RAPID COMMUNICATIONS

PHYSICAL REVIEW B VOLUME 43, NUMBER 10 1 APRIL 1991

Metal-semiconductor transition in partially compensated Ge:Sb

J. C. Phillips
AT&T Bell Laboratories, Murray Hill, New Jersey 07974
(Received 21 December 1990; revised manuscript received 28 January 1991)

The conductivity exponent $\alpha = d\ln\sigma/d\ln(n_D - n_c) \approx 0.5$ for uncompensated shallow donor-impurity bands, but it crosses over to $\alpha \approx 1.0$ for partially compensated impurity bands with $K = n_A/n_D$. I calculate K_c, the crossover compensation, and obtain a result in excellent agreement with experiment.

Metal-semiconductor transitions have become more interesting since the discovery of high-T_c superconductivity (HTSC) in layered cuprates.[1] These materials, however, are extremely complex, and it is never certain whether the disorder that is responsible for their high room-temperature resistivities (comparable to metallic glasses, where homogeneity is also questionable) is random or microscopically homogeneous. The situation in Si:P impurity bands is much more favorable. The chemical interactions between impurities in Si:P can be extremely small, and with modern semiconductor technology, it is possible to produce Si:P which is more nearly homogeneous than any other material which exhibits a metal-semiconductor transition. Other factors, such as the ability to shift the dopant concentration n_D relative to the critical transition concentration n_c, by means of uniaxial strain on a single sample, are also very favorable for precise determination of the microscopic character of the transition. It may be fair to say that any theoretical technique which cannot explain the metal-semiconductor transition for relatively simple Si:P impurity bands is unlikely to explain the origin of anomalous electronic properties (such as HTSC) in other much more complex materials (such as layered cuprates) which are also vicinal to this transition.

The central facts are the following.[1] When Si is doped with donors only (P, As, Sb, or P+As), the zero-temperature conductivity $\sigma(0)$ is given by

$$\sigma(0) \propto [(n_D/n_c) - 1]^{\alpha}, \tag{1}$$

where n_D is the donor concentration and n_c is the critical donor concentration separating metallic and insulating states. The best experimental value for α is 0.51(5). However, when the donors are partially compensated by acceptors, α rapidly crosses over from ~ 0.5 to ~ 1.0. If $K = n_A/n_D$ measures the degree of compensation, the crossover value[2] where $\alpha \approx 0.7$ is $K \approx 0.05$, which is small, and perhaps surprisingly so, even for what may possibly be an interference effect.

The theoretical situation is the following. Almost all theories which have discussed this transition approached from the metallic side have used a perturbative treatment of strong scattering in the context of one-component Fermi-liquid theory (FLI). Such perturbative treatments of localization and Coulomb interaction effects are valid for homogeneous well-ordered quasicrystals with narrow diffraction bands, such as *I*-Al-Li-Cu. There the localization effects are of order 10%, while the Coulomb interaction effects are of order 0.1%, but there is little indication of a metal-insulator transition.[3] To obtain a formula resembling Eq. (1), one can extrapolate perturbation theory beyond its region of applicability,[1,4] but this yields $\alpha = 1$ for $d = 3$ dimensions. (This is the value also obtained by classical scaling theory.[5]) A recent attempt[6] to combine scaling theory with one-component Fermi-liquid theory disappointingly led to a discontinuous conductivity $\sigma_0 = 0$ for $n < n_c$ and $\sigma_0 = \sigma_{\min}$ for $n > n_c$, which corresponds to $\alpha = 0$ and is essentially Mott's minimum conductivity picture.[7] The discouraging result here is that technical refinements (replacement of Born scattering by diffusive scattering) in one-component Fermi-liquid theory with interactions (FLI) produce *worse* results[8] than were previously obtained in the one-electron approximation.[4]

The obvious conclusion which can be drawn from this brief discussion, or from a more lengthy review of perturbative one-component Fermi-liquid theories,[9] is that the problem here is fundamentally not technical but conceptual. The fundamental assumption that perturbation theory[4] or even convergent one-component Fermi-liquid theory can predict the correct value for α may be incorrect. Even classical scaling, which gives $\alpha = 1$, may be correct in most materials [such as ^{1}Si or Ge with partially compensated shallow impurities, or deep impurities (Si:Au)] only by accident. Until we understand the origin of $\alpha = 0.5$ in the ideal case, Si:P, we do not understand subtle quantum interactions near metal-semiconductor transitions.

Because the approach of perturbative one-component Fermi-liquid theory (or FLI) produces such disappointing results, it is natural to discard this assumption in favor of nonperturbative two-component Fermi-liquid theory (FLII). The two components here are the obvious ones: localized states (which may or may not be magnetic, and which are the only states present for $n_D < n_c$) coexist for $n_D > n_c$ with extended states, and at the outset are nonperturbatively separable from them.[9,10] The separability is discussed from the point of view of set theory, which is intrinsically nonperturbative, and the set-theoretical description in FLII is not derivable from increasing interaction strengths[6] in FLI. Of course, increasing the carrier charge e does not increase the ratio of electron-electron interaction energies (e^2) to electron-ion interaction energies (Ze^2), and it actually *decreases* the ratio of collective plasma oscillator energies ($\hbar\omega_p$, $\omega_p^2 = 4\pi Ne^2/m$) to

electron-ion interaction energies. It is primarily the randomness of the electron-ion interactions that produces localization, not strong electron-electron interactions; the Mott-Anderson transition in impurity bands is different from the Wigner transition of a free-electron gas. The impurity-band transition may well be explicable entirely within the context of the Hartree-Fock one-electron approximation without many-electron dynamical correlations.

The question of whether extended states can coexist with and be separated from localized states depends not on the strength of electron-electron interactions, but rather on dimensionality d. In general, the central limit theorem tell us that the effects associated with randomness scale with $N^{d/2}$, where N is the number of donor electrons and d is the dimensionality. This feature applies not only to the quantum-mechanical impurity band problem[9,10] but also to the classical random-field Ising model.[11] In the latter, phase transitions to ordered states are suppressed in lower dimensionalities by the formation of domains bounded by domain walls with surface filling factors proportional to N^{d-1}. For impurity bands the quantum-mechanical limits on measuring momentum with electrodes separated by sample dimensions similarly introduce the same factor N^{d-1}. The same marginal dimensionality d_n arises in both cases when $d-1=d_m-1=d_m/2$, or

$$d_m=2. \tag{2}$$

For $d>d_m$, localized and extended states can coexist and be separated.[9,10] This is the case for Si:P.

If extended states exist, what is their effect on the conductivity? Because the states have energy E and an average direction $\hat{\mathbf{e}}$, we can define $k^2=m^*(E-E_0)/\hbar^2$ and $\mathbf{k}=k\hat{\mathbf{e}}$. Here $\mathbf{k}$ corresponds to a Bloch vector in the sense that its component parallel to the applied field measures the *average* wavelength of the extended state in that direction. Near $E=E_0$,

$$dN_e/dE=dN_e/dk\,(dE/dk)^{-1} \propto k^{d-1}k^{-1}=k^{d-2} \tag{3}$$

so that

$$dN_e/dE \propto (E-E_0)^m,\quad m=(d-2)/2. \tag{4}$$

For E near E_0, the density of localized states

$$dN_l/dE \gg dN_e/dE \tag{5}$$

so that the scattering rate Γ or $1/\tau$ is nearly constant. This gives

$$\alpha=(d-2)/2 \tag{6}$$

or $\alpha=0.5$, in agreement with the experimental result (1).

This simple argument rests on nothing more complicated than a quantum-mechanical analogue of the set-theoretic reasoning which is agreed to be seminal in classical random-field Ising models.[11] Because of its nonperturbative character, however, we would like to test the reasoning against the observed crossover[2] of α from 0.5 ($K=0$) to 1.0 ($K\sim 1$) near 0.7 ($K=0.05$). To do this we need to modify our set-theoretic approach to include the effects of compensation.

In Si:P with homogeneous samples we assume that the P is randomly distributed, and this may be the case for both Sb and the background acceptors A in partially compensated Ge:Sb as well.[2] We divide the Sb donors into two types, "pure" donors which contain no acceptor atoms A in their nearest-neighbor sphere (type D_0), and compensated donors which contain one or more A atoms in this sphere (type D_1). The nearest-neighbor sphere is defined as the volume which lies within r of the central atom, with $0\le r\le 1.2r_0$, where r_0 is the average donor spacing. One can estimate (for instance, from random-sphere packing models) that for $d=3$ this sphere contains on the average about

$$N_1\sim 10 \tag{7}$$

atoms. The fraction f_0 of D_0 atoms is given by the Poisson distribution as

$$f_0=e^{-N_1K}. \tag{8}$$

We assume that electrons in extended states percolate through the crystal along paths containing pure donors D_0. There are a number of possible reasons for this assumption, but the simplest one is that an electron is an extended state passing through a D_1 compensated donor may exchange with an electron localized on the near-neighbor A^- acceptor.[12] This exchange scattering interrupts the phase coherence of the extended state, because the phase of the localized A^- electron is random. Alternatively, we have mixed a localized state (A^-) into the donor extended state, contrary to our original prescription which separated extended and localized states. Localization can also occur because of donor clustering, and that is why sample homogeneity is necessary to observe $\alpha\approx 0.5$. Compositional inhomogeneities can always broaden the transition and give $\alpha\approx 1.0$.

The necessary condition for D_0 percolation in the mean-field approximation[13] is that on the average each D_0 atom has at least two (and probably three) D_0 neighbors in the first-neighbor sphere, i.e.,

$$f_0=e^{-N_1K}=2.5(5)/N_1, \tag{9}$$

which with $N_1=10$ gives the percolation threshold

$$K_c=0.14(2). \tag{10}$$

By interpolating linearly on (K,α) between (0, 0.5) and (0.14, 1), we obtain a crossover near

$$(K_1,\alpha_1)=[0.06(1),0.7], \tag{11}$$

in excellent agreement with the sparse experimental data[2] with $K_1=0.05(5)$.

The set-theoretic, or topological, approach embodied in FLII seems to be poorly suited to calculation, while perturbation theory, combined with scaling arguments,[5] as in FLI, produces very lengthy calculations, which are widely recognized.[14,15] Thus, if one measures the significance of a theoretical paper by the number of equations it contains, FLI would seem to be much more powerful than FLII.

However, one of the characteristics of set theory is that in the few problems not soluble by analytic methods it is sometimes able to derive a few important results. This is because when analytic methods are applied to infinite sets they may produce counting errors (for instance, by including amplitudes but not phases) which set theory avoids. The present example of α and K_1, both of which necessarily involve measuring (or counting) infinite sets of basis states associated with randomly distributed impurities, seems to fit this description. Phase also plays an important role in superconductivity, and I have argued elsewhere that loss of the phase coherence of extended states by scattering into localized states can explain the linear dependence on temperature of the normal state resistivity of layered high-T_c cuprates.[16]

---

[1]R. F. Milligan and G. A. Thomas, Annu. Rev. Phys. Chem. **36**, 139 (1985); G. A. Thomas, M. Paalanan, and T. F. Rosenbaum, Phys. Rev. B **27**, 3897 (1983).

[2]G. A. Thomas, Y. Ootuka, S. Katsumoto, S. Kobayashi, and W. Sasaki, Phys. Rev. B **25**, 4288 (1982).

[3]K. Kimura *et al.*, J. Phys. Soc. Jpn. **58**, 2472 (1989).

[4]K. F. Berggren, J. Phys. C **15**, L45 (1982); N. F. Mott and M. Kaveh, J. Phys. C **15**, L697 (1982); **15**, L707 (1982).

[5]E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).

[6]C. Castellani, G. Kotliar, and P. A. Lee, Phys. Rev. Lett. **59**, 323 (1987). This paper contains a large number of surprising results. It predicts, for example, that γ, the strength of the linear term in the specific heat, approaches infinity on the metallic side of the transition as n is decreased to n_c, as *all* the electrons are localized as *free* spins. This prediction has not been confirmed by experiment; M. A. Paalanen, J. E. Graebner, R. N. Bhatt, and S. Sachdev, Phys. Rev. Lett. **61**, 597 (1988). The approximations made in this paper try to describe the correlations concealed in N^2 off-diagonal matrix elements by a few constants which describe the energy dependence of N diagonal renormalization factors.

[7]N. F. Mott, Philos. Mag. **26**, 1015 (1972).

[8]In discussing the experimental situation with $\alpha=0.5$ in Si:P, Ref. 5 makes two hopeful suggestions: (a) Spin-orbit scattering distorts the experimental data somehow, or (b) $\alpha\approx 1$, or at least $\alpha\geq\frac{2}{3}$, "the rigorous limit." Neither of these suggestions can be correct. The spin-orbit coupling strength scales as r_D^{-3}, where r is the impurity radius, and also as Z^3, where Z is the effective charge of the impurity. For Si:Au, $r\approx a$ (the atomic radius) and $Z\approx 10^2$, so that spin-orbit scattering might be important. (Inhomogeneous clustering is probably even more important.) But for Si:P, $r\approx 10a$, $Z\sim 1$, and λ/W (where W is the impurity bandwidth) is negligibly small, especially when one notices that the conduction-band edge in Si is *nondegenerate*. Case (b) is already ruled out by $\alpha=0.51(5)$ experimentally (Ref. 1). In case (b) the authors of Ref. 5 also suppose that a series expansion in $d-2$ might give a continuous transition, as in Ref. 3. But, as in Ref. 3, not only must such a series expansion be carried beyond its range of applicability ($d-2=1$!), but here the first-order term alone is inadequate. The second- and higher-order terms are not calculated, even as to sign, essentially because of technical difficulties associated with FLI diffusive basis functions for $d=2$.

[9]J. C. Phillips, Philos. Mag. **47**, 407 (1983); Philos. Mag. B **58**, 361 (1988). Note that the phase of the extended state is defined only parallel to the applied field, not normal to it. Thus the usual factor n (the total number of conductive electrons) in the conductivity is replaced by $(n_l+n_e)^{2/3}n_e^{1/3}$ and the usual factor $n\approx(E_F-E_0)^{3/2}$ is replaced by $n_e^{1/3}\approx(E_F-E_0)^{1/2}$.

[10]J. C. Phillips, Solid State Commun. **47**, 191 (1983). The coexistence of exponentially localized and extended states for $d>2$ was demonstrated by construction for a special case (the Anderson model) by J. T. Chayes, L. Chayes, J. R. Franz, J. P. Sethna, and S. A. Trugman, J. Phys. A **19**, L1173 (1986). However, this analysis was not formulated in the set-theoretic language of Refs. 9–11. Axiomatic set theory, as described by M. Tiles, *The Philosophy of Set Theory* (Blackwell, Oxford, 1989), pp. 118–134, is especially relevant to this problem because it shows that constructive algorithms, while useful for well ordering (alphabetizing) a set, are unnecessary for separability, which is the central issue in discussing the behavior of $\alpha(K)$.

[11]Y. Imry and S.-K. Ma, Phys. Rev. Lett. **35**, 1399 (1975).

[12]I restrict the exchange interactions to nearest neighbors because the exchange interaction (which in a free-electron gas is long range in the Hartree-Fock approximation) is actually short range when dynamical correlations are included [E. P. Wigner, Phys. Rev. **46**, 1002 (1934)]. This is the only point where many-electron effects might be important. However, note that near the metal-insulator transition most of the states are localized, $N_1(D_1)\approx 8$, $N_1(D_0)\approx 2$. For such localized states the interaction is inherently short range, so that inclusion of dynamical correlations may actually have only a small effect on α or K_c.

[13]S. Kirkpatrick, Rev. Mod. Phys. **45**, 574 (1973).

[14]M. Lakner and H. v. Löhneysen, Phys. Rev. Lett. **63**, 648 (1989); **64**, 108 (1990).

[15]J. C. Phillips, Phys. Rev. Lett. **64**, 107 (1990).

[16]J. C. Phillips, Phys. Rev. B **40**, 7348 (1989); **41**, 8968 (1990). It is also instructive to compare the description of coherent extended percolative states, as done here by set theory, with previous treatments (Refs. 3 and 5) which make no spatial analysis of their basis states. Another instructive comparison is with spin glasses. There I showed [J. C. Phillips, Phys. Rev. B **31**, 3179 (1985)] that separation of the smallest rings into frustrated $\{f\}$ and consistent $\{c\}$ led to an estimate of the spin-glass transition temperature kT_g of a simple cubic model in good agreement with the results of numerical simulations ($T_g/J=1.25$ (est.), 1.2(2) (sim.) [A. T. Ogielski and I. Morgenstern, Phys. Rev. Lett. **54**, 928 (1985)]). This separation parallels the present $\{D_0\}$, $\{D_1\}$ separation.