_{opt}/kT)$ IR absorption coefficients
E_2	= activation energy to impurity band	$\Delta\nu_{FW}$	= full width of Raman Spectrum defined by eq. (2.1)
N	= number of impurities in a cluster	θ	= scattering angle in eq. (2.1)
W_N	= energy absorbed by cluster	E_3	= energy between neutral

and ionized impurity sites.
E_1 = activation energy to conduction band
$\sigma_1, \sigma_2, \sigma_3$ = conductivity coefficients defined by eq. (3.3)
m = coefficient in eq. (3.4)
β = exponent in eq. (3.4)
α = function defined on p. 249
κ = screening wave vector
χ = $(2k_F/\kappa)$
F = $[\ln(1+\chi)]/\chi$
B = coefficient in eq. (3.6)
τ_t = relaxation time
W_p = $[4\pi e^2 n/m^*\epsilon]^{1/2}$
γ = $\equiv E_F\hbar/W_p$
c = coefficient defined by eq. (3.7)
ν = valley degeneracy
S_0 = parameter depending on effective mass
$\rho(H,T)$ = field and temperature dependent resistivity
$\Delta\rho$ = change in ρ with field
A_1 = coefficient defined by eq. (3.8)
τ_{in} = inelastic scattering time
H_c = $\equiv \hbar c/(4eD\tau_{in})$
$G(h)$ = integral defined by eq. (3.10)
h = $\equiv g\mu_B H/kT$
A_c = coefficient defined by eqs. (3.11) and (3.12)
A = $\equiv A_1 + A_c$
L = characteristic length in eq. (3.14)
$\rho(T,E)$ = temperature and electric field dependent resistivity, eq. (3.14)
U = Coulomb energy
$\langle|\psi(0)|^2\rangle_{E_F}$ = wave function probability at the nucleus averaged over states at E_F.
ψ_a, ψ_b = electron wave functions
A = constant in eq. (5.1)
c = specific heat
γ = coefficient of T in eq. (5.2)
a = coefficient of T^3 in eq. (5.2)
χ_0 = coefficient defined by eq. (6.4)
ξ = exponent in eq. (6.4)
ν = exponent in eq. (6.5)
σ_0 = coefficient defined by eqs. (6.5), (7.6)
σ_B = conductivity from Boltzmann transport theory
S_F = Fermi surface area
C = constant in eq. (7.4)
Δ = mobility gap
t = power of $(n - n_c)$ term in percolation theory
C' = constant in eq. (7.6)
q = electron wavevector
q_c = wavevector cut-off value at $\sim 1/\xi$
$\Psi(r)$ = electron wavefunction
$\phi_k(r)$ = Bloch wavefunction
$\chi(r)$ = envelope function
a, b = effective mass radii in eq. (8.3)
χ_l, χ_t = components of r along axes of mass tensor in eq. (8.3)
S = magnitude of the stress
C = elastic constant
E_0 = impurity band energy
Ξ = conduction band deformation potential
$U_0(r)$ = electrostatic interaction of outer electron with neutral donor
$U_a(r)$ = second-order perturbation term, eq. (8.7)
$\mathscr{H}$ = system Hamiltonian
$N(0)$ = electron density of states defined by eq. (9.1)
T_0 = characteristic temperature defined by eq. (9.2)
$a_\perp$ = Bohr orbit radius perpendicular to H
$a_\parallel$ = Bohr orbit radius parallel to H.

Part I – General overview

1. *Introduction*

The metal–insulator (MI) transition in disordered systems has been a subject of interest to physicists for many years, going back to the classic papers by Mott (1949, 1956) and by Anderson (1958), which emphasize, respectively, the role of electron correlation and disorder in the phenomenon of electron localization. The basic idea is quite simple, and we take the Mott picture to begin with – consider an ensemble of (neutral) hydrogen atoms held rigidly in a regular array (e.g., simple cubic) in space. If the lattice spacing is small compared to the Bohr radius, so that electron wavefunctions overlap, we expect for this one electron per atom case, following standard band theoretical considerations, that in the ground state the 1s electron band is half filled and the system is a metal at $T = 0$. However when the density is very low, so that electrons on different atoms do not overlap significantly, it costs a finite amount of energy for electrons to hop from their host nucleus to another because of electron–electron repulsion [the Hubbard U (Hubbard 1964a, b)], and the system is insulating at $T = 0$. Thus the system undergoes an MI transition as a function of density on account of electron correlation. This is schematically shown in fig. 1, which depicts the one-electron energy diagram as a function of density (Bhatt and Rice 1981). At low density the occupied level is the 1s state at -1 Ry, while if an extra electron is added to the system it forms a negative hydrogen ion (H^-), which is known to be bound by only 0.055 Ry. As the density is increased, these levels broaden into bands but the system remains insulating so long as a gap exists between the two. At a critical density n_c, the gap vanishes, and the system becomes metallic at $T = 0$, describable by Fermi liquid theory (this is not strictly true for disordered systems, as will be seen later). Clearly in such a system the gap, and hence the insulating state, is due to electron interactions.

Anderson in 1958 showed that the presence of a sufficiently large random potential, even in the absence of electron interactions, leads to localized electron states and hence an insulator at $T = 0$. In such non-interacting systems, the insulator–metal transition is viewed not as a closing of a true one-electron energy gap, but as that of a mobility gap. The mobility gap is the distance of the Fermi level from a critical energy (mobility edge), at which the one-electron states become delocalized. The relevance of the Anderson idea to

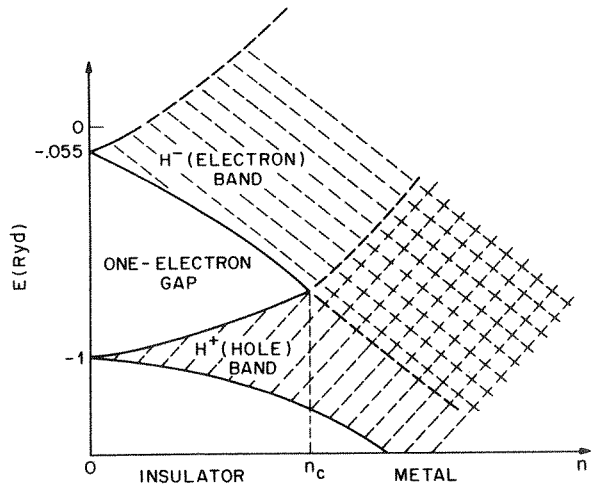

Fig. 1. Schematic diagram of the electron (H^-) and hole (H^+) bands (also referred to as the upper and lower Hubbard bands) for a lattice of hydrogen atoms, as a function of density; a useful diagram for locating n_c though incomplete in its description of random systems such as doped semiconductors.

our hydrogenic system immediately becomes clear if the hydrogen atoms are distributed not on a lattice, but randomly, leading to an effective random potential in which the electrons move. Which of the two effects – disorder or electron interactions – is dominant near the MI transition is an issue debated many times and we have come a full circle to the belief that both of them are necessary to describe the experimental results.

Whatever the detailed differences in various systems from our simple hydrogenic model, one simple formula known as the Mott criterion, relating the critical density at the MI transition n_c to the Bohr radius, a^*, illustrates the universality of the phenomenon:

$$n_c^{1/3} a^* \simeq 0.25. \tag{1.1}$$

Fig. 2 (from Edwards and Sienko 1978) depicts how well eq. (1.1) is satisfied by a variety of systems undergoing an MI transition, with n_c's varying over eight orders of magnitude.

In this chapter we will consider the interplay between the effects of Anderson localization and Coulomb interactions as they appear in the measured properties of doped semiconductors. Doped semiconductors are excellent materials to use for this purpose since they are well characterized and the

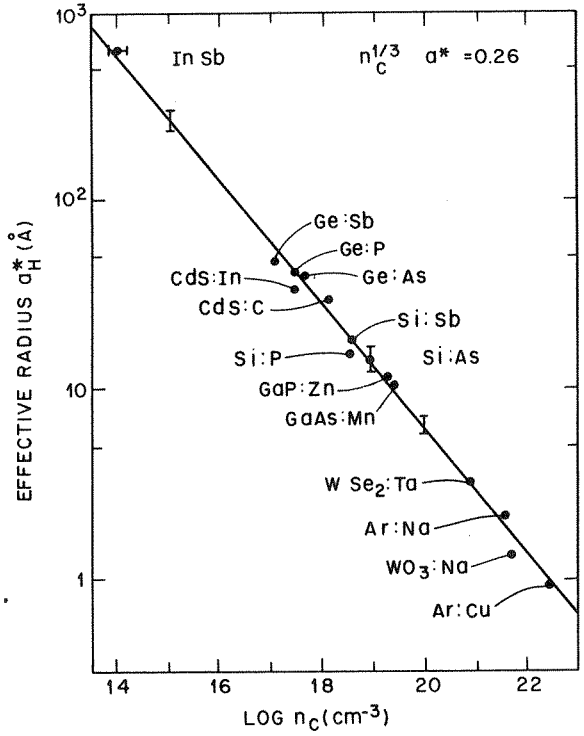

Fig. 2. Mott criterion illustrated for a variety of systems (Edwards and Sienko 1978).

addition of selected impurities allows a controlled progression of properties as the material is taken through the MI transition. The characteristically low electronic energy allows electronic effects of interest to be separated from lattice-dominated processes, although in order to probe the region near the transition, experiments must be done a very low temperatures or, as in the case of optical studies, at far infrared (IR) frequencies. Fortunately, the recent advances in far IR detection and the availability of He-3 dilution refrigerators ameliorates these experimental problems.

In Part I, we will present an overview of experiments close to but outside the immediate region of the MI transition. The optical, transport, magnetic and dielectric properties of doped semiconductors are strongly influenced by both the disorder and the correlations and they often can be analyzed directly in terms of Coulomb interactions and/or localization models. In Part II, we will study the critical region of the transition, with particular emphasis on the

behavior of the zero-temperature conductivity and recently developed stress tuning techniques.

2. Optical properties of doped semiconductors

2.1. Far infrared studies

Doped semiconductors have been widely studied to probe the nature of the MI transition in disordered systems (Mott 1974, Friedman and Tunstall 1978). The outer electrons of shallow donor states have a large effective Bohr radius which

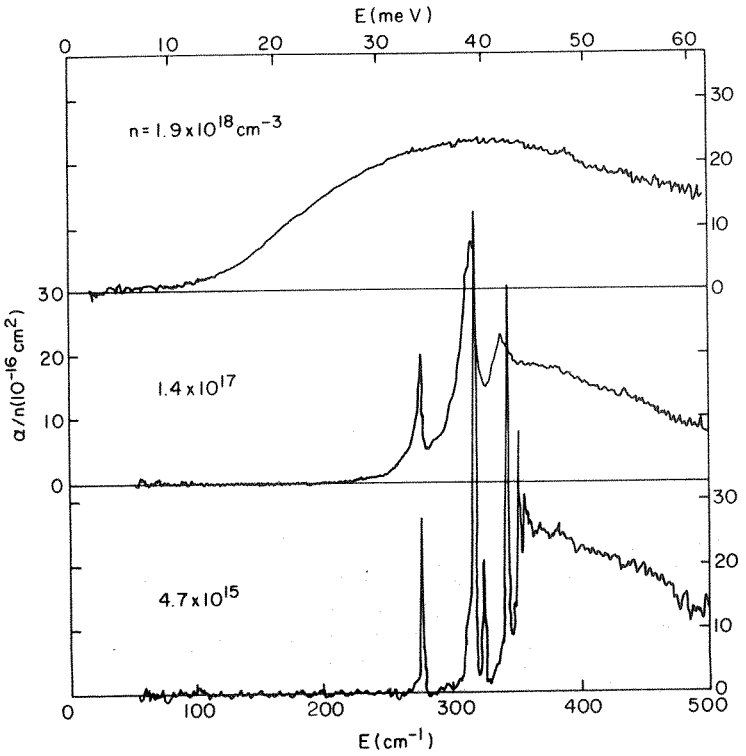

Fig. 3. The actual absorption spectrum of a random system in contrast to model for a periodic system shown in fig. 1. The normalized infrared absorption coefficients for three Si : P samples are shown at $T \simeq 2$ K. The sharp lines characteristic of isolated donors broaden into a continuum at high donor concentrations (Thomas et al. 1981a).

encompasses many lattice sites, making the discreteness of the lattice unimportant in describing the interactions among electrons. At low concentrations, n, of randomly distributed donors there is negligible overlap of the hydrogenic wave functions of these donor electrons and the material is an insulator at temperature $T = 0$ K. In far IR studies this low n region produces sharp absorption lines arising from transitions between bound states, followed by a broad absorption band at high photon energies (that is, energies above ~ 46 meV in Si : P, for example) due to transitions to the conduction band states (Fisher and Ramdas 1969, and references therein).

Figure 3 shows the normalized absorption coefficient measured by Thomas et al. (1981a) as a function of photon energy for three samples of Si : P with differing donor densities, n. Samples were at $T \simeq 2$ K and uncompensated. The low density sample is dominated at low energies by the sharp atomic-like lines. These narrow lines, characteristic of isolated donors, are seen to broaden appreciably in the next sample with $n = 1.4 \times 10^{17}\text{cm}^{-3}$ and disappear completely at $n = 1.9 \times 10^{18}\text{cm}^{-3}$. The initial broadening has been seen in both n-Si, p-Si and GaAs (Thomas et al. 1981a, and references therein) and for Si : P has been interpreted as arising from donor pairs.

At higher donor concentrations, overlap among nearby electrons causes the formation of electron energy bands. The band model proposed by Hubbard (1964a, b) considers Coulomb interactions between electrons only when they reside on the same atom, neglecting longer range forces. For small values of n this model predicts a filled lower (hole) band and empty upper (conduction) band, also simply called the lower and upper Hubbard bands. In terms of donor states these are referred to as the D^+ and D^- bands – the bands of states involved in removal or addition of an electron (e.g., Fritzsche 1978). In the density range of the two more heavily doped samples shown in fig. 3, dc measurements indicate that the conductivity can be parameterized as electron activation to an impurity band with activation energy E_2 (Fritzsche 1955, 1978, 1980, Fritzsche and Cuevas 1960a, b, Mott 1974; also see section 3.1). E_2 decreases with increasing n as the impurity band edge moves down toward the donor ground state. Within the Hubbard model increasing n causes the D^- states to broaden and push the band edge down. This same impurity overlap broadening and lowering of the band edge has been assumed by many workers as the factor responsible for the broad, featureless IR spectrum seen for heavily doped samples shown in fig. 4 (see Capizzi et al. 1980, and references therein).

A band model, although useful for a qualitative picture of why there must be a MI transition, is really a one-electron approximation for uniform systems. In the insulating phase the disorder is large and many-electron (e.g., exciton) effects must be taken into account. A recent interpretation of the optical data shows that the broadening in the far IR for increasing n is a result of the growth in intensity of absorption from random clusters (Capizzi et al. 1979, 1980, Thomas et al. 1981a). The analysis uses a statistical theory of optical

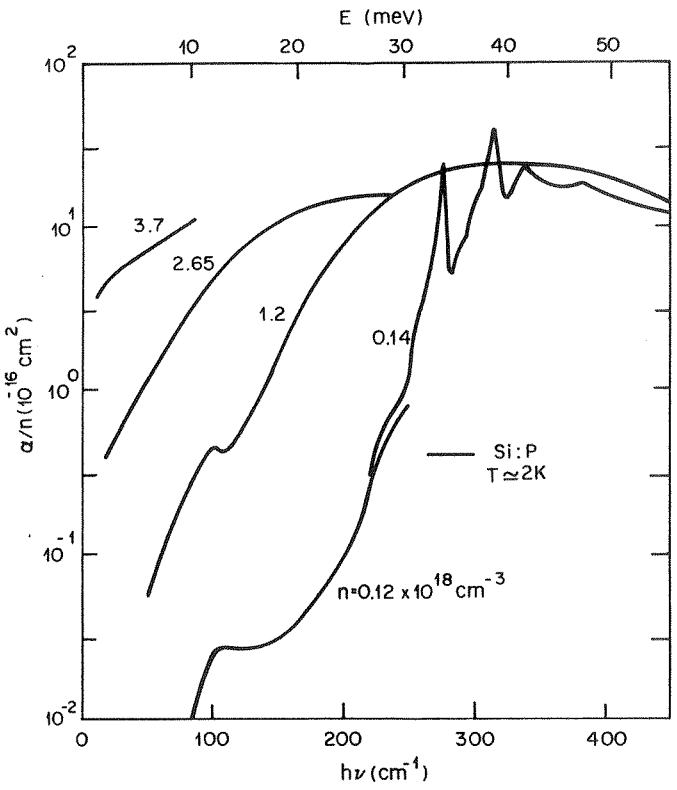

Fig. 4. Absorption coefficients for heavily doped Si : P samples. The lines characteristic of close donor pairs disappear as the concentration is increased. The most heavily doped sample is just on the insulating side ($n_c = 3.74 \times 10^{-18}\text{cm}^{-3}$) (Thomas et al. 1981a).

absorption by donor clusters of various sizes, similar to that used by Bhatt and Rice (1979) to explain the absorption of expanded fluid mercury. In fig. 4 we see the growth of the broad absorption band with increasing n (Thomas et al., 1981a). For the lowest donor density shown absorption from close donor pairs can still be seen. At higher densities the analysis must consider larger clusters. The absence of data at high photon energies for the samples with larger n results from the increasing metallic-like reflectivity of the samples as the critical concentration $n_c = 3.74 \times 10^{-18}\text{cm}^{-3}$ is approached. In the cluster calculations the assumption is made that a cluster of N donors will absorb at an energy W_N with an absorption coefficient $\alpha(W_N) \propto n^N$ as given by the Poisson distribution probability of a cluster of size N. The theoretical absorption coefficient shown as the dotted line in fig. 5. The values $R \equiv n^{-1/3}$. The distance $R = 130\,\text{Å}$ is about twice that at either the peak or average in the Poisson distribution of nearest neighbors.

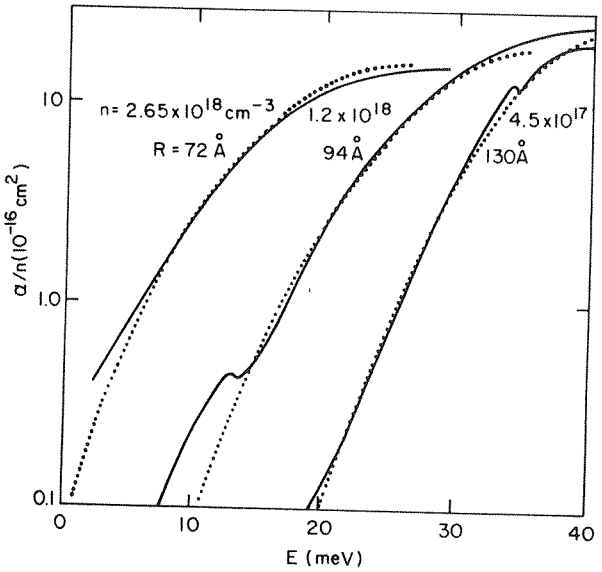

Fig. 5. The absorption coefficients of heavily doped Si : P samples are compared with theoretical curves (dotted lines) based on a cluster calculation. Here $R \equiv n^{-1/3}$ (Thomas et al. 1981a).

An interpretation of the far IR absorption based on Hubbard bands was presented by Narita and Kobayashi (1980) for Ge : Sb as shown in fig. 6. Here the absorption coefficient is shown for four different temperatures. They interpreted the transparent low energy region of the 1.6 K spectrum as resulting from the Hubbard gap and the rise in absorption above ~ 2.2 meV from transitions between the lower and upper Hubbard band. Analysis of the data shows that it fits the following form:

$$\alpha(h\nu) = \alpha_0(h\nu) + \alpha_1 \exp(-\epsilon_{\text{opt}}/kT).$$

Here $\alpha_0(h\nu)$ represents the transitions from the lower to upper Hubbard band. They interpret the second term as evidence of thermally populated localized states in the upper band, from which transitions occur to even higher levels. However, as mentioned earlier, the picture in terms of Hubbard bands does not take into account excitonic states below the band gap. The data of fig. 6

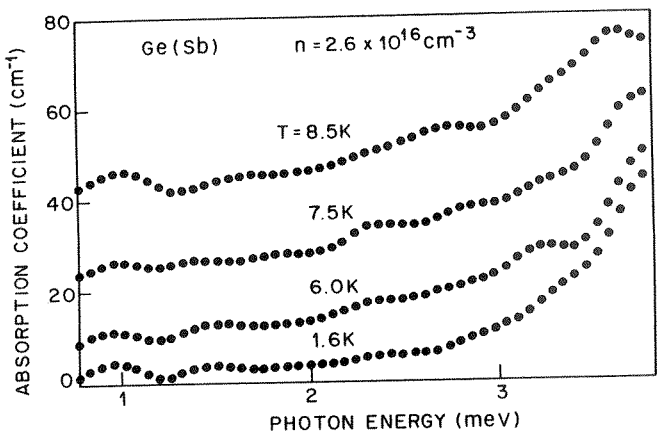

Fig. 6. Absorption coefficients of a Ge : Sb sample at four different temperatures. The data have been interpreted in terms of Hubbard bands (Narita and Kobayashi 1980).

can also be interpreted in terms of excitonic states due to large random clusters.

2.2 *Raman spectroscopy*

The existence of localized, singly occupied donor ground states has been inferred from Raman excitation of these ground states to excited states on the same donor (Doehler 1975, Doehler et al. 1975, Jain et al. 1976). The IR spectrum of Si : P (fig. 3) shows no hint of isolated donors at $n = 1.8 \times 10^{18}\text{cm}^{-3}$, whereas a line from inter-ground-state transitions still appears in the Raman spectrum of a sample at $n = 3.7 \times 10^{18}\text{cm}^{-3}$, just below the critical concentration n_c (Rosenbaum et al., 1980b, Jain et al. 1976). Similarly, the electron spin resonance (ESR) hyperfine split donor signal in Si : P is washed out by $\sim 10^{18}\text{cm}^{-3}$ (Maekawa and Kinoshita 1965). As with the far IR data, the idea of randomly distributed clusters appears germane here.

Figure 7 shows the Raman spectrum of Ge : As for several values of n obtained by Doehler (1975). The peak, Raman shifted by $\sim 35\,\text{cm}^{-1}$, can still be seen in the sample with $n = 4.4 \times 10^{17}\text{cm}^{-3}$, above n_c for Ge : As. Jain et al. (1976), however, report that a similar Raman spectrum for donor states in Si : P disappears as n goes through n_c.

Another application of Raman spectroscopy to the study of electronic properties of doped semiconductors is spin flip Raman spectroscopy, SFRS. In this technique a magnetic field, H, causes an energy level to be split by the

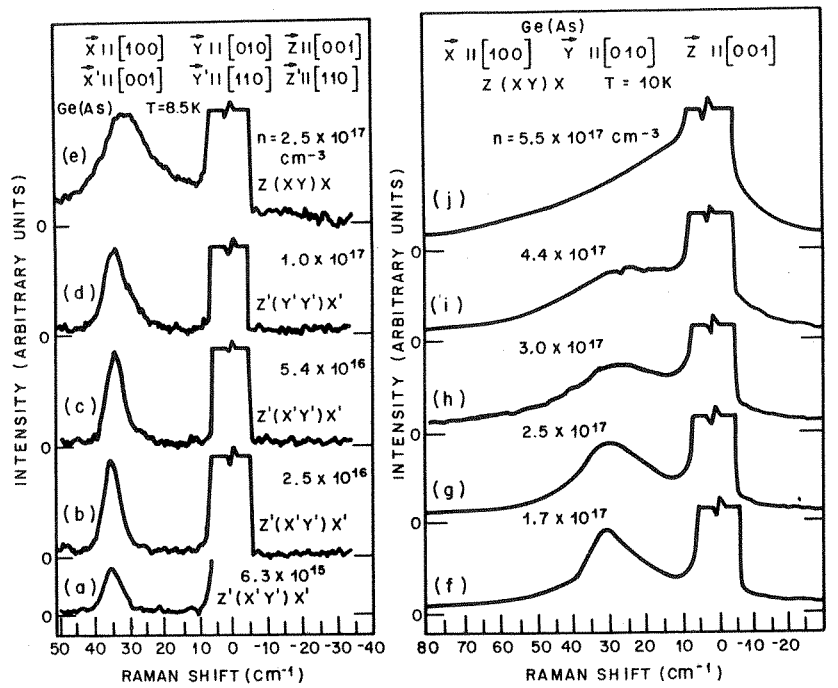

Fig. 7. Raman spectrum of Ge : As. A peak is still visible for the sample with $n = 4.4 \times 10^{17}\text{cm}^{-3}$ which is above n_c (Doehler 1975).

Zeeman energy, $g\mu_B H$, into two levels. An electron at one of these levels absorbs a photon of energy E, bringing it to a virtual energy state, whereupon the electron falls back to the other spin flipped low level with the emission of a photon with energy $E \pm g\mu_B H$. Information on electron dynamics is contained in the SFRS spectra, as discussed by Geschwind et al. (1980) who used the technique to study n-CdS. For mobile carriers the CdS SFRS lineshape is a collisionally narrowed Doppler line. The resulting Lorentzian lineshapes can therefore be related to the diffusion constant or in turn the resistivity, ρ, and carrier concentration, n. Geschwind et al. (1976, 1979, 1980) show that for n-CdS using 4880 Å light the full width at half intensity is given by

$$\Delta\nu_{\text{FW}} = \frac{4.2 \times 10^5}{\rho n^{1/3}} \sin^2(\theta/2)\,\text{cm}^{-1}, \tag{2.1}$$

where θ is the scattering angle. Using the measured value of ρ and n, eq. (2.1)

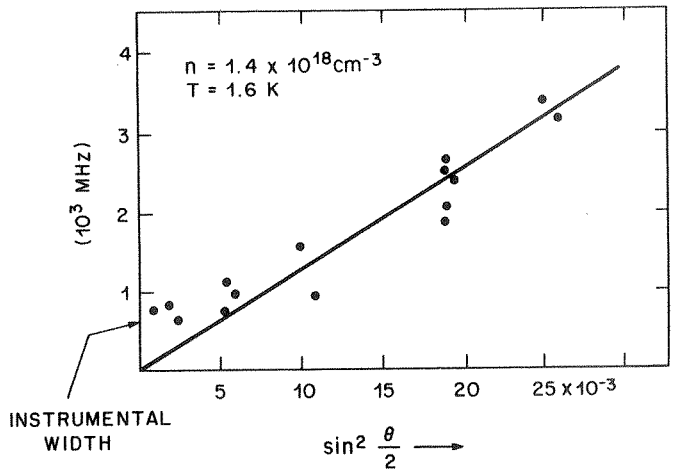

Fig. 8. Linewidth from the spin flip Raman spectrum of n-CdS plotted against $\sin^2(\theta/2)$. The solid line is the best fit to the data and is in qualitative agreement with eq. (2.1) (Geschwind et al. 1980).

gives theoretical values of $\Delta\nu_{FW}$ which are generally within a factor of 2 of the measured linewidths. Since ρ is determined from the bulk sample and the SFRS probes only a small portion of the sample ($\sim 100\,\mu$m in size), macroscopic inhomogeneities may account for the discrepancy. The SFRS lineshape may, in fact, be a reasonable method for probing inhomogeneities in n on this size scale.

The angular dependence of eq. (2.1) is investigated in fig. 8 for n-CdS at 1.6 K (Geschwind et al., 1980). We see that for this metallic sample the linewidth is proportional to $\sin^2(\theta/2)$ for small angle scattering, indicating diffusive behavior.

3. Transport properties

In a series of papers, Mott (1967, 1972, 1981a, b) has put forward the notion of a zero temperature minimum metallic conductivity, σ_M, which in three dimensions is given by

$$\sigma_M = 0.026\, e^2/\hbar d_c, \tag{3.1}$$

where d_c is the average spacing between impurities at n_c. Experimental estimates of σ_M for several doped semiconductors are compared with eq. (3.1)

in fig. 9 (Fritzsche 1978). These estimates are derived from the transport properties of insulating samples (see sections 3.1 and 3.2). We will also consider the role of electron interactions in the transport properties of metallic samples at low temperature (sections 3.3 and 3.4) and briefly examine nonlinear effects on both sides of the MI transition (section 3.5). In Part II, we will return to eq. (3.1) and the current debate about the existence of σ_M.

3.1. The insulating phase

In addition to being a minimum conductivity for metals, Mott has claimed that the same σ_M enters the high T conductivity expression for non-metallic systems where conduction occurs via electrons excited to the mobility edge, E_c. Mott (1967, 1972, 1974, 1978) finds that for the appropriate concentration region

$$\sigma = \sigma_M \exp[-(E_c - E_F)/kT], \tag{3.2}$$

where E_F is the Fermi level. Experimentally other activation energies are also observed as discussed below. The results supporting eq. (3.2) and σ_M for metals are illustrated in fig. 10 which shows the temperature dependence of the resistivity, $\rho \equiv 1/\sigma$, for Ge:Ga (Fritzsche 1955, 1978). For this system $n_c \sim$

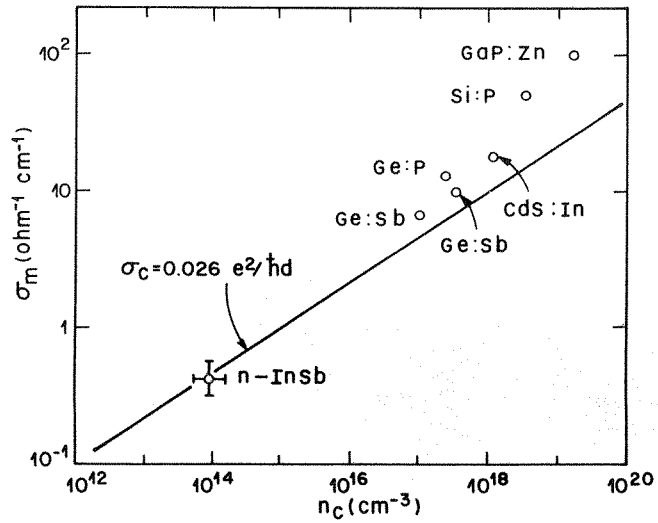

Fig. 9. Historically interesting plot of the estimated values of $\log\sigma_M$ versus $\log n_c$ for various systems. The solid line is a plot of eq. (3.1). In no case is there evidence of a discontinuity in σ at $T = 0$ K (Fritzsche 1978).

$10^{17}cm^{-3}$ and the bottom four curves show the approximately temperature independent resistivity characteristic of metals at low T. As acceptor concentration decreases toward n_c, $\rho \to \rho_M \simeq 0.1\,\Omega\,cm$. Similarly, the lightly doped samples ($n \sim 3$ to $6 \times 10^{16}cm^{-3}$) at high temperatures converge to roughly the same resistivity as predicted by eq. (3.2). This region is characterized by the activation energy $E_c - E_F$ (E_2 in Fritzsche's notation) into an impurity band.

At lower T these same samples are believed to conduct via phonon assisted hopping between neutral and ionized impurity sites with an activation energy E_3 resulting from the random fields of the compensating impurities (Mott 1974). At least a small amount of compensation is required in order for there to be some unoccupied impurity sites (Mott and Twose 1961).

At high T these systems show another activation energy, E_1, to the conduction band. For samples with concentrations just below n_c Mott has suggested that the low T conductivity proceeds via variable range hopping (Mott and

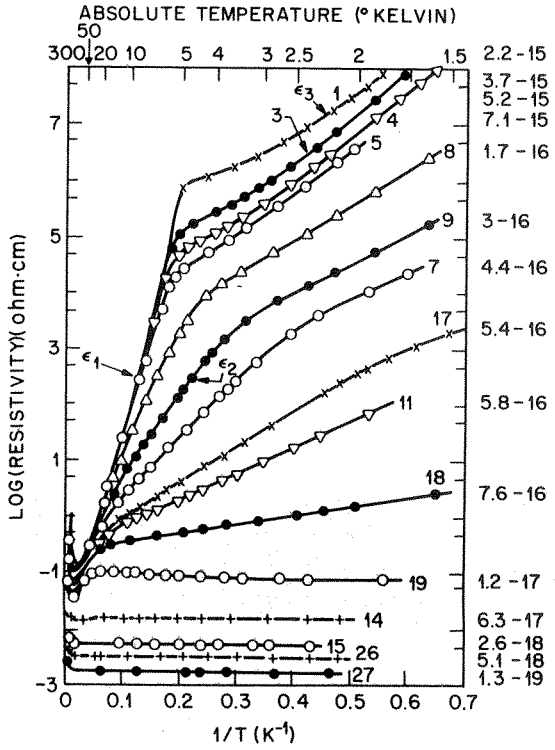

Fig. 10. Plot of $\log\rho$ versus $1/T$ for Ge:Ga. The sample with a donor concentration of $1.3 \times 10^{19}cm^{-3}$ is labeled as 1.3-19. Sample 19 is just on the metallic side of the transition ($n_c \sim 10^{17}cm^{-3}$). Samples with lower donor concentrations show evidence of activation energies as discussed in the text (Fritzsche 1955, 1978).

Davis 1979) which yields $\log(\rho) \propto T^{-1/4}$. The essence of this process is that electrons may preferentially hop to impurity sites which are close in energy to the site they have left, but not necessarily close spatially. Sample 18 in fig. 10 is characterized by variable range hopping at low T. Ignoring the variable range hopping region, these systems have a conductivity given by

$$\sigma = \sigma_1 \exp(-E_1/kT) + \sigma_2 \exp(-E_2/kT) + \sigma_3 \exp(-E_3/kT), \qquad (3.3)$$

where Mott finds $\sigma_2 = \sigma_M$.

3.2. *The Hall effect*

Fig. 11 shows the Hall coefficient as a function of $1/T$ for this same set of samples (Fritzsche 1955, 1978). The transitions between various activation

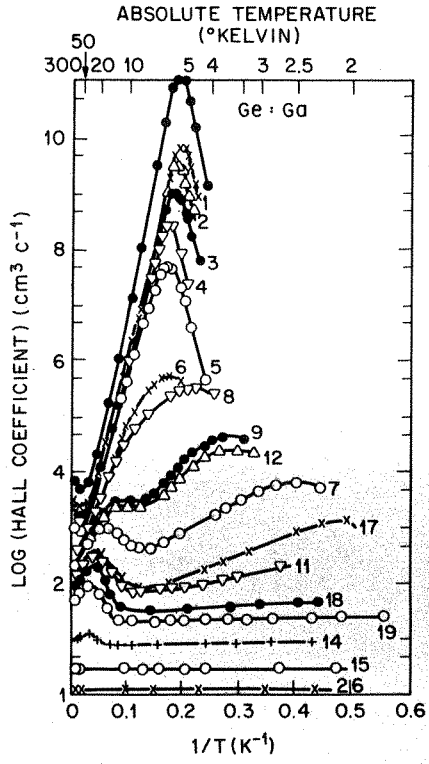

Fig. 11. The log of the Hall coefficient is plotted against $1/T$ for the same samples as in fig. 10. Peaks in the Hall coefficient indicate transition between temperature regions where different activation energies dominate (Fritzsche 1955, 1978).

energy regions is interpreted as the peak in the Hall coefficient. For example, the peak in the top curve at ~ 5 K represents the transition from E_1 to E_3 conduction (compare this with the top curve of fig. 10). At higher concentrations two peaks occur corresponding to the three conduction regions (E_1, E_2 and E_3).

Fritzsche (1978) defines two metallic conduction regions in doped semiconductors. The first starts at the MI transition ($n = n_c$) and conduction is assumed to occur in an impurity band. In this region the low T conductivity increases very rapidly with small increases in n and the Hall mobility rises with n, as shown in fig. 12 for the case of Si : P at 4.2 K (Yamanouchi et al. 1967). The authors suggest that this rise in mobility continues until the Fermi level passes into the conduction band. Further increases in n produce the drop in Hall mobility characteristic of metals.

3.3. *Barely metallic samples: Temperature dependence*

The T-dependence of σ has yielded important information regarding the relative strengths of Coulomb interaction and localization effects. Originally, millidegree measurements of $\sigma(T)$ for metallic Si : P and Ge : Sb (Ootuka et al. 1976, 1979, Sasaki 1980a, b) and n-InSb (Morita et al. 1980) showed a log T dependence which was not readily interpretable in terms of either type of

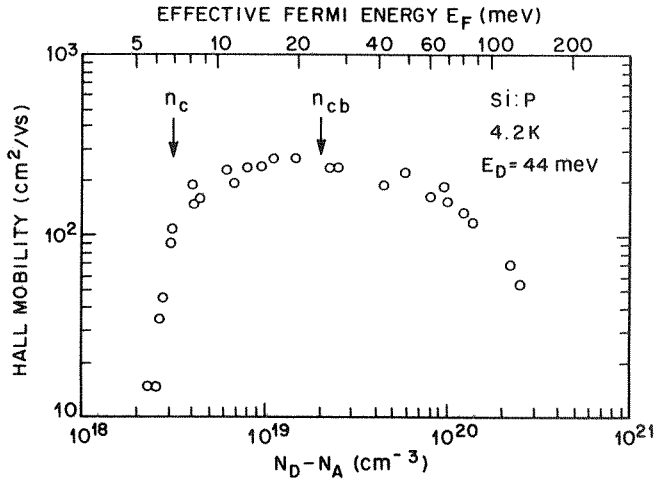

Fig. 12. Log–log plot of Hall mobility against donor concentration for Si : P. The mobility continues to rise as the concentration exceeds n_c. n_{cb} marks the concentration which was assumed to move the Fermi level into the conduction band (Yamanouchi et al. 1967).

theory. More recent measurements on Si : P (Rosenbaum et al. 1981a, b, 1983), Ge : Sb (Thomas et al. 1982b) and n-InSb (Morita et al. 1982) show a power law dependence for $\sigma(T)$. Following Rosenbaum et al. (1981a) we write $\sigma(T)$ as follows:

$$\sigma(T) = \sigma(0) + mT^{\beta}, \tag{3.4}$$

where $\sigma(0)$ is the $T = 0$ K conductivity and the T^{β} term comes from Coulomb interactions. The data for Si : P ($T < 5$ K), n-InSb ($T < 100$ mK), and Ge : Sb ($T < 150$ mK) can be fit to the form of eq. (3.4) using $\beta = \frac{1}{2}$ and a value of m which changes sign with concentration. Measurements on amorphous $Ge_{1-x}Au_x$ alloys (Dodson et al. 1981) yield a positive T correction to $\sigma(0)$ which we find can be fit to $\beta = \frac{1}{2}$, although the data was originally interpreted as fitting $\beta = \frac{1}{3}$.

The Coulomb interactions model with electron–electron and electron–hole scattering in the presence of random impurities predicts a $T^{1/2}$ correction to σ (Altshuler and Aronov 1979a, b, Lee and Ramakrishnan 1982). In the limit of weak scattering $k_F l \gg 1$, where k_F is the Fermi wave vector and l is the mean free path, they find

$$\sigma(T) = \sigma(0) + \alpha\left[\tfrac{4}{3} - 2F\right]\sqrt{T}, \tag{3.5}$$

where

$$\alpha = 0.025\left[T_F\left(\frac{m^*D}{\hbar}\right)^3\right]^{-1/2}\sigma(0).$$

Here T_F is the Fermi temperature, m^* the effective mass, and D the diffusion constant. The dimensionless term F results from the Hartree interaction and is a function of $x = (2k_F/K)^2$, where K is the screening wave vector. It is given by $F = x^{-1}\ln(1+x)$, which can range from 0 to 1. In the Si : P, Ge : Sb and n-InSb studies most samples had a donor concentration which yielded negative values of m ($F > \frac{2}{3}$). However, for samples very close to the transition, screening breaks down (K^{-1} diverges) and $F \to 0$, causing m to change sign. Also, the magnitude of m (determined mostly by α) grows as $n \to n_c$, since α is proportional to $\sigma(0)^{-1/2}$. Although there is good qualitative agreement of m with n, the best fits to the data differ from theory by a factor of 2.6 for Ge : Sb and ~ $\frac{1}{8}$ for Si : P. Part of the disagreement may result from the fact that the theories do not include either the anisotropy or the effect of intervalley scattering, which occurs at an unknown rate in Ge and Si (Rosenbaum et al. 1981a, b, Thomas et al. 1982b). Recent calculations by Bhatt and Lee (1982) consider the effects of intervalley scattering, valley degeneracy and valley anisotropy. The anisotropy turns out to be the most important correction and

with its inclusion agreement with theory is found both for the single valley semiconductor n-InSb (Morita et al. 1982) and the many-valley Si : P and Ge : Sb.

In Ge : Sb the simple $T^{1/2}$ dependence for $\sigma(T)$ is restricted to very low temperatures (< 150 mK). The same dependence extends up to ~ 5 K in Si : P, partly because the electronic and lattice characteristic energies are greater in Si than in Ge. Thomas et al. (1982b) find that the inclusion of a linear T term provides a good fit to the Ge : Sb data up to 500 mK. They write

$$\sigma(T) = \sigma(0) + mT^{1/2} + BT. \tag{3.6}$$

The BT term arises from localization theory (Gorkov et al. 1979) assuming an energy relaxation time τ_ϵ of the form $1/\tau_\epsilon = cT^2$. Such a form is expected if the energy relaxation proceeds via electron–electron scattering. τ_ϵ is then given (Quinn and Ferrill 1958) by

$$\frac{1}{\tau_\epsilon} = 4\left(\frac{\pi}{4}\right)^3 \frac{e^2 k_F \gamma}{\sqrt{3}\,\hbar\epsilon}\left(\frac{k_B T}{E_F}\right)^2 = cT^2,$$

where ϵ is the dielectric constant and

$$\gamma = E_F/\hbar W_p, \quad W_p = \sqrt{\frac{4\pi e^2 n}{m^*\epsilon}}\,.$$

B from eq. (3.6) is related to c as follows:

$$B = S_0 \nu \frac{e^2}{2\pi^2\hbar}\sqrt{\frac{c}{D}}\,, \tag{3.7}$$

where D is the diffusion constant, ν the valley degeneracy ($\nu = 4$ for Ge) and S_0 a parameter which depends on the effective mass tensor.

The fit to eq. (3.6) is shown in fig. 13 for a sample with a net donor density $n = 3.3 \times 10^{17}\,\mathrm{cm}^{-3}$ ($n_c \sim 1.5 \times 10^{17}\,\mathrm{cm}^{-3}$ for Ge : Sb) and compensation $\lesssim 5\%$ (Thomas et al. 1982b). Here the abscissa is linear in $T^{1/2}$. The dashed–line shows the best fit using only a $T^{1/2}$ term ($B = 0$). The fit can be extended to ~ 500 mK if both the $T^{1/2}$ and T terms are included. The solid line is the sum of these two terms (dashed–dotted lines).

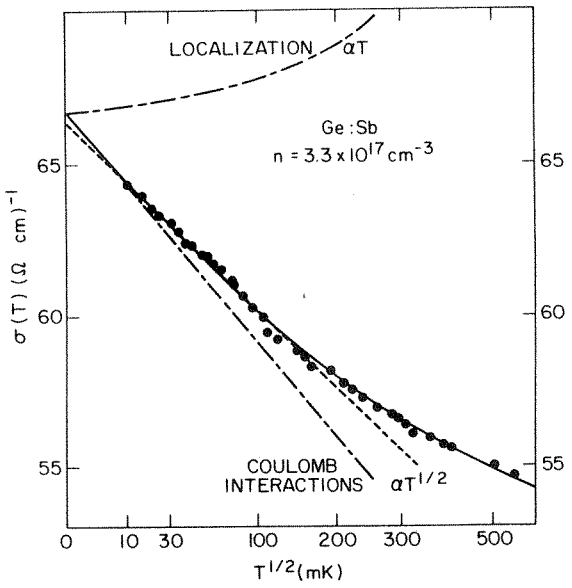

Fig. 13. $\sigma(T)$ plotted against $T^{1/2}$ for a barely metallic sample of Ge : Sb. The data have been analyzed in terms of a linear T localization term and a $T^{1/2}$ Coulomb interaction term. The two lower dashed lines assume different values of $\sigma(0)$. The solid line is a best fit to the data using both the T and $T^{1/2}$ terms as in eq. (3.6) (Thomas et al. 1982b).

3.4. *Barely metallic samples: Magnetic field dependence*

Early work on metallic n-Si and n-Ge (Alexander and Holcomb 1968, and references therein) have generally shown a negative magnetoresistance, although the magnitude and sign of $\Delta\rho/\rho$ change with concentration, temperature and field strength. More recently, extensive studies of the low T (< 1 K) magnetoresistance have been made on Si : P (Sasaki 1980a, Rosenbaum et al. 1981b), Ge : Sb (Ootuka et al. 1979, Sasaki 1980b) and n-InSb (Morita et al. 1980, 1982). The behavior of the magnetoresistance has been interpreted as arising from donor spins (Alexander and Holcomb 1968, Mott 1974, Sasaki 1980a, b, Ootuka et al. 1979), localization (Kawabata 1980, 1981), Coulomb interactions (Altschuler et al. 1980a, b, Lee and Ramakrishnan 1982) or a combination of localization and Coulomb interactions (Rosenbaum et al. 1981b, Chui et al. 1981, Morita et al. 1982).

Within the framework of Anderson localization (Abrahams et al. 1979), Kawabata (1980) predicts a negative magnetoresistance which varies as H^2 at

low field and $H^{1/2}$ at high field. Specifically he finds

$$\Delta\rho/\rho(0,0) = -0.031\left[\rho(0,0)e^2/\hbar\right]\left(\frac{eH}{\hbar c}\right)^{1/2} \equiv A_l H^{1/2}, \tag{3.8}$$

for $H \gg H_c$, where

$$\Delta\rho/\rho(0,0) = [\rho(H,T) - \rho(0,T)]/\rho(0,0), \quad A_l = -0.918\rho(0,0)$$

for H in kOe and ρ in Ω cm, and $4eH_c/\hbar c \equiv (D\tau_{in})^{-1}$. Here D is the diffusion constant and τ_{in} is the inelastic scattering time. Estimates of H_c for Si : P, Ge : Sb and n-InSb indicate that the condition $H \gg H_c$ is easily satisfied experimentally.

Using the Coulomb interactions model, eq. (3.5) can be extended to include the effect of magnetic fields (Rosenbaum et al. 1981b, 1983, Lee and Ramakrishnan 1982). In the presence of H, eq. (3.5) becomes

$$\rho(H,T) = \rho(0,0) - \alpha\left[\tfrac{4}{8} - F\right]T^{1/2} + \alpha F T^{1/2} G(h)/G(0), \tag{3.9}$$

where

$$G(h) = \int_0^\infty dw \frac{\partial^2}{\partial w^2}\left(\frac{w}{e^w - 1}\right)\left[(w+h)^{1/2} + (|w-h|)^{1/2}\right], \tag{3.10}$$

and $h = g\mu_B H/kT$. They find $G(h)/G(0) \approx 0.77\sqrt{h}$ for $h \gg 1$ and $G(h)/G(0) \approx 1 + \sigma(h^2)$ for $h \ll 1$. This yields a magnetoresistance which is always positive. For $g\mu_B H \gg kT$ eq. (3.9) becomes

$$\Delta\rho/\rho(0,0) = -\alpha\left(\tfrac{4}{3} - F\right)T^{1/2}/\rho(0,0) + 0.77\alpha F(g\mu_B H/k)^{1/2}/\rho(0,0) \tag{3.11}$$

$$\equiv -BT^{1/2} + A_c H^{1/2}, \tag{3.12}$$

where the term A_c is independent of T. The magnetoresistance can now be written as

$$\Delta\rho/\rho(0,0) = (A_l + A_c)H^{1/2} - BT^{1/2}$$

$$\equiv AH^{1/2} - BT^{1/2}, \tag{3.13}$$

where the localization and interaction effects are assumed to be additive. Such

an analysis of the low T magnetoresistance was first done on Si : P and Ge : Sb by Rosenbaum et al. (1981b) and subsequently on granular aluminum (Chui et al. 1981) and n-InSb (Morita et al. 1982). In fig. 14 we show the results for n-InSb, where ρ is plotted against $\sqrt{H}$ for two different temperatures. As in the Si : P and Ge : Sb cases, ρ has a $\sqrt{T}$ dependence at constant field values and the slopes of the two curves in fig. 14 are essentially independent of temperature and field direction as predicted by eq. (3.13). From the slopes of the ρ versus $H^{1/2}$ and $T^{1/2}$ curves the values of A and B can be obtained. Subtracting the calculated value of A_l from A, the values of A_c and B are compared with theory (eq. 3.11). The qualitative agreement with theory is quite good, although the actual values of the parameters differ by a factor of 2 to 3. One common problem encountered in interpreting the low T magnetoresistance of all three materials is that the largest effects occur near the transition where $k_F l \sim 1$, contrary to the conditions of the theory, $k_F l \gg 1$.

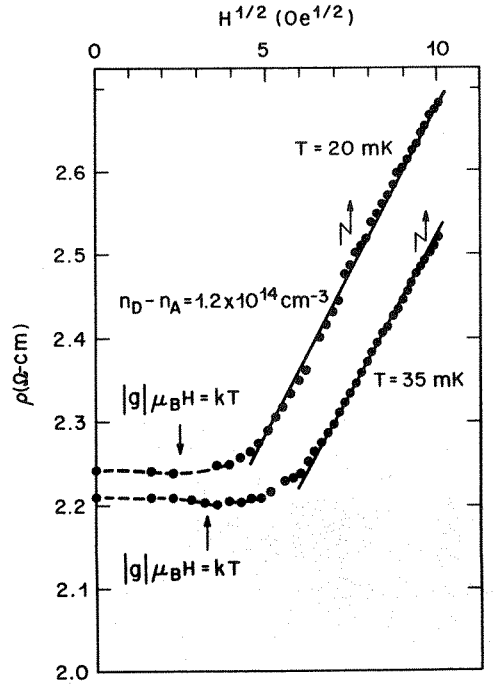

Fig. 14. Resistivity of a metallic n-InSb sample plotted against $H^{1/2}$. At high fields ρ is proportional to $H^{1/2}$ with a slope independent of T, as shown in eq. (3.13) (Morita et al. 1982).

3.5. *Electric field effects near the transition*

At millidegree temperatures the resistivity of Si : P samples near n_c shows a marked electric field dependence not related to lattice heating. This has been observed for samples with $n < n_c$ (Rosenbaum et al. 1980b) and $n > n_c$ (Rosenbaum et al. 1983). The data for samples on the metallic side is not understood; however, on the insulating side the results can be interpreted in terms of a hopping model. For small electric fields E the model predicts (Hill 1971, Pollak and Riess 1976)

$$\rho(T,E) \propto \rho(T)\exp(eEL/k_BT), \tag{3.14}$$

where L is a characteristic length. Figure 15 shows $\ln(\rho/\rho_M)$ plotted against E/T, where $\rho_M \equiv 1/\sigma_M$ conveniently scales the data. Using the line for $T = 392$ mK yields $L = 16\,\mu$m compared with the Bohr radius of the isolated donor of ~ 16 Å. Eq. (3.14) is calculated for $eEL \ll k_BT$ and this condition may not be well satisfied for the lower values of T, as the poorer fits suggest.

At present, the large value of L implied by the data is not understood. L is much larger than either the average donor spacing $n^{-1/3} = 63$ Å or estimates of the localization length $\xi \sim 10$ Å for amorphous semiconductors (Mott 1974, Mott and Davis 1979). Scaling theories of localization (see Part II) suggest a divergent length scale near the MI transition. However, even estimates of this are ~ 250 times too small (Capizzi et al. 1980, Hess et al. 1982, Rosenbaum et al. 1983).

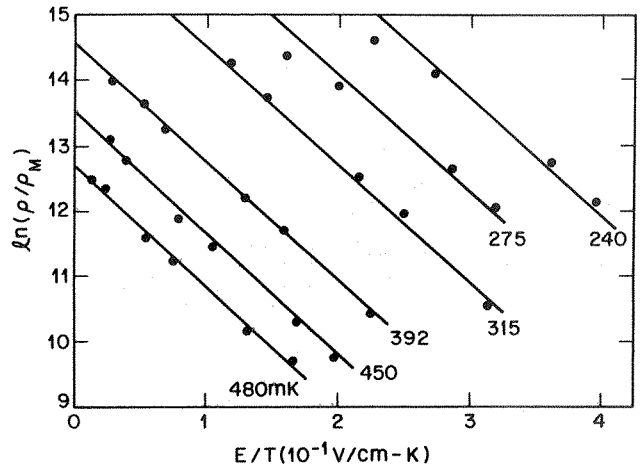

Fig. 15. Plot of $\ln(\rho/\rho_M)$ versus E/T at several temperatures for a barely insulating Si : P sample. Here $\rho_M \equiv 1/\sigma_M$ and E is the electric field (Rosenbaum et al. 1980b).

4. *Magnetic properties*

In this section, we discuss a number of magnetic field related phenomena, such as electron spin resonance, Knight shift, spin susceptibility and Faraday rotation. Extensive reviews of magnetic phenomena can be found in Alexander and Holcomb (1968), Holcomb (1978) and Tunstall (1980).

4.1. *Electron spin resonance studies*

The effects of donor clustering seen in the far IR absorption studies are also apparent in the electron spin resonance (ESR) spectra of donor electrons in Si : P (Fletcher et al. 1954, Slichter 1955, Feher et al. 1955, Maekawa and Kinoshita 1965). At low donor concentrations ($n \lesssim 10^{17}$cm^{-3}) the spectrum is

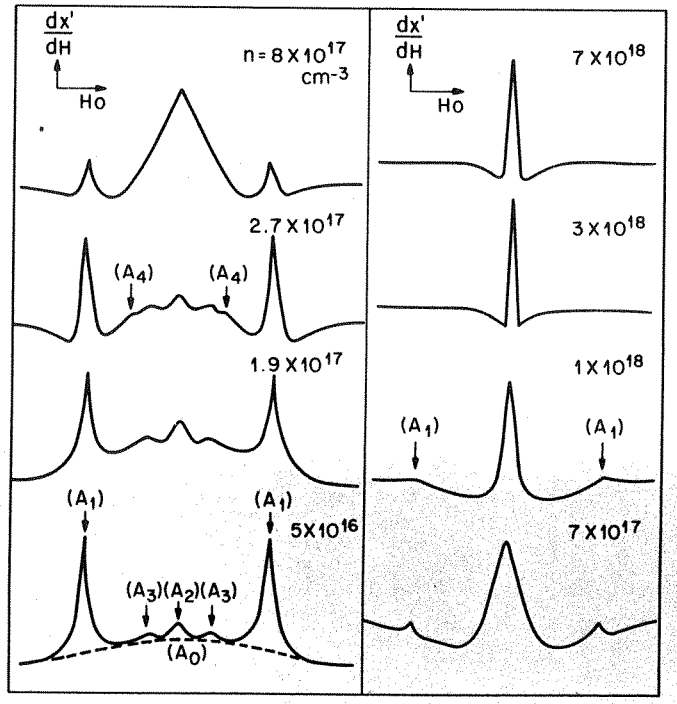

Fig. 16. ESR spectra of donor electrons in Si : P for samples with various donor concentrations. The hyperfine doublet, A_1, which dominates at low concentrations is characteristic of isolated donors, however signals from donor pairs, A_2, and triplets, A_3, are also visible. At high concentrations the only signal which remains is a single narrow line characteristic of delocalized spins (Maekawa and Kinoshita 1965).

dominated by two hyperfine lines of the 100% abundant ^{31}P nucleus ($I = \frac{1}{2}$). Increasing n introduces signals from donor pairs and larger donor clusters. In fig. 16 we show a series of ESR spectra of the donor electrons in Si : P at various values of n. The curves are taken from Maekawa and Kinoshita (1965). We calculated the values of n using their reported values of ρ at 300 K and the neutron activation calibration curve for Si : P (Mousty et al. 1974). The donor concentration determined in this fashion is somewhat larger than the values they report. For example, where they reported an excess donor concentration of $1.7 \times 10^{18}\text{cm}^{-3}$, we find $3 \times 10^{18}\text{cm}^{-3}$.

In Sample 1 the predominant resonance, A_1, is due to isolated ^{31}P atoms. Even at this low concentration, where the average spacing between impurities is ~ 300 Å, signals from donor pairs, A_2, and triplets, A_3, can be seen (Slichter 1955, Holcomb 1978). The central resonance, which rapidly grows and narrows with increasing n, is evidence of interaction among electrons (Holcomb 1970, 1978, Holcomb and Rehr 1969). The narrowest line occurs at a value of n far below the MI transition in Si : P. The observation that charge delocalization occurs at a value of n at least 3 times that of spin delocalization has been interpreted as evidence of correlation effects (Holcomb 1978). The exchange interaction J between neighboring spins can be large enough for the electrons to average the hyperfine interaction of many ^{31}P nuclei, producing the large narrow resonance. In order for charge transport to occur, however, an additional electron must be placed on the donor atom, which would cause one electron to be raised by the Coulomb energy, U, to a D^- state. As long as $J < U$ this Coulomb barrier prevents charge transport at low T.

4.2. *Spin susceptibility*

The spin susceptibility, χ_s, has been measured in doped semiconductors by a variety of techniques: direct static measurements (Sonder and Stevens 1958), integration of ESR spectra (Quirt and Marko 1971, 1972, 1973, Jerome et al. 1965, Ue and Malkawa 1971), low T SQUID magnetometer measurements (Andres et al. 1981) and Faraday rotation (Kummer et al. 1978, Walstedt et al. 1979, Geschwind et al. 1980).

In fig. 17 we see the T dependence of χ_s^{-1} from the ESR integration results of Quirt and Marko (1973) for Si : P with $n = 5.9 \times 10^{18}\text{cm}^{-3}$. Although this sample is clearly in the metallic range as far as conductivity studies suggest ($n_c = 3.74 \times 10^{18}\text{cm}^{-3}$), the strong T dependence of χ_s^{-1} cannot be explained by the T-independent Pauli susceptibility expected for degenerate electrons. It is not until higher concentrations ($n \sim 1\text{–}2 \times 10^{19}\text{cm}^{-3}$) that χ_s is independent of T at low temperatures, the suggestion being that the T dependence arises from correlation effects (Holcomb 1978).

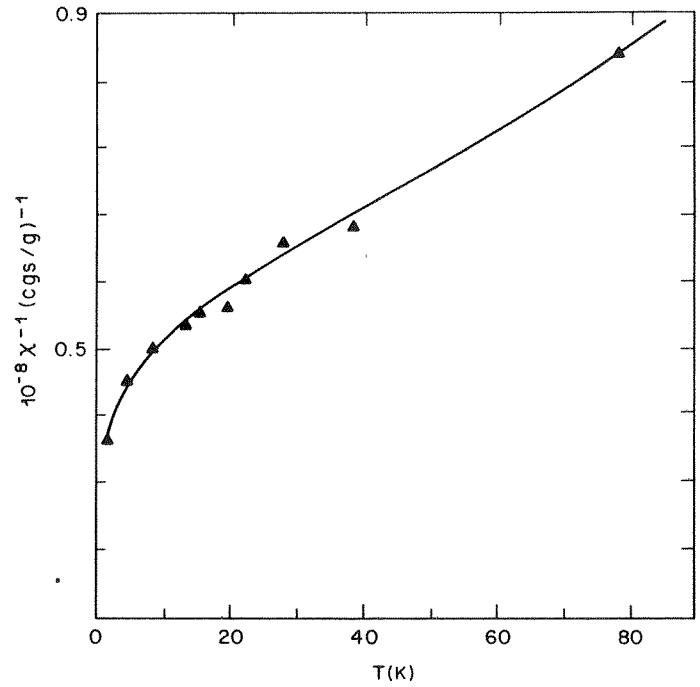

Fig. 17. The temperature dependence of the inverse of the spin susceptibility of Si : P with $n = 5.9 \times 10^{18}\text{cm}^{-3}$. This metallic sample does not show the T-independence of χ expected for degenerate electrons (Quirt and Marko 1973).

Figure 18 shows the results of Andres et al. (1981) for Si : P on the insulating side. These measurements were extended to both low T and low field. The unexpected leveling off of χ_s below ~ 30 mK is roughly independent of concentration and remains unexplained. The data of higher T, however, has been adequately explained using two different types of cluster calculations which involve exchange coupling between the many-valley orbital ground states of the P atoms. The first is a pair calculation which has been modified to include larger clusters (Bhatt and Rice 1980) while the second is a computer simulation cluster analysis (Walstedt et al. 1979). The dashed lines in fig. 18 are the results of these calculations. The solid lines indicate the slope for Curie law behavior. The deviation from Curie law behavior is also seen in fig. 19 which shows χ_s versus n for two temperatures. Note that $n_c = 3.74 \times 10^{18}\text{cm}^{-3}$ is well off the scale in this plot. The deviation from linear Curie law behavior ($\chi_s \propto n$) is most dramatic at low T, but the data is still in good agreement with

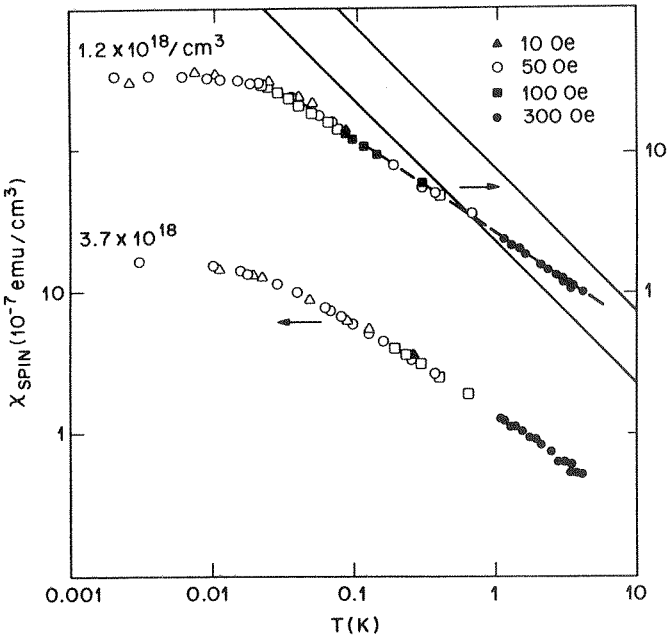

Fig. 18. The temperature dependence of the spin susceptibility for insulating samples of Si : P. The dashed lines are the results of cluster calculations while the solid line is the Curie law for free spins (Andres et al. 1981).

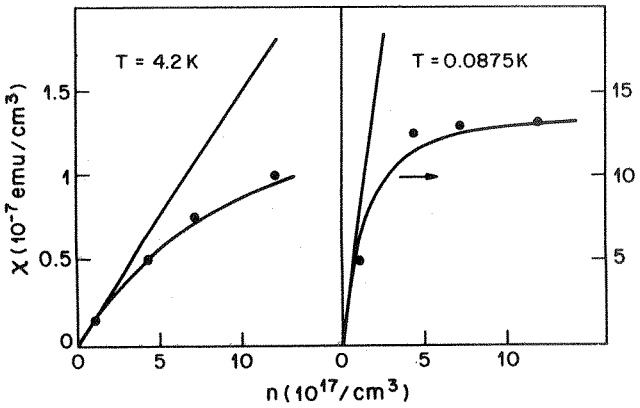

Fig. 19. Donor spin susceptibility plotted against donor concentration for insulating samples. The straight line is the Curie law behavior, while the curved line is based on the modified pair approximation (Andres et al. 1981).

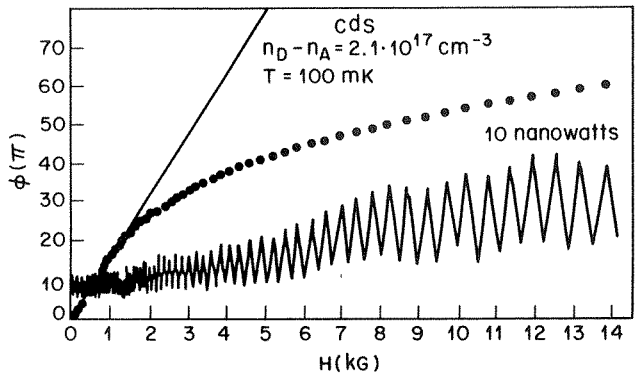

Fig. 20. Recorder tracing of the Faraday rotation of linearly polarized light transmitted through CdS as a function of magnetic field. The dots show the accumulated rotation which is linear in H at small field with a slope proportional to χ_s (Geschwind et al. 1980).

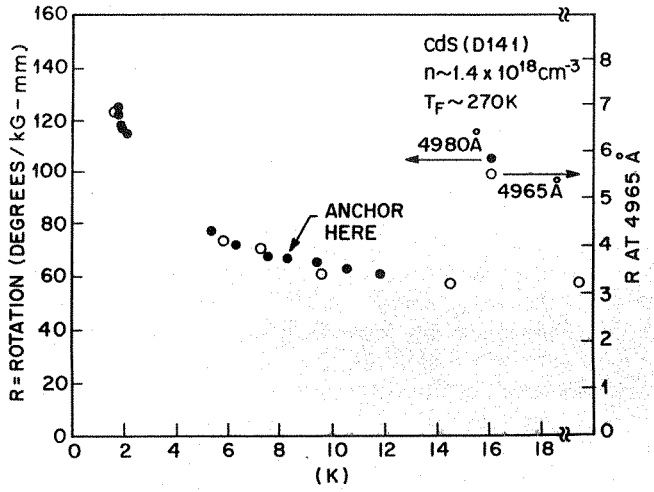

Fig. 21. The temperature dependence of the Faraday rotation for a barely metallic sample of CdS with a Fermi temperature of 270 K. Two wavelengths were used and normalized to each other at 8.5 K (Geschwind et al. 1980).

the calculated value of χ_s using the modified pair approximation (curved line).

A different approach to measuring χ_s is by observing the Faraday rotation of linearly polarized light as it passes through a sample in a magnetic field. Romestain et al. (1975) have shown that the rate at which the rotation changes with field is proportional to χ_s at small fields. The technique selectively measures χ_s of the donor in very small concentrations even though magnetic impurities may be present, an advantage over methods which measure the bulk susceptibility. This technique has successfully been applied to n-CdS at low T (Kummer et al. 1978, Walstedt et al. 1979, Geschwind et al. 1980). Figure 20 shows the Faraday rotation Φ as a function of magnetic field. The recorder trace shows the signal transmitted through an analyser, where each positive peak corresponds to an additional π rotation of the plane of polarization. The accumulated rotation is shown as the dots on the graph, the slope of which is constant at small H and proportional to χ_s.

The temperature dependence of the Faraday rotation for a barely metallic sample of CdS is shown in fig. 21. Here two different wavelengths were used and the data made to coincide at 8.5 K. The temperature dependence of χ_s (proportional to the rotation) is similar to the data of Quirt and Marko (1973) for Si : P shown in fig. 17.

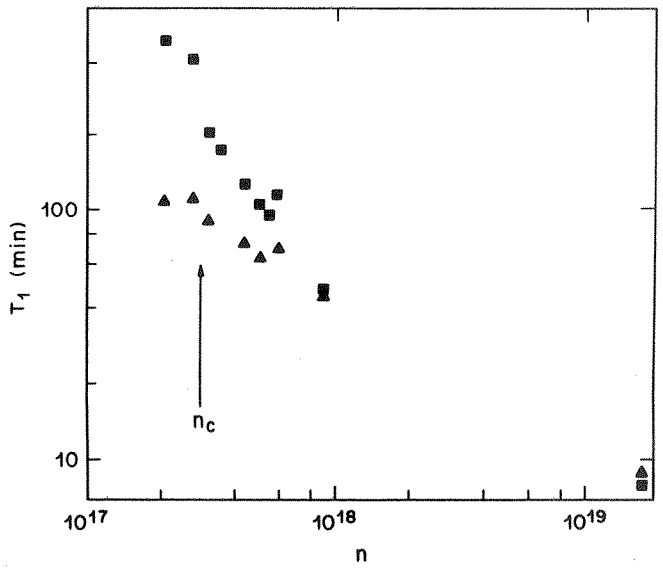

Fig. 22. Nuclear spin–lattice relaxation time, T_1, plotted against donor concentration for Ge : As. Both samples are metallic (Tunstall and Deshmukh 1979).

4.3. Nuclear magnetic resonance studies

The nuclear spin–lattice relaxation time, T_1, can be used to provide information about electron spin motion since it monitors the strength of the interaction between the electron and nuclear spins. Spin motion, however, may not be the same as charge motion. The concentration dependence of T_1 for Ge : As (Tunstall and Deshmukh 1979) is shown in fig. 22. These samples are above the MI transition and the only temperature dependence of T_1 is the usual T^{-1} found for degenerate systems (i.e., only electrons within kT of the Fermi level participate in the process). An interesting feature of the data is the increase in T_1 with field, a process which grows as $n \to n_c$. Tunstall (1980) suggests that this quenching of the spin–lattice relaxation may be due to local moments.

A related topic is the Knight shift, K, which depends on both χ_s and the electron wave function probability at the nucleus $\langle|\psi(0)|^2\rangle_{E_F}$ averaged over

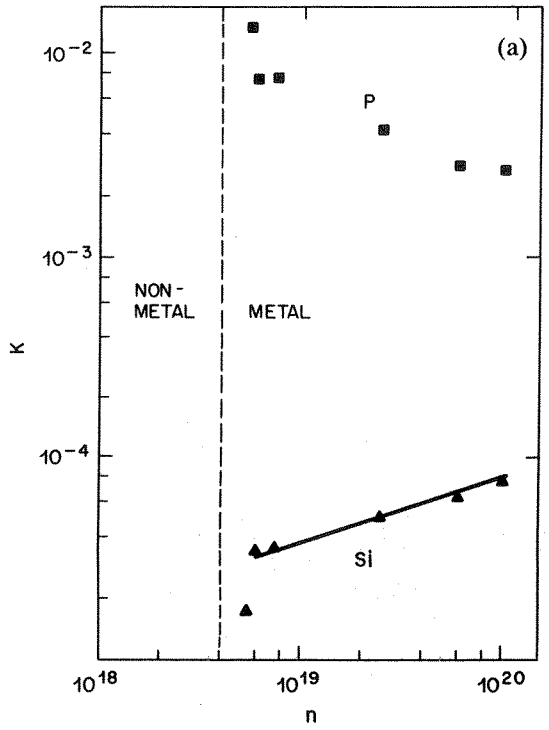

Fig. 23. Knight shifts as a function of donor concentration (a) for ^{29}Si and ^{31}P in Si : P samples (Sasaki et al. 1974) and (b) for ^{73}Ge in Ge : As samples (Deshmukh and Tunstall 1976).

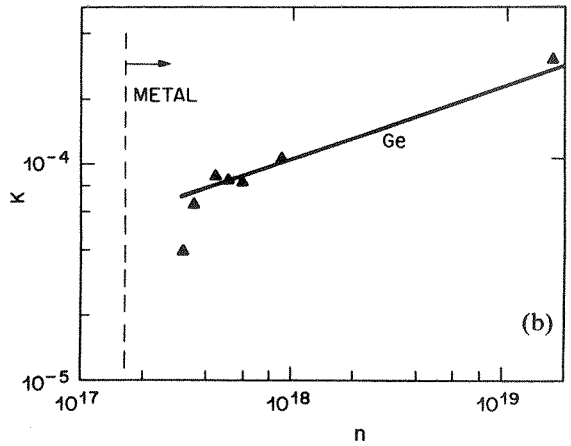

Fig. 23b.

electron states at the Fermi level, as shown below:

$$K = (8\pi/3)\Omega\langle|\psi(0)|^2\rangle_{E_F}\chi_s. \tag{4.1}$$

Here Ω is the atomic volume. Figure 23 shows the donor density dependence of the Knight shifts for ^{29}Si and ^{31}P for Si : P samples (Sasaki et al. 1974) and ^{73}Ge for Ge : As samples (Deshmukh and Tunstall 1976). Both the Si and Ge hosts show an n dependence of roughly $K \propto n^{1/3}$ (solid line), which is basically the dependence of χ_s for these densities. Presumably the donor wave function at the host site changes very little until just above n_c. However, the Knight shift for ^{31}P increases dramatically as $n \to n_c$, which must mean that the wave function at the donor site grows rapidly as n approaches n_c from above. The results for the hosts are consistent with the T_1 measurements. K and T_1 are related by the Korringa relation, $T_1Tk^2 =$ constant (see for example Holcomb 1978) and in this region $K \propto n^{1/3}$, $T_1 \propto n^{-2/3}$ and $T_1 \propto 1/T$.

5. *The specific heat*

Mott and Davis (Mott 1972, 1981b, Mott and Davis 1975) have shown that $\sigma(0)$ can be written as

$$\sigma(0) = A\left\langle\left|\int\psi_a\frac{\partial}{\partial x}\psi_b\,dx^3\right|^2\right\rangle[N(E_F)]^2, \tag{5.1}$$

where ψ_a and ψ_b are electron wave functions within the random potential. Here A is a constant, $N(E_F)$ is the density of states at the Fermi energy E_F and the bracketed term is a matrix element. Within the framework of Anderson localization (1958), $\langle\ \rangle$ vanishes for $n < n_c$ while the density of states remains finite.

One way to check this result is by measurements of the specific heat, c, which give indications of the way in which $N(E_F)$ changes with donor

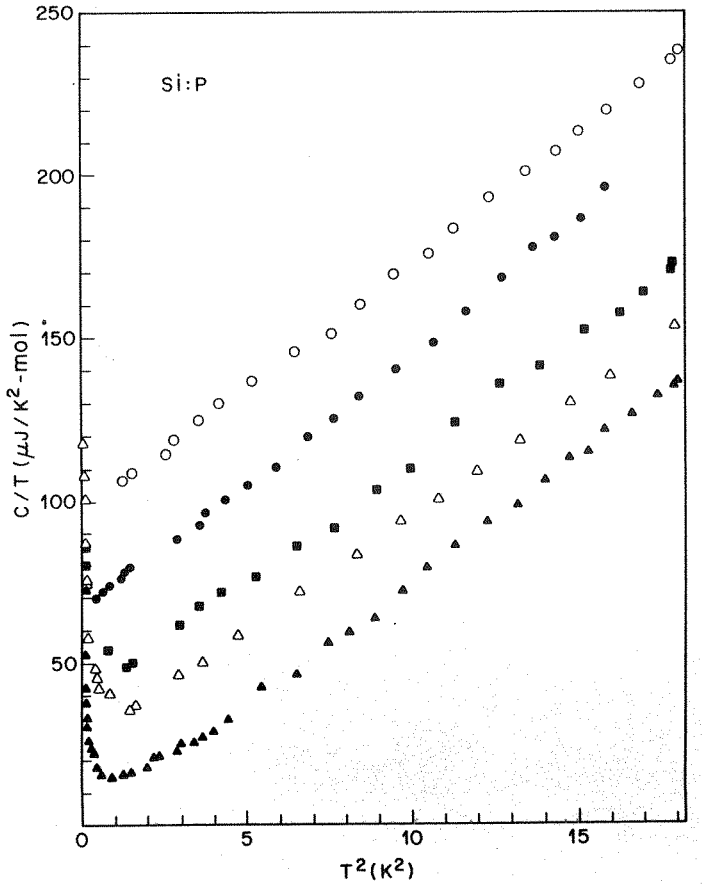

Fig. 24. The specific heat c divided by the absolute temperature is plotted against T^2 for Si : P samples above and below the transition. The intercept of a linear fit gives γ which is proportional to the density of states at the Fermi level. These results replotted in fig. 25 show that γ is continuous through the transition (Kobayashi et al. 1979).

concentration. At low temperatures we can write

$$c = \gamma T + aT^3, \tag{5.2}$$

where the T and T^3 terms are the electronic and lattice specific heats, respectively. Within the free-electron model we have (Sasaki 1980b)

$$\gamma = (\pi^2 k_B^2/3) N(E_F). \tag{5.3}$$

Kobayashi et al. (1977, 1979) and Thomas et al. (1981b) have measured the

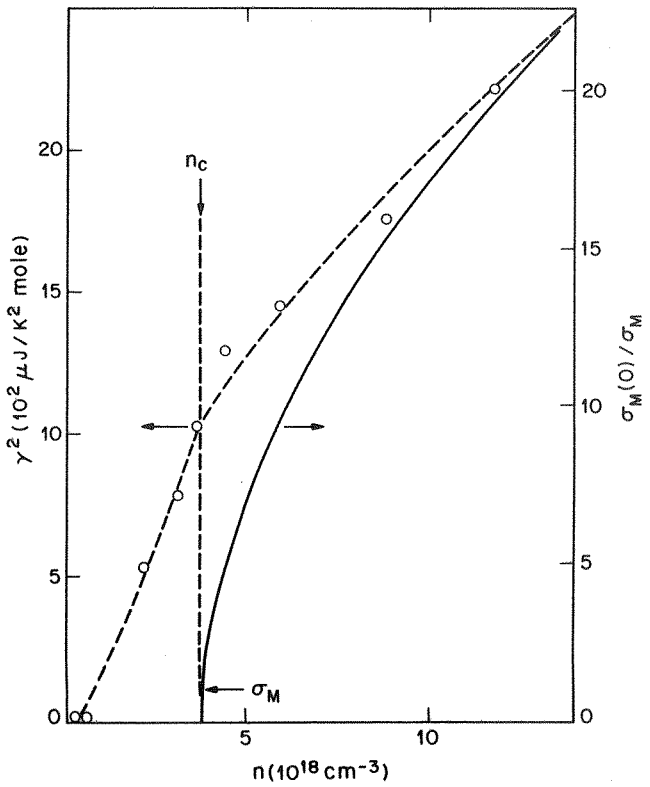

Fig. 25. The square of γ, as defined in eq. (5.2), is plotted against donor concentration for Si : P samples. The dashed line for $n > n_c$ is based on the free electron form, while for $n < n_c$ it is a guide to the eye. The solid line shows the behavior of $\sigma(0)/\sigma_M$ versus n in the metallic region (Thomas et al. 1981b).

low temperature specific heat of Si : P. In fig. 24 we show c/T plotted against T^2 for several samples with different values of n. The intercepts give γ while the slopes yield a. The slopes are essentially all the same since the lattice specific heat is primarily determined by the host material. The key feature of the data is that γ is continuous through the MI transition. The results obtained by Thomas et al. are in essential agreement with those of Kobayashi et al., but yield values of γ which are slightly smaller. In addition, the values of n used by Thomas et al. are from the neutron activation analysis of Mousty et al. (1974). They define $n_c = 3.74 \times 10^{18} cm^{-3}$, at the point where $\sigma(0) \to 0$. Kobayashi et al., however, use $n_c = 3.2 \times 10^{18} cm^{-3}$, determined from the concentration where the activation energy $E_2 \to 0$ (Yamanouchi et al. 1967). Recalibration brings the two sets of data into agreement on n_c.

Comparing eqs. (5.1) and (5.3), we see that the theory can be checked by comparing γ^2 and $\sigma(0)$ as functions of n. Figure 25 shows the results for the data of Thomas et al. (1981b). The solid curve shows the behavior of $\sigma(0)$ as found in this and earlier work (Yamanouchi et al. 1967, Rosenbaum et al. 1980b, Ootuka et al. 1980). The dashed line for $n > n_c$ is a fit to the free electron form, $\gamma = \gamma_0 (n/n_c)^{1/3}$ (Sullivan and Seidel 1968). For $n < n_c$ the values of γ^2 drop off more rapidly with n as expected for insulators. The dashed line in this region is a guide to the eye.

These results confirm the arguments of Mott and Anderson. $N(E_F)$ does not drop to zero at the transition. Rather, it varies slowly with n, in marked contrast to the abrupt disappearance of $\sigma(0)$ for $n < n_c$.

6. *Divergence of the dielectric susceptibility*

The critical behavior of the donor dielectric constant, ϵ, as $n \to n_c$ from the insulating side provides great insight into the physics of the MI transition. Castner and co-workers (Castner et al. 1975, 1980, Castner 1980, Bethin et al. 1974) have measured ϵ for n-Ge and n-Si using a low frequency capacitance bridge technique and temperatures down to 11 mK. Figure 26 shows the results for Sb, P and As doped Si samples as a function of $n \equiv N_D - N_A$. Here ϵ_X is the dielectric constant of the impurity found by subtracting the host value $\epsilon_{Si} = 11.4$ from the measured total dielectric constant. As $n \to n_c$ there is clearly a steep rise in ϵ_X.

In terms of the donor polarizability, α_D, the Clausius–Mosotti expression gives ϵ_X as

$$\epsilon_X = (1 + 8\pi n \alpha_D/3)/(1 - 4\pi n \alpha_D/3). \tag{6.1}$$

A polarization catastrophe occurs when $4\pi n \alpha_D/3 = 1$, which within the effective mass approximation occurs at a critical concentration n_c given by

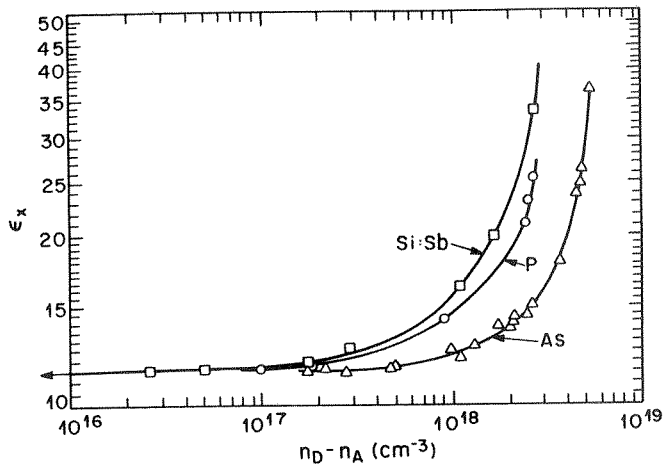

Fig. 26. The donor dielectric constant ϵ_χ, for Sb, P and As impurities in Si is plotted against the net donor concentration. ϵ_χ diverges as n_c is approached (Castner 1980).

$n_c^{1/3}a^* = 0.376$ (see for example Castner 1980, Fritzsche 1978). Castner et al. (1975) find, however, that ϵ_χ diverges even more rapidly than predicted by the Clausius–Mosotti relationship eq. (6.1).

Another means of obtaining ϵ and the dielectric susceptibility, χ, has been by integrating the far IR absorption coefficient $\alpha(\omega)$ over ω and using the Kramers–Kronig relation (Townsend 1978, Capizzi et al. 1980). The susceptibility is related to the dielectric constant, ϵ, by $\epsilon = \epsilon_{Si} + 4\pi\chi$. In terms of the zero frequency refractive index $n(0) = \epsilon^{1/2}$ we get

$$4\pi\chi = [n(0)]^2 - \epsilon_{Si}, \tag{6.2}$$

where $n(0)$ is determined by the Kramers–Kronig relation

$$n(0) = n_{Si} + c/\pi \int_0^{\omega_\infty} [\alpha(\omega)/\omega^2]\, d\omega. \tag{6.3}$$

Here c is the speed of light. Figure 27 shows the results (open circles) for Si : P (Capazzi et al. 1980). The spectrometer used for this experiment measured $\alpha(\omega)$ in the energy range $\hbar\omega = 2.5$ to 58 meV. To perform the integration indicated in eq. (6.3), a high frequency Drude tail ($\sim 1/\omega^2$) was added to the data. χ was insensitive to the precise form of the tail.

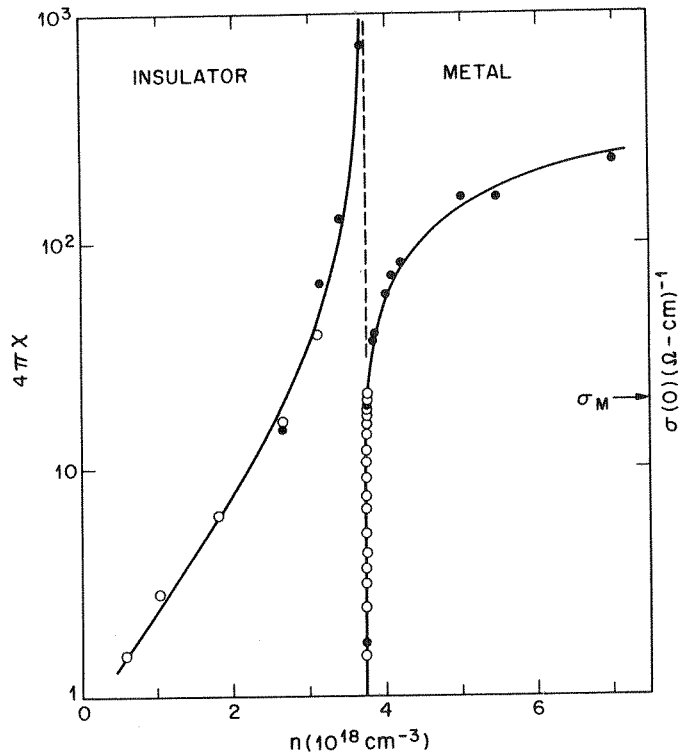

Fig. 27. Divergence of the donor dielectric susceptibility, $4\pi\chi$, and conductivity $\sigma(0)$ extrapolated to $T = 0$ as the MI transition is approached from below and above, respectively. The system is Si : P. The open circles for $4\pi\chi$ are from IR studies while the closed circles are from direct measurements at 400 MHz. The open circles for $\sigma(0)$ were obtained for samples under stress. The solid lines for $4\pi\chi$ and $\sigma(0)$ are best fits using eq. (6.4) and eq. (6.5) respectively (Hess et al. 1982).

In addition to the IR data, the closed circles show direct measurements of χ using a resonant transmission cavity at 400 MHz and temperatures down to 20 mK (Hess et al. 1982). These additional measurements were required because the IR measurements are difficult to make near the MI transition due to the nearly metallic reflectivity of the sample and the appreciable absorption which occurs at frequencies below the measured range.

The data is found to fit the form

$$4\pi\chi = \chi_0 (n_c/n - 1)^{-\xi}, \tag{6.4}$$

with $\chi_0 = 7.0$ and $\xi = 1.15$. Similarly the zero-temperature conductivity $\sigma(0)$ fits the form

$$\sigma(0) = \sigma_0(n/n_c - 1)^{\nu}, \tag{6.5}$$

with

$$\sigma_0 = 260(\Omega\,\mathrm{cm})^{-1} \quad \text{and} \quad \nu = 0.55.$$

Both forms are suggested by scaling theories of localization (Abrahams et al. 1979, Imry 1980) which predict a divergent length scale as $n \to n_c$ from above or below. Most of the theoretical and experimental work has concentrated on the approach to n_c from above, which we will now present in Part II.

Part II – "Tuning" the MI transition

As we have seen in Part I, the MI transition can be probed through many different properties – optical, transport, magnetic, dielectric, etc. – each of which give somewhat different insights into the details of the transition. Perhaps the most illuminating is the dc conductivity, as suggested by the name: Metal–Insulator Transition.

7. *The conductivity story*

The metallic phase is characterized by a finite dc conductivity at $T = 0$, $\sigma(0)$, while $\sigma(0)$ is zero in the insulator. Thus $\sigma(0)$ is in some ways akin to an "order parameter" characterizing a phase transition. (It should be noted that there is a transition only at absolute zero, somewhat like a ferromagnetic transition which occurs strictly only in zero magnetic field.) Deep in the metallic phase, $\sigma(0)$ is given by the Boltzmann transport theory:

$$\sigma(0) = \sigma_B = \frac{e^2 S_F l}{12\pi^3 \hbar}, \tag{7.1}$$

where S_F is the Fermi surface area, and l the (elastic) mean free path. Equation (7.1) is just the familiar form $\sigma_B = ne^2\tau/m$ generalized to arbitrary Fermi surfaces. For a spherical Fermi surface, $S_F = 4\pi(3\pi^2 n)^{2/3}$, where n is the electron density, so defining $d \equiv n^{-1/3}$, an average distance between

electrons, we get

$$\sigma(0) = \left(\frac{\pi}{3}\right)^{1/3} \frac{e^2}{\pi\hbar d} \frac{l}{d}. \tag{7.2}$$

Ioffe and Regel (1960) pointed out that for a Boltzmann theory to be valid, l must be greater than the distance between the scattering centres, thus eq. (7.2) has a lower limit

$$\sigma_{IR} \sim e^2/3\hbar d. \tag{7.3}$$

This idea was formalized into the concept of a "minimum metallic conductivity", σ_M, by Mott in 1972. Mott argued that as the MI transition was approached from the metallic side, $\sigma(0)$ decreased with l, till $l \approx d$ and then the system became insulating with an abrupt drop in $\sigma(0)$, from a minimum value σ_M to zero (fig. 28a). σ_M was given by:

$$\sigma_M = C(e^2/\hbar d), \tag{7.4}$$

with C a numerical factor estimated by Mott to be ~ 0.025–0.1, quite a bit lower than the constant in σ_{IR} (eq. 7.3). This, Mott claimed, was due to a reduction in the density of states at the Fermi level, compared to free electron estimates. Experimental support for this value came from the prefactor of the temperature-activated conductivity in the insulating phase (Fritzsche 1978), where $\sigma(T) \approx \sigma_M \exp(-\Delta/2T)$, with the mobility gap Δ varying with n (see section 3.1).

For many years, Mott's claim was not seriously challenged except by Cohen and co-workers (Webman et al. 1975) who viewed the metal–insulator transition as classical percolation of metallic regions in an insulating medium,

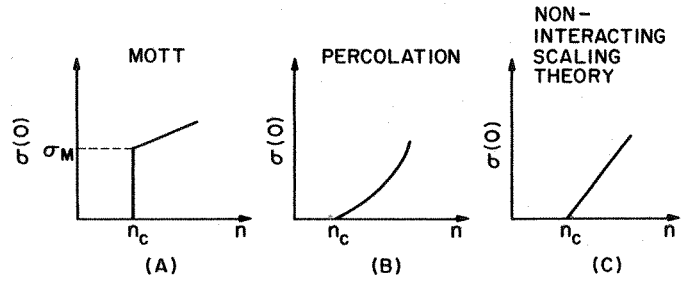

Fig. 28. Various proposed scenarios for the conductivity onset: (a) Mott, (b) Classical Percolation and (c) Scaling theory.

instead of an intrinsically quantum phenomenon. Percolation theory implied a soft onset $\sigma(0) \propto (n - n_c)^t$ with $t \approx 1.8$, (fig. 28b), and experimental support for this picture was scant.

In 1979, Abrahams, Anderson, Licciardello and Ramakrishnan, following the work of Wegner (1976) and Thouless (1977), put forth a scaling theory of localization for non-interacting electrons in disordered systems, which put the MI transition in the context of the general theory of phase transitions. Their picture was based on the assumption that conductance was the only relevant parameter describing the phenomenon of localization close to the transition. Applied to three-dimensional systems, their results implied that the dc conductivity near the MI transition would be given by a form like eq. (7.4) but with the interparticle spacing d replaced by the localization length ξ, which is the only relevant length near n_c. Thus $\sigma(0) = 0$ for $n < n_c$, while above n_c,

$$\sigma(0) = C'(e^2/\hbar\xi). \tag{7.5}$$

The localization length diverges at the transition, $\xi \sim |n - n_c|^{-\nu}$, so that $\sigma(0)$ has a continuous, albeit critical, onset (see fig. 28c):

$$\sigma(0) = \sigma_0(n/n_c - 1)^{\nu}, \tag{7.6}$$

where $\sigma_0 \sim \sigma_M$ within a factor of two or so. ν was estimated to be 1.

On the insulating side, the approach to the critical point (see section 6) is characterized by a diverging polarizability (dielectric constant). For large wavevectors q (short distances), the system behaves like a metal, so $\epsilon(q) \sim 1/q^2$. This behavior is cut off at $q_c \sim 1/\xi$ so the dielectric constant should go as

$$\epsilon(0) \sim 1/q_c^2 \sim \xi^2 \sim (n_c/n - 1)^{-2\nu}. \tag{7.7}$$

The experimental situation prior to 1982 is typified by the results of Si : P (Rosenbaum et al. 1980a) shown in fig 29. The data do not exhibit any discontinuous change in $\sigma(0)$, and can be fit over the range $1 < n/n_c < 1.5$ (and surprisingly, even further), with a critical form (eq. 7.7). However, it should be noted that the jump implied by Mott (σ_M) is extremely small, and data below σ_M are scant with uncertainties in n of the order of the variation of n. Thus they are not invulnerable to the criticism that density inhomogeneities may be playing a major role. Further, a fit to the critical form gave $\nu = 0.55 \pm 0.1$ and $\sigma_0 \simeq 13\sigma_M$, both in apparent disagreement with the scaling theory. Parallel work on amorphous Au–Ge alloys (Dodson et al. 1981) also yielded samples with $\sigma(0) < \sigma_M$, but problems of calibration of the density and of inhomogeneities in that system were no better, perhaps worse. Clearly a much finer search around n_c was called for to settle the issue of the existence of σ_M. Other nagging questions included why the critical region ($\sigma < \sigma_0$ or σ_M) was so small,

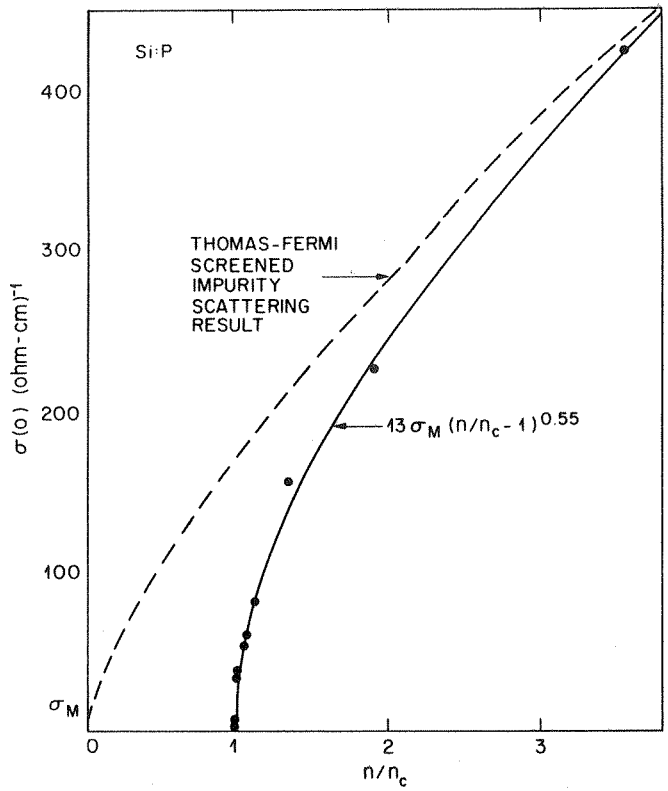

Fig. 29. Low temperature (mK) dc conductivity $\sigma(0)$ of phosphorus doped silicon, as a function of phosphorus concentration. $\sigma_M \simeq 20\ (\Omega\,\mathrm{cm})^{-1}$, $n_c \simeq 3.7 \times 10^{18}\,\mathrm{cm}^{-3}$ (data from Rosenbaum et al. 1980a).

and what role electron–electron interactions played if a scaling description was valid.

8. *Tuning the metal–insulator transition*

Around the time the experimental investigations were at this stalemate, it was realized that there existed, at least for the doped semiconductor system, an alternate probe which allows a study of the critical region with greater control and precision and less ambiguity. It was a classic case of the mountain coming to Muhammad since Muhammad couldn't go to it (i.e., varying n_c through n,

the density of a given sample, since the converse was not fruitful). This unique versatility of "tuning" the MI transition [by a factor of two (Bhatt 1981) or even more] by application of modest stresses readily available in the laboratory, is a consequence of the strong coupling of the impurity band to non-cubic deformations. In order to understand the phenomenon, however, we review some basics concerning the impurity band.

8.1. *Shallow impurities in semiconductors*

When a group V element (donor, e.g., P, As, Sb) is substituted for the host element in crystalline Si or Ge, four of the valence electrons go into the chemical bonding and the fifth sees a singly charged ion screened by the large dielectric constant ϵ of the semiconductor ($\epsilon_{Si} = 11.4$, $\epsilon_{Ge} = 15.4$). Neglecting the short range part of the potential (due to atomic shells and lattice relaxation) one obtains the effective mass approximation (ema) (Kohn 1957) in which the electronic wavefunction is expressed as:

$$\Psi(r) = \phi_{k_0}(r)\chi(r), \tag{8.1}$$

i.e., a product of the Bloch wave ϕ_k at the conduction band minimum ($k = k_0$) of the host semiconductor and an envelope χ which satisfies a hydrogenic wave equation:

$$\left(-\frac{\hbar^2}{2m^*}\nabla^2 - \frac{e^2}{\epsilon r}\right)\chi(r) = E\chi(r). \tag{8.2}$$

Here m^* is the conduction band mass (in general a tensor, with different elements along its major and minor principal axes). When m^* is a scalar (e.g., for a band minimum at $k_0 = 0$ as in GaAs), eq. (8.2) is the hydrogen atom wave equation with a 1s ground state and the usual set of excited bound states (2s, 2p, 3s etc.) below the continuum; the "Bohr radius" $a^* = \epsilon\hbar^2/m^*e^2$ is typically 30–100 times the normal hydrogen atom, while the "Rydberg" $\mathrm{Ry} = m^*e^4/2\epsilon^2\hbar^2$ is 300–1000 times smaller than that of hydrogen. For Si and Ge, the conduction band masses are anisotropic, and the 1s ground state wavefunction is no longer the spherically symmetric exponential form $\exp(-r/a^*)$. However, it is well represented by the variational form (Kohn and Luttinger 1955, Kittel and Mitchell 1955):

$$\chi(r) \sim \exp\left[-\left(\frac{x_\ell^2}{b^2} + \frac{x_{t1}^2 + x_{t2}^2}{a^2}\right)^{1/2}\right], \tag{8.3}$$

where x_ℓ and x_{ti} ($i = 1, 2$) are the components of r along the major and minor axes of the mass tensor and a and b are variational parameters often referred

to as the transverse and longitudinal radii of the wavefunction. The effective mass radii in Si are $a \approx 24$ Å, $b \approx 14$ Å and in Ge $a \approx 62$ Å, $b \approx 22$ Å (note the large anisotropy). Since the donor wavefunction radii are much larger than the nearest neighbor distance of the host lattice (~ 2.4 Å), the discreteness of the positions for a substitutional impurity is unimportant, and we may assume that the impurities are randomly distributed in a continuum. This property will come in use later.

Both Si and Ge really have multiple conduction band minima (called valleys): in Si there are six equivalent minima, along the positive and negative x, y and z axes [(100) directions], roughly 85% the way to the Brillouin zone boundary; Ge has four minima in the (111) directions at the zone boundary. For multiple minima (valleys), there is a hydrogenic set of impurity levels derived from each, which are orthogonal to each other because of the orthogonality of the Bloch waves ϕ_{k_i} (eq. 8.1). The envelope wavefunctions are all of the form (8.3), only the axes (x_ℓ and x_t) are rotated along the axes of the appropriate valley – i.e., the anisotropic wavefunctions "point" in different directions. Thus, within the effective mass approximation, for a v-fold degenerate conduction band minimum there is a v-fold degenerate ground state. For most donors, however, this degeneracy is split by the short range (atomic/lattice deformation) potentials of the impurity known as the central cell, and the ground state is usually the non-degenerate, symmetric combination of the v valleys:

$$\Psi_{\mathrm{sym}} = \sum_{i=1,v} \chi_i(r)\phi_{k_i}(r). \tag{8.4}$$

[Ge : Sb and (interstitial) Si : Li however, have unusually small splittings and are well approximated by degenerate em wavefunctions for most purposes.] Usually the lowering of the energy of the symmetric state is accompanied by a "sucking in" of the envelope function by the central cell. The envelope functions for the other states (which consist of a triply degenerate and a doubly degenerate set in Si, and one triply degenerate set in Ge), however, are not much affected because these "antisymmetric" states have zero amplitude at the impurity site. Thus the effective radii of the ground state wavefunction are smaller, sometimes by as much as 30%, than the excited 1s states.

Non-cubic deformations break the degeneracy of the conduction band minima in these cubic semiconductors and thus couple directly to the donor wavefunction and hence the critical density n_c. The relatively modest external stresses needed are easily understood because the relevant energy scale is the impurity band energy $\sim 10^{-2}$eV rather than $\sim$ eV typical of most electronic bands. We may obtain an order of magnitude estimate of the stress needed as

$$S \sim CE_0/\Xi \tag{8.5}$$

where C is an elastic constant ($\sim 10^{12}$ dyne cm^{-2}), E_0 the impurity band energy ($\sim 10^{-2}$ eV) and Ξ the conduction band deformation potential (typically ~ 10 eV). Putting these numbers in eq. (8.5) yields $S \sim 10^{9}$ dyne cm^{-2} $\sim$ 1 kbar, a stress easily obtainable in the laboratory.

8.2. *Effective mass donors*

The coupling of n_c with uniaxial stress is largest for the effective mass case, where the strain splits the degeneracy of the donor ground state. Thus the donor wavefunction, which has the freedom of choosing between a number of differently oriented ground states (see fig. 30a) in the unstressed case, is forced into the lowest one of them in the high stress limit (fig. 30b). (In the case of Si, the high stress limit has two lowest valleys related by inversion symmetry; however the envelope functions are identical, so there is no freedom of orientation left.) In the absence of stress, the impurity band is thus formed by a random, isotropic system of donors with nearly isotropic wavefunction overlap. (This is because of the orientational degree of freedom. Strictly speaking, there is only cubic symmetry – isotropy would require a continuous, infinitely degenerate orientational degree of freedom.) In the high stress limit, on the other hand, the donor system, though random, has highly anisotropic wavefunctions (fig. 30c). However, it was noted (Bhatt 1981) that by a scale contraction along the transverse axes (or expansion along the longitudinal axis) of the wavefunction (eq. 8.3), which makes the wavefunction isotropic in the new coordinates (fig. 30d), the system can be mapped on to an essentially isotropic system, with a different density. This is because: (a) the randomness of the distribution of the donor nuclei is retained *with no preferred direction* in the transformed frame, despite the anisotropic scaling; and (b) the dominant (exponential) dependence of quantities determining n_c e.g., hopping integrals, is isotropic; the residual anisotropy in prefactors of the exponential has only a small effect in determining n_c.

By comparing the Mott criteria (eq. 1.1) for the two almost isotropic systems – the stress free case, and the high stress limit in the transformed frame, one may obtain the ratio of the critical density in the two cases, independent of the value of the constant on the right side of eq. (1.1). For Ge and Si, the results thus obtained are:

$$\text{Ge:}\quad \frac{[n_c]^{1\text{ valley}}}{[n_c]^{4\text{ valleys}}} \approx 2.2; \qquad \text{Si:}\quad \frac{[n_c]^{2\text{ valleys}}}{[n_c]^{6\text{ valleys}}} \approx 1.5. \tag{8.6}$$

The former result compares well with the experimentally obtained value of 2.1 in Ge:Sb by Cuevas and Fritzsche (1965), while in Si the comparison is yet to be done – the high diffusion rate of the effective mass donor Li in Si at room temperatures destroys the quenched nature of the system and complicates the issue.

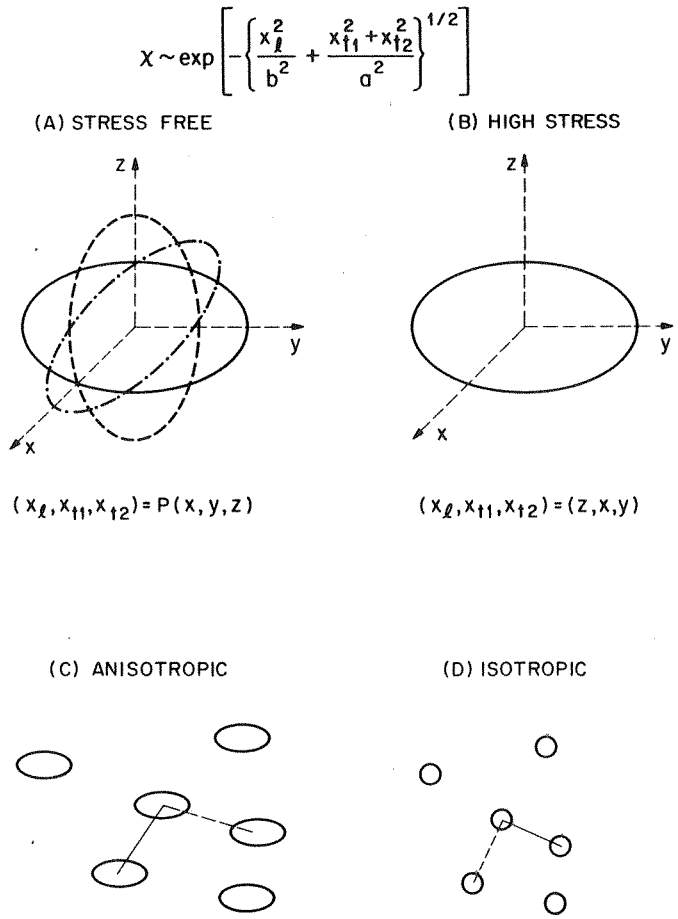

Fig. 30. (a) The three possible orientations of the spheroidal effective mass donor wavefunctions in Si at zero stress are suppressed to a single orientation at high stress (b). Mapping of the anisotropic random system in the high stress limit (c) into an isotropic random system (d). It should be noted that the distribution in (d) remains isotropic, though nearest (solid line) and next-nearest (dashed line) neighbors may get interchanged.

8.3. *Moderate central cells*

In the case of most donors (e.g., Si : P, on which detailed measurements have been done), the orientational degeneracy is lifted even at zero stress by the

central cell potential and the ground state wavefunction is the symmetric combination eq. (8.4). Further, interference between the different Bloch waves leads to rapidly oscillating (on the scale of the Si lattice spacing) factors (Andres et al. 1981, Bhatt 1982), which make a simple application of the Mott criterion uncertain, and a detailed calculation of the one-electron bands (fig. 1) has to be done. Such a calculation for the case of the Mott transition has been done recently (Bhatt 1982), using parameters appropriate for Si : P. The change in n_c with stress is, in this case, related to the mixing of the excited 1s states with the symmetric ground state by the non-cubic deformation. Since the excited states are more extended, this causes a broadening of the occupied (hole) band, as well as a lowering of the empty (electron) band. Both these effects lead to a reduction in n_c, which is signalled by a crossing of the two bands. Note that this is the *opposite* of the effective mass case where n_c at high stresses is larger than at zero stress.

The first ingredient of the calculation is setting up the pseudopotential for an electron in the D^- band (equivalent to H^-), scattering off a lattice of neutral hydrogenic donors. This is done using the method of polarized orbitals, which views it in a sort of Born–Oppenheimer approach – scattering of a slow "outer" electron by a positive point charge surrounded by a fast moving "inner" electron which adjusts adiabatically to the slow outer one. The potential is then obtained as a sum of the electrostatic (Hartree) interaction of the outer electron with the neutral donor $U_0(r)$ and a perturbation series arising from the modification of the wavefunction of the occupied 1s state (virtual excitation to higher levels) due to the outer electron. For example, the leading (second order) term of the perturbation series is

$$U_2(r) = -\left\langle \chi(r)\left|V(r,r')\frac{1}{\mathscr{H}-E_g}V(r,r')\right|\chi(r')\right\rangle, \tag{8.7}$$

where E_g and $\chi(r')$ are the ground state energy and wavefunction of the internal electron, $\mathscr{H}$ the Hamiltonian of the system, and

$$V(r,r') = -\frac{e^2}{\epsilon r} + \frac{e^2}{\epsilon|r-r'|} \tag{8.8}$$

is the perturbation due to the outer electron. It can be shown (Bhatt and Rice 1981) that truncating the perturbation series at this term leads to a potential which reproduces the correct asymptotic behavior both for $r \to 0$ and $r \to \infty$, and for an isolated H^- yields a binding energy within 0.01 Ry of the exact result, which is extremely good for a one-electron description. Thus this method provides a correct description in the low density limit, and complements methods such as the local density functional which are expected to be good at high densities. Further, the method is readily generalized to the case

where a central cell is present. A simplified cellular method – approximating the Wigner–Seitz cell by a sphere of equal volume – is used to locate the band edges. (The top of the band corresponds to a zero of the wavefunction at the sphere boundary, while the bottom is given by a zero derivative boundary condition.) The D^+ (hole) band is calculated using a tight-binding approximation for the occupied ground state orbitals. For the D^+ band there is a large reduction in the bandwidth, by a factor of almost three, compared to the hydrogenic case because of phase mismatch of the different Bloch waves in the ground state wavefunction at different donor sites. In addition, spin-flip scattering for an antiferromagnetic ground state as exists in the insulating state, leads to a further 25% reduction in the bandwidth. When all these effects are included, the one-electron bands calculated for Si : P are shown in fig. 31, for various values of compressional stress along a (110) direction. The calculated change in n_c is about 3% for $S \approx 2$ kbar, which is consistent with recent experimental estimates of 10–15% shift in n_c for 10 kbar stress (Paalanen et al.

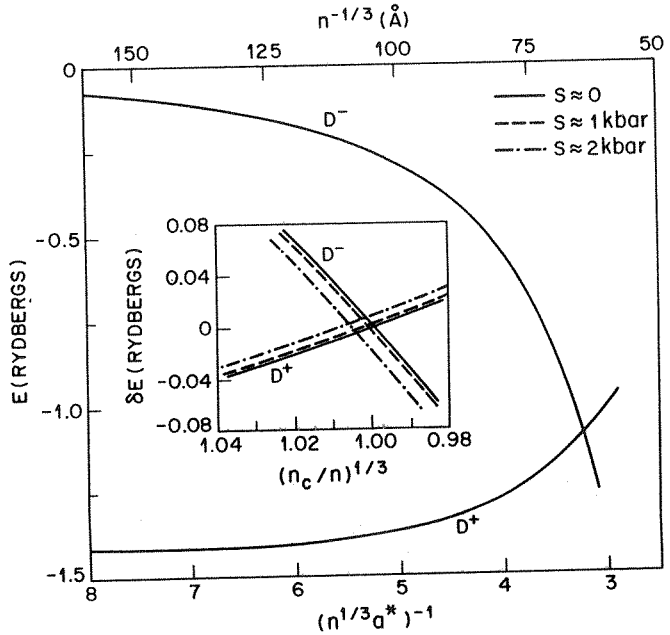

Fig. 31. The bottom of the electron (D^-) and top of the hole (D^+) band for Si : P as a function of donor separation $d \equiv n^{-1/3}$. Region near n_c is enlarged in the inset, for different values of stress S (Bhatt 1982).

1982). As stated before, n_c is reduced by application of stress, and thus an insulator near the MI transition should transform into a metal.

9. *Tuning n_c in practice: Squeezing at low temperature*

The ideas outlined in the previous section were implemented in the experiment by Paalanen et al. (1982) in which uncompensated samples of Si : P with dopant densities just below n_c were subjected to uniaxial stress S of up to 10 kbar at millikelvin temperatures. The necessity of low temperature is clear – the MI transition strictly takes place only at $T = 0$. One might naively guess that experimentally it would be sufficient to have $T \ll \Delta$, the one-electron mobility gap (fig. 31); for samples 1% away from n_c, Δ, is about 3 K. However, as shown by Altshuler et al. (1979a, b, 1980a), electron–electron interactions in disordered Fermi liquids with diffusive propagators lead to strong singularities near the Fermi level even for weak disorder. For example, the density of states in three dimensions has a singularity of the form

$$N(E) \sim N(0)\left[1 + (E/E_0)^{1/2}\right], \tag{9.1}$$

which has been seen in tunnelling experiments (Dynes and Garno 1981, McMillan and Mochel 1981), while the temperature dependence of the dc conductivity is given by

$$\sigma(T) = \sigma(0)\left[1 + (T/T_0)^{1/2}\right] \tag{9.2}$$

which has been verified in a number of systems, including Si : P (Rosenbaum et al. 1981a, see section 3.3). The temperature T_0 is given by

$$T_0 \sim E_F\left(\sigma_{IR}/\sigma(0)\right)^3, \tag{9.3}$$

where E_F is the Fermi energy (~ 100 K for Si : P near n_c), σ_{IR} the Ioffe–Regel value (eq. 7.3) and $\sigma(0)$ the zero temperature conductivity. For $\sigma(0) \sim \sigma_M$ (n within 1% of n_c), $T_0 \sim 100$ mK, so that for the temperature-dependent correction in eq. (9.2) to be small, we need $T \ll 100$ mK. This emphasizes the necessity of millikelvin temperatures (and proper extrapolation to absolute zero), which is abundantly clear in the experimental results in the stress dependence of the conductivity in Si : P taken at successively lower temperatures (fig. 32).

The measurements were made on samples roughly a centimetre long and a fraction of a millimetre in the other directions. Originally two sections from a single wafer were used, but measurements have since been extended to samples from three different boules with slightly different densities. Samples were mounted in a pressure device operated by liquid ^{4}He as illustrated in the inset

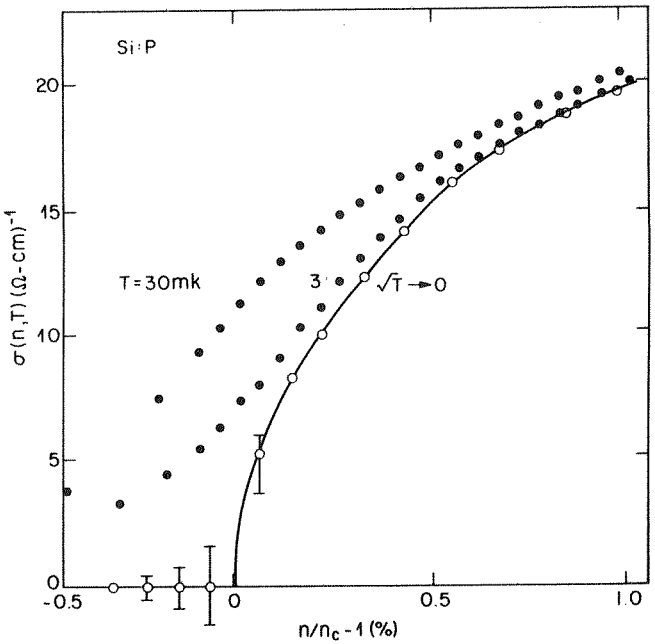

Fig. 32. Dc conductivity of a sample of Si : P near n_c as a function of uniaxial stress at three different temperatures (Thomas et al. 1983).

of fig. 34, and the stress measured capacitively at one end of the sample. Measurements were done using both a four-probe and a capacitance bridge (for dielectric constant in the insulating phase) technique. Low frequencies (10–10^3 Hz) and power levels below 10^{-15} W were used to ensure linear response.

In a small interval around a finite critical stress S_c (when $n_c = n$), the variation of n_c with S may be approximated by a linear form. Thus the onset of $\sigma(0)$ with S would yield the critical exponent ν (eq. 7.6). However, due to rounding at low conductivity [$\sigma < 5\,(\Omega\,\mathrm{cm})^{-1}$], which varied with sample, the critical exponent of the intrinsic dependence is not clear from fig. 32. (The downward curvature implies that $\nu < 1$, though.) Thus an extrapolation to $T = 0$ has to be done by fitting an appropriate form such as eq. (9.2), which yields $\nu \simeq 0.5 \pm 15\%$. Alternatively, one may plot σ^2 versus S at the lowest temperature of 3 mK (fig. 33); the linear behavior above the tail immediately implies a critical exponent of 0.5 for σ.

Note that the data include zero temperature conductivities well below Mott's minimum value σ_M, and are therefore clearly inconsistent with this possibility shown in fig. 28a. Further, the exponent for the critical onset disagrees with

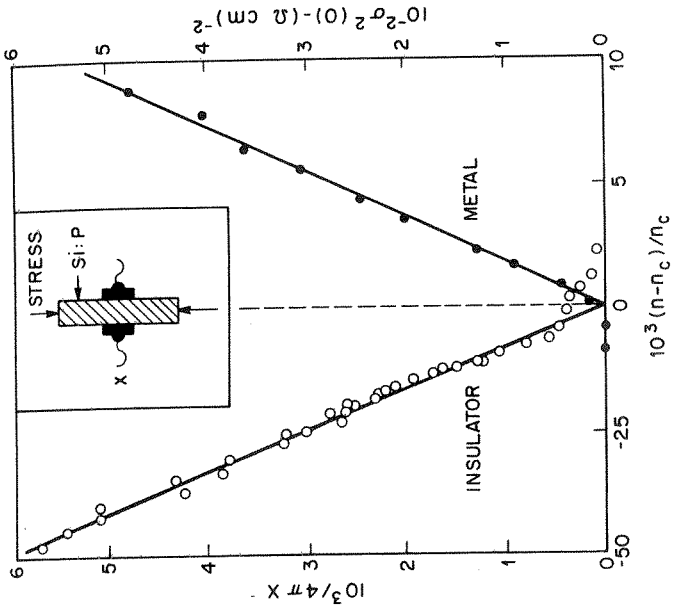

Fig. 33. Inverse donor polarizability (closed circles) and square of conductivity (open circles) as a function of stress near the critical value (Paalanen et al. 1983).

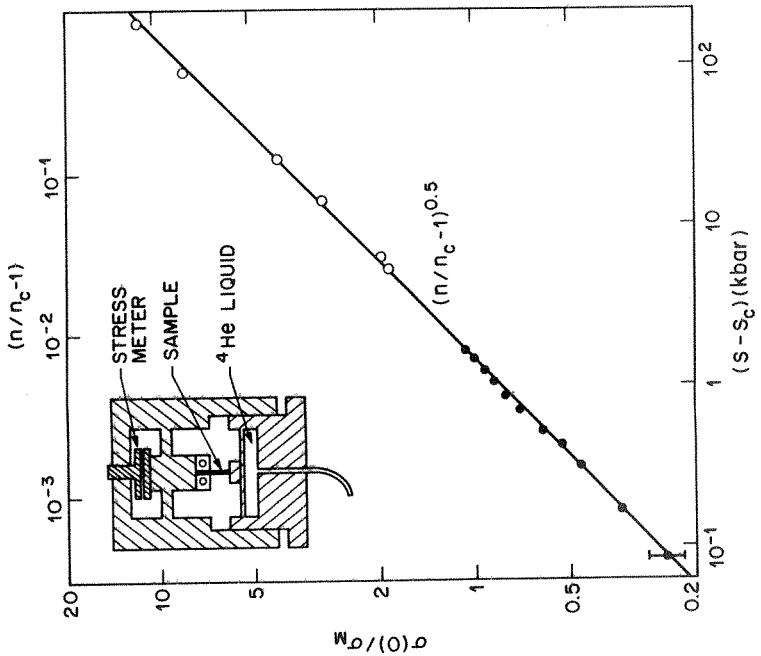

Fig. 34. Conductivity versus "distance" from transition in Si:P (density, upper scale; stress, lower scale) on a double logarithmic plot (Paalanen et al. 1982).

that predicted by the scaling theory for the non-interacting case. However, it does agree with earlier measurements (fig. 29). This is shown in fig. 34, which plots both sets of data in a double logarithmic plot covering over two orders of magnitude in conductivity – the data show no evidence for any discontinuity or strong anomaly at σ_M and add confidence in the obtained value of ν.

Finally, fig. 33 also shows the inverse donor dielectric susceptibility in the insulating phase versus S; the linear variation implies a divergence of the dielectric constant as n_c is approached from below by a critical form with an exponent $\zeta \simeq 1$. This behavior was observed in earlier experiments further from the transition (Capizzi et al. 1980), and is consistent with the simple argument given above, according to which $\zeta = 2\nu$.

10. Tuning with magnetic field

Application of a magnetic field has also been used to induce a MI transition in barely metallic samples (see Robert et al. 1980, and references therein, Mott

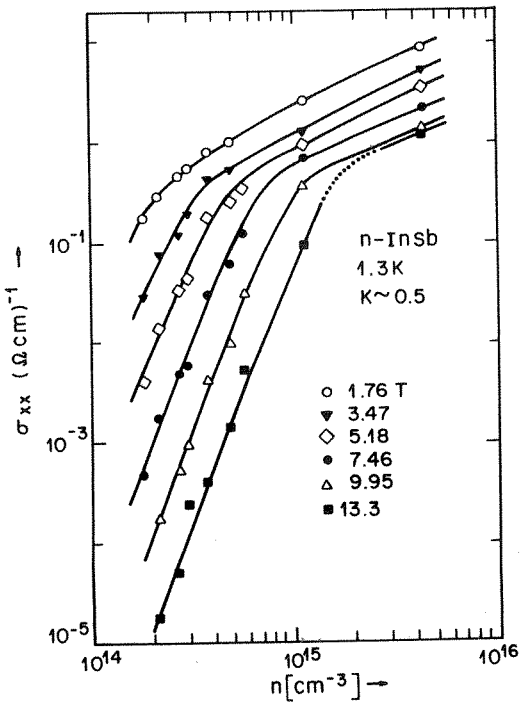

Fig. 35. The influence of applied magnetic fields on the MI transition in n-InSb samples is shown in this log–log plot of conductivity versus donor concentration. Increased field strength causes n_c, as measured by the knee of the curve, to move to higher values (Ishida and Otsuka 1977b).

1981b). The shrinking of the electron orbit due to the field may be sufficient to reduce orbit overlap to the point where the electrons become localized. In fig. 35 we see data from Ishida and Otsuka (1977b) for n-InSb. The knee in the curve is interpreted as marking the transition. The most heavily doped sample stays metallic even at a field of 13.3 T while the sample with excess donor concentration of $\sim 10^{15}\text{cm}^{-3}$ passes through the transition at ~ 10 T. They find that the condition for the MI transition to occur is $(N_D a_\perp^2 a_\parallel)^{1/3} \simeq 0.26$, where N_D is the donor concentration and $a_\perp$ and $a_\parallel$ are the Bohr orbit radii perpendicular and parallel to the field. Unfortunately, the temperature is not low enough for an evaluation of the exponent ν. [This presents a genuine problem since systems for which magnetic tuning is feasible (e.g., InSb) naturally have a small microscopic energy scale $\lesssim 1$ meV, and hence need extremely small T]. Robert et al. (1980) have performed similar experiments on n-InSb but with the addition of hydrostatic pressure. They find that a given example will become non-metallic at a lower field with pressure than it does without pressure.

11. Concluding remarks

We have seen that the variation of the MI transition density n_c with stress in n-doped semiconductors can be theoretically understood, and is a powerful, experimentally accessible tool for studying the behavior of systems near n_c. The onset of the zero temperature conductivity determined by this technique is found to be of a continuous, but critical form $\sigma \sim (n - n_c)^\nu$ with $\nu \approx 0.5$, with no discontinuity or strong anomaly at Mott's "minimum metallic conductivity" σ_M. The exponent ν does not agree with that obtained within a scaling theory for non-interacting electrons (or with the percolation exponent for conductivity). This discrepancy is supported by the divergence of the static dielectric constant in the insulator which is characterized by an exponent $\zeta \approx 1.0 \approx 2\nu$. One may speculate that this indicates the relevance of electron interaction effects. [In the opposite limit of electron correlations with no disorder, the transition is known to be first order ($\nu = 0$): the experimental results lie midway between the two extremes.] Scaling theories for interacting electrons have been put forward by McMillan (1981) and very recently by Grest and Lee (1983); the latter calculation gives $\nu = 0.6$ and $\zeta/\nu = 2\frac{1}{3}$, not far from the experimental results in Si : P.

Further, the continuity between the $\sigma < \sigma_M$ and $\sigma > \sigma_M$ results (fig. 34) suggests that the critical region is larger than σ_M, perhaps $\sim \sigma_{IR}$ (eq. 7.3). In fact, a reanalysis of the scaling picture in the weak scattering limit for many-valley systems (Bhatt and Ramakrishnan 1983) yields a conductivity scale $\sigma_0 \sim 12\sigma_M$ for Si, in excellent agreement with the prefactor seen experimentally (fig. 29). This restores the critical region from its assumed anomalous

narrowness of $\sim 10^{-2}$ to a more reasonable value. Additional evidence in favor of the larger scale is that it corresponds to the value when strong departures occur from the conductivity calculated either (a) using the Kubo formula with a Fermi surface density of states obtained from the experimental low temperature electronic specific heat (Thomas et al. 1981b), or (b) that deduced for a free electron gas with Thomas–Fermi screened impurities (fig. 29).

Very recently, precision tunnelling and transport measurements have been carried out in amorphous Nb_xSi_{1-x} alloy system (Hertel et al. 1983) as a function of Nb concentration. They too obtain a continuous onset with $\sigma(0)$ values much below σ_M, but with an exponent $\nu \simeq 1.0$, perhaps because their system is partly compensated. This larger value of the conductivity exponent had been indicated in compensated semiconductors (Thomas et al. 1982a) as

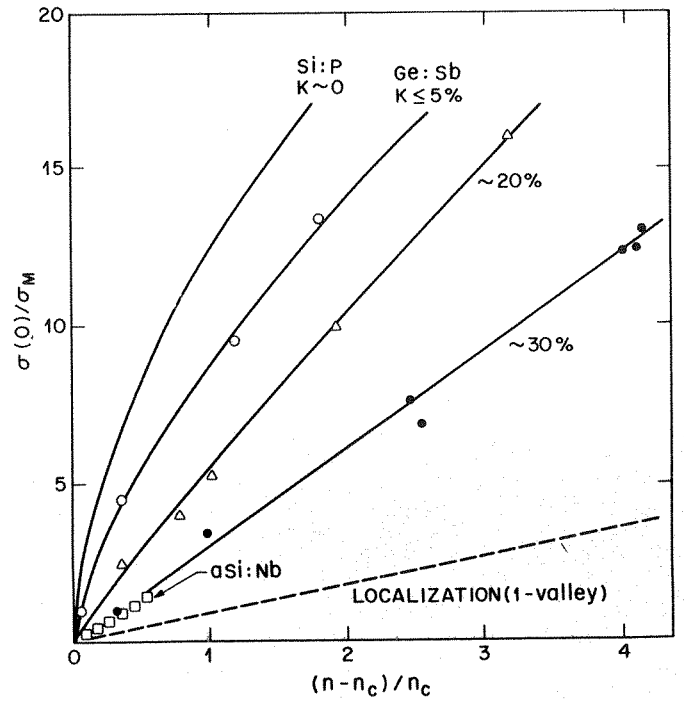

Fig. 36. Plot of $\sigma(0)/\sigma_M$ against $(n - n_c)/n_c$ for the Si : P system (see fig. 27), the Ge : Sb system with various degrees of compensation and αSi : Nb. Increasing compensation, pushes the data closer to the localization result (dashed line) (Thomas et al. 1982a).

well as in amorphous Au_xSi_{1-x} (Nishida et al. 1982), though these measurements did not probe as close to the transition (see fig. 36). It would clearly be of great interest to do stress tuning experiments in compensated semiconductors to see if there is a crossover from interaction dominated to disorder dominated behavior (as the apparent discrepancy in the exponent ν would suggest).

References

Abrahams, E., P.W. Anderson, D.C. Licciardello and T.V. Ramakrishnan, 1979, Phys. Rev. Lett. **42**, 693.
Alexander, M.H., and D.F. Holcomb, 1968, Rev. Mod. Phys. **40**, 815.
Altshuler, B.L., and A.G. Aronov, 1979a, Zh. Eksp. Teor. Fiz. **77**, 2028 [Soviet Phys. JETP **50**, 968].
Altshuler, B.L., and A.G. Aronov, 1979b, Solid State Commun. **36**, 115.
Altshuler, B.L., A.G. Aronov and P.A. Lee, 1980a, Phys. Rev. Lett. **44**, 1288.
Altshuler, B.L., D. Khmelnitskii, A.I. Larkin and P.A. Lee, 1980b, Phys. Rev. **B22**, 5142.
Anderson, P.W., 1958, Phys. Rev. **109**, 1492.
Andres, K., R.N., Bhatt, P. Goalwin, T.M. Rice and R.E. Walstedt, 1981, Phys. Rev. **B24**, 244.
Bethin, J., T.G. Castner and N.K. Lee, 1974, Solid State Commun. 14, 1321.
Bhatt, R.N., 1981, Phys. Rev **B24**, 3630.
Bhatt, R.N., 1982, Phys. Rev. **B26**, 1082.
Bhatt, R.N., and P.A. Lee, 1982, Phys. Rev. Lett. **48**, 344.
Bhatt, R.N., and T.V. Ramakrishnan, 1983, Phys. Rev. **B28**, 6091.
Bhatt, R.N., and T.M. Rice, 1979, Phys. Rev. **B220**, 466.
Bhatt, R.N., and T.M. Rice 1980, Phil. Mag. **B42**, 859.
Bhatt, R.N., and T.M. Rice, 1981, Phys. Rev **B23**, 1920.
Capizzi, M., G.A. Thomas, F. De Rosa, R.N. Bhatt and T.M. Rice, 1979, Solid State Commun. **31**, 611.
Capizzi, M., G.A. Thomas, F. De Rosa, R.N. Bhatt and T.M. Rice, 1980, Phys. Rev. Lett. **44**, 1019.
Castner, T.G., 1980, Phil Mag. **B42**, 873.
Castner, T.G., N.K. Lee, G.S. Cieloszyk and G.L. Salinger, 1975, Phys. Rev. Lett. **34**, 1627.
Castner, T.G., N.K. Lee, H.S. Tan, L. Moberly and O. Symko, 1980, J. Low Temp. Phys. **38**, 447.
Chui, T., P. Lindenfeld, W.L. McLean and K. Mui, 1981, Phys. Rev. Lett. **47**, 1617.
Cuevas, M., and H. Fritzsche, 1965, Phys. Rev. **139**, A1628; Phys. Rev. **137**, A1847.
Deshmukh, V.G.I., and D.P. Tunstall, 1976, J. Physique **C4**, 329.
Dodson, B.W., W.L. McMillan, J.M. Mochel and R.C. Dynes, 1981, Phys. Rev. Lett. **46**, 46.
Doehler, J., 1975, Phys. Rev. **B12**, 2917.
Doehler, J., P.J. Colwell and S.A. Solin, 1975, Phys. Rev. Lett. **34**, 584.
Dynes, R.C., and J.P. Garno, 1981, Phys. Rev. Lett. **46**, 137.
Edwards, P.P., and M.J. Sienko, 1978, Phys. Rev. **B17**, 2575.
Feher, G., R.C. Fletcher and E. A. Gere, 1955, Phys. Rev. **100**, 1784.
Fisher, P., and A.K. Ramdas, 1969, Physics of the Solid State, eds. S. Balakrishna, M. Krishnamurti and B. Ramachandra (Academic Press, London) p. 149.
Fletcher, R.C., W.A. Yager, G.L. Pearson and F.R. Merritt, 1954, Phys. Rev. **95**, 844.
Friedman, L.R., and D.P. Tunstall, eds., 1978, The Metal Non-Metal Transition in Disordered Systems (SUSSP Publ., Edinburgh).
Fritzsche, H., 1955, Phys. Rev. **99**, 406.
Fritzsche, H., 1978, in: The Metal Non-Metal Transition in Disordered Systems, eds. L.R. Friedman and D.P. Tunstall (SUSSP Publ., Edinburgh) p. 193.
Fritzsche, H., 1980, Phil. Mag. **B42**, 835.
Fritzsche, H., and M. Cuevas, 1960a, Phys. Rev. **119**, 1238.
Fritzsche, H., and M. Cuevas, 1960b, Proc. Intern. Conf. Semiconductors, Prague, p. 222.
Geschwind, S., R. Romestain and G.E. Devlin, 1976, J. Physique, **37**, C4, 313.
Geschwind, S., R. Romestain and G.E. Devlin, 1979, in: Proc. 14th Intern. Conf. on the Physics of Semiconductors, Edinburgh, 1978, Inst. Phys. Conf. Ser. No. 43, ed. B.L.H. Wilson (The Institute of Physics, Bristol) p. 1013.
Geschwind, S., R.E. Walstedt, R. Romestain, V. Narayanamurti, R.B. Kummer, R. Feigenblatt and G. Devlin, 1980, Phil. Mag. **B42**, 961.
Gorkov, L.P., A.I. Larkin and D. Khmelnitskii, 1979, JETP Lett. **30**, 228.
Grest, G.S., and P.A. Lee, 1983, Phys. Rev. Lett. **50**, 693.
Hertel, G., D.J. Bishop, E.G. Spencer, J.M. Rowell and R.C. Dynes, 1983, Phys. Rev. Lett. **50**, 743.
Hess, H.F., K. DeConde, T.F. Rosenbaum and G.A. Thomas, 1982, Phys. Rev. **B25**, 5578.
Hill, R.M., 1971, Phil. Mag. **24**, 1307.
Holcomb, D.F., 1970, Radiat. Eff. **2**, 15.
Holcomb, D.F., 1978, in: The Metal Non-Metal Transition in Disordered Systems, eds. L.R. Friedman and D.P. Tunstall (SUSSP Publ., Edinburgh) p. 251.
Holcomb, D.F. and J.J. Rehr, 1969, Phys. Rev. **183**, 773.
Hubbard, J., 1964a, Proc. R. Soc. London Ser. **A277**, 237.
Hubbard, J., 1964b, Proc. R. Soc. London Ser. **A281**, 401.
Imry, Y., 1980, Phys. Rev. Lett. **44**, 469.
Ioffe, A.F., and A.R. Regel, 1960, Prog. Semicond. **4**, 237.
Ishida, S., and E. Otsuka, 1977a, J. Phys. Soc. Jpn. **42**, 542.
Ishida, S., and E. Otsuka, 1977b, J. Phys. Soc. Jpn. **43**, 124.
Jain, K., S. Lai and M. Klein, 1976, Phys. Rev. **B13**, 5448.
Jerome, D., C. Ryter and J.M. Winter, 1965, Physics **2**, 81.
Kawabata, A., 1980, Solid State Commun. **34**, 431.
Kawabata, A., 1981, J. Phys. Soc. Jpn. **50**, 2461.
Kittel, C., and A.H. Mitchell, 1955, Phys. Rev. **96**, 1488.
Kobayashi, N., S. Ikehata, S. Kobayashi and W. Sasaki, 1977, Solid State Commun. **24**, 67.
Kobayashi, N., S. Ikehata, S. Kobayashi and W. Sasaki, 1979, Solid State Commun. **32**, 1174.
Kohn, W., 1957, Solid State Physics, Vol. 5, eds. F. Seitz and D. Turnbull (Academic Press, London) pp. 257–320.
Kohn, W., and J.M. Luttinger, 1955, Phys. Rev. **98**, 915.
Kummer, R.B., R.E. Walstedt, S. Geschwind, V. Narayanamurti and G.E. Devlin, 1978, Phys. Rev. Lett. **40**, 1098.
Lee, P.A., and T.V. Ramakrishnan, 1982, Phys. Rev. **B26**, 4009.
Maekawa, S., and N. Kinoshita, 1965, J. Phys. Soc. Jpn. **20**, 1447.
McMillan, W.L., 1981, Phys. Rev. **B24**, 2739.
McMillan, W.L., and J. Mochel, 1981, Phys. Rev. Lett. **46**, 556.
Mooij, J.H., 1973, Phys. Stat. Sol. (a) **17**, 521.
Morita, S., T. Fukase, Y. Isawa, S. Ishida, Y. Taksuti and N. Mikoshiba, 1980, Proc. 15th Intern. Conf. Phys. Semiconductors, Kyoto, 1980, J. Phys. Soc. Jpn. **49**, 379.
Morita, S., Y. Isawa, T. Fukase, S. Ishida, Y. Koike, Y. Taksuti and N. Mikoshiba, 1982, Phys. Rev., **B25**, 5570.
Mott, N.F., 1949, Proc. Camb. Phil. Soc. **32**, 281.
Mott, N.F., 1956, Proc. Phys. Soc. London **A62**, 416.
Mott, N.F., 1967, Adv. Phys. **16**, 49.
Mott, N.F., 1972, Phil. Mag. **26**, 1015.
Mott, N.F., 1974, Metal–Insulator Transitions (Taylor and Francis, London).
Mott, N.F., 1978, in: The Metal Non-Metal Transition in Disordered Systems, eds. L.R. Friedman and D.P. Tunstall (SUSSP Publ., Edinburgh) p. 149.

Mott, N.F., 1981a, Phil. Mag. **B43**, 941.
Mott, N.F., 1981b, Phil. Mag. **B44**, 265.
Mott, N.F., and E.A. Davis, 1979, Electronic Processes in Non-Crystalline Materials, 2nd Ed. (Oxford Univ. Press).
Mott, N.F., and W.D. Twose, 1961, Adv. Phys. **10**, 107.
Mousty, F., P. Ostoja and L. Passari, 1974, J. Appl. Phys. **45**, 4576.
Narita, S., and M. Kobayashi, 1980, Phil. Mag. **B42**, 895.
Nishida, N., M. Yamaguchi, T. Furubayashi, K. Morigaki, H. Ishimoto and K. Ono, 1982, Solid State Commun. **44**, 305.
Ootuka, Y., S. Ikehata, S. Kobayashi and W. Sasaki, 1976, Solid State Commun. **20**, 441.
Ootuka, Y., S. Kobayashi, S. Ikehata, W. Sasaki and J. Kondo, 1979, Solid State Commun. **30**, 169.
Ootuka, Y., F. Komori, Y. Monden, S. Kobayashi and W. Sasaki, 1980, Solid State Commun. **36**, 827.
Paalanen, M.A., T.F. Rosenbaum, G.A. Thomas and R.N. Bhatt, 1982, Phys. Rev. Lett. **48**, 1284.
Paalanen, M.A., T.F. Rosenbaum, G.A. Thomas and R.N. Bhatt, 1983, Phys. Rev. Lett. **51**, 1896.
Pollak, M., and I. Riess, 1976, J. Phys. **C9**, 2339.
Quinn, J.J., and R.A. Ferrell, 1958, Phys. Rev. **112**, 812.
Quirt, J.D., and J.R. Marko, 1971, Phys. Rev. Lett. **26**, 318.
Quirt, J.D., and J.R. Marko, 1972, Phys. Rev. **B5**, 1716.
Quirt, J.D., and J.R. Marko, 1973, Phys. Rev. **B7**, 3842.
Robert, J.L., A. Raymond, R.L. Aulombard and C. Bousquet, 1980, Phil. Mag. **B42**, 1003.
Romestain, R., S. Geschwind and G.E. Devlin, 1975, Phys. Rev. Lett. 35, 803.
Rosenbaum, T.F., K. Andres, G.A. Thomas and R.N. Bhatt, 1980a, Phys. Rev. Lett. **45**, 1723.
Rosenbaum, T.F., K. Andres and G.A. Thomas, 1980b, Solid State Commun. **35**, 663.
Rosenbaum, T.F., K. Andres, G.A. Thomas and P.A. Lee, 1981a, Phys. Rev. Lett. **46**, 568.
Rosenbaum, T.F., R.F. Milligan, G.A. Thomas, P.A. Lee, T.V. Ramakrishnan, R.N. Bhatt, K. DeConde, H. Hess and T. Perry, 1981b, Phys. Rev. Lett. **47**, 1758.
Rosenbaum, T.F., R.F. Milligan, M.A. Paalanen, G.A. Thomas, R.N. Bhatt and W. Lin, 1983, Phys. Rev. **B27**, 7509.
Sasaki, W., 1980a, Phil. Mag. **B42**, 725.
Sasaki, W., 1980b, Proc. 15th Intern. Conf. Physics Semiconductors, Kyoto, p. 31; J. Phys. Soc. Jpn. **49** Suppl. A, 31.
Sasaki, W., S. Ikehata and S. Kobayashi, 1974, J. Phys. Soc. Jpn. 36, 1377.
Slichter, C.P., 1955, Phys. Rev. **99**, 479.
Sonder, E., and D.K. Stevens, 1958, Phys. Rev. **110**, 1027.
Sullivan, P.F., and G. Seidel, 1968, Phys. Rev. **173**, 679.
Thomas, G.A., M. Capizzi, F. DeRosa, R.N. Bhatt and T.M. Rice, 1981a, Phys. Rev. **B23**, 5472.
Thomas, G.A., Y. Ootuka, S. Kobayashi and W. Sasaki, 1981b, Phys. Rev. **B24**, 4886.
Thomas, G.A., Y. Ootuka, S. Katsumoto, S. Kobayashi and W. Sasaki, 1982a, Phys. Rev. **B25**, 4288.
Thomas, G.A., A. Kawabata, Y. Ootuka, S. Katsumoto, S. Kobayashi and W. Sasaki, 1982b, Phys. Rev. **B26**, 2113.
Thomas, G.A., M.A. Paalanen and T.F. Rosenbaum, 1983, Phys. Rev. **B27**, 3897.
Thouless, D.J., 1977, Phys. Rev. Lett. **39**, 1167.
Townsend, P., 1978, J. Phys. **C11**, 1481.
Tunstall, D.P., 1980, Phil. Mag. **B42**, 735.
Tunstall, D.P., and V.G.I. Deshmukh, 1979, J. Phys. **C12**, 2295.
Ue, H., and S. Maekawa, 1971, Phys. Rev. **B3**, 4232.
Walstedt, R.E., R.B. Kummer, S. Geschwind, V. Narayanamurti and G.E. Devlin, 1979, J. Appl. Phys. **50**, 1700.
Webman, I., J. Jortner and M.H. Cohen, 1975, Phys. Rev. **B8**, 2885.
Wegner, C., 1976, Z. Phys. **B25**, 327.
Wegner, C., 1979, Z. Phys. **B35**, 207.
Yamanouchi, C., K. Mizuguchi and W. Sasaki, 1967, J. Phys. Soc. Jpn. **22**, 859.