

SEMICONDUCTOR-METAL TRANSITIONS

T. M. Rice

Bell Laboratories

Murray Hill, New Jersey 07974, U.S.A.

ABSTRACT

The theory of the semiconductor-metal transition is reviewed. The study of electron-hole fluid has shed new light on the problem and revived the old idea that the long-range Coulomb force makes the semiconductor-metal transition intrinsically first order. Recent theoretical work supports the idea of separate liquid-gas and metal-nonmetal critical points in the electron-hole fluid in germanium.

I. INTRODUCTION

The theory of transitions between metallic and nonmetallic systems has been the focus of increasing attention in recent years. It is now some thirty years since the pioneering papers by Landau and Zeldovich,¹⁾ and by Mott²⁾ which posed the problem. Since that time a wide variety of systems have been studied experimentally and theoretically. In this brief review we shall not attempt to cover the diverse systems of interest and instead refer the interested reader to the excellent book on the subject by Mott.³⁾ Rather, the focus is on the close relationship between the study of a high density fluid of electrons and holes created by a nonequilibrium external source and the metal-semiconductor transition as an indirect band gap passes through zero. The study of the former system, which is the main topic of this conference, has shed new light on some of the old questions on the latter topic. No attempt will be made to review the electron-hole fluid in detail here. Such reviews can be found elsewhere.⁴⁻⁶⁾

In particular, in the original papers of Landau and Zeldovich¹⁾ and Mott²⁾ the role of the Coulomb force at the metal-semiconductor transi-

tion was stressed. Peierls (quoted in ref. 1) and Mott²⁾ argued that the long-range nature of the Coulomb force would cause a transition between the metallic and semiconducting phases to be intrinsically first order at zero temperature. The argument can be simply stated. In the semiconducting phase an electron and a hole attract each other with a long-range Coulomb potential which always has a bound state. On the other hand in a metallic phase with a finite density of free carriers, the long-range Coulomb force is screened and replaced by a short-range force which will not have a bound state for sufficiently strong screening. Therefore the metallic phase is stable only with a finite density of electrons and holes and it follows that the transition is first order.

In discussing the subsequent work it is helpful first to relate the two problems — the semiconductor-metal transition and the electron-hole fluid. This is the topic of the next section. In Section III the theory is presented for a low density fluid of electrons and holes and related to the earlier work on the excitonic insulator. In Section IV the theory of the metallic electron-hole liquid is reviewed and a general discussion of the high density regime is given. The transition between the two regions especially as a function of temperature is the subject of Section V. At present there are very interesting open questions on the transition from the low density of exciton gas which is weakly ionized to the high density plasma which is strongly ionized.

II. THE SEMICONDUCTOR-METAL TRANSITION AND THE ELECTRON-HOLE LIQUID

We shall begin by discussing the relationship of the theory of the electron-hole liquid to the problem of the semiconductor-metal transition. In the first case one studies the energy of fixed number of electrons and holes interacting with Coulomb forces. The Hamiltonian of the system can be described using the effective mass approximation

$$\begin{aligned}
H = & - \sum_i \frac{\nabla_i^2}{2m_e} - \sum_j \frac{\nabla_j^2}{2m_h} + \frac{1}{2} \sum_{ij} \frac{e^2}{\kappa |\vec{r}_i^e - \vec{r}_j^e|} + \frac{1}{2} \sum_{ij} \frac{e^2}{\kappa |\vec{r}_i^h - \vec{r}_j^h|} \\
& - \sum_{ij} \frac{e^2}{\kappa |\vec{r}_i^e - \vec{r}_j^h|}
\end{aligned} \tag{2.1}$$

where m_e and m_h are the electron and hole effective masses and κ is the static dielectric constant. The assumption is made that in the steady state the rates at which electrons and holes are being generated and are recombining are sufficiently slow and may be neglected. Thus one assumes that there is true thermodynamic equilibrium of a model system described by eq. (2.1). The thermodynamic variables are the density and temperature.

In the semiconductor-metal transition the simplest case to study is that of an indirect gap semiconductor in which the energy gap is varied through zero. The background dielectric constant κ is due to vertical transitions and need not vary appreciably as the indirect band gap varies through zero. By varying the band structure the chemical potential, or the energy required to add an electron and hole to the system, is varied. In this system the number of electrons in the conduction band and the number of holes in the valence band are clearly not conserved. The relationship between the two systems was pointed out by Halperin and Rice,⁷⁾ namely that they are both described by the Hamiltonian (2.1) but with different thermodynamic variables.

There are, of course, other types of semiconductor-metal transitions involving direct gap materials or the overlap of Mott-Hubbard bands or disordered materials. The key assumption to which we will be restricted, is that the background dielectric constant κ and effective masses m_e and m_h remain finite at the transition. This can be true even in direct gap materials if the valence and conduction bands belong to different irreducible representations such that the interband matrix element of the momentum operator vanishes at the symmetry point. Dis-

ordered materials such as Si:P, where the P sites are frozen in fixed random positions in the lattice, are also excluded from this brief review.

III. LOW DENSITY GAS OF EXCITONS: THE EXCITONIC INSULATOR

It was pointed out, some years ago, by Knox⁸⁾ that the binding energy of an exciton, E_x , could become greater than the band gap, E_{gap} , in an indirect gap semiconductor leading to an instability of the ordinary ground state of the crystal with respect to the spontaneous formation of excitons. The value of E_x is given simply by a modified hydrogenic formula

$$E_x = \frac{\mu_0}{m\kappa^2} \text{ (rydbergs)} \quad (3.1)$$

where $\mu_0^{-1} = m_e^{-1} + m_h^{-1}$. As discussed above, both μ_0 and κ may depend sensitively on the direct energy gap, but not on an indirect energy gap. The theoretical description of the new distorted phase, caused by the spontaneous formation of excitons was first developed by des Cloizeaux⁹⁾ and Keldysh and Kopayev¹⁰⁾ and subsequently discussed by many others. It has been reviewed by Halperin and Rice^{7,11)} and by Kohn.¹²⁾ Consider the simplest case of a valence band with a single maximum at the zone center ($\vec{k}=0$) and a conduction band with a minimum at the zone boundary ($\vec{k}=\vec{w}$). If $E_{\text{gap}} < E_x$, excitons are present and the Hartree-Fock approximation was used to describe the distorted state and the expectation value

$$\langle b_{\vec{k}+\vec{w},\sigma}^+ a_{\vec{k},\sigma'} \rangle \neq 0 \quad (3.2)$$

for some σ and σ' . The distorted state, generally known as the excitonic insulator, has a period in real space which is just double the period of the undistorted lattice. In the Hartree-Fock picture, the one electron states of the distorted crystal are made up of linear

combinations of wave functions of wave vectors \vec{k} and $\vec{k}+\vec{w}$ from the valence and conduction bands, respectively, of the nondistorted crystal. If the expectation value (3.2) is real the distorted phase is a charge density or spin density wave. If (3.2) is imaginary, orbital anti-ferromagnetic or spin current states are possible.⁷⁾ The energies of these four possibilities are split by interband scattering effects when the electron and hole are on the same site. In the simplest model the spin density wave state is lowest.

Let us examine these results from the point of view of a gas of excitons. Let $A_{\vec{w}}^+$ be the creation operator for an exciton with wave vector \vec{w} ,

$$A_{\vec{w}}^+ = \sum_{\vec{k}, \sigma, \sigma'} f_{\sigma\sigma'}(\vec{k}) b_{\vec{k}+\vec{w}, \sigma}^+ a_{\vec{k}, \sigma'} \quad (3.3)$$

where $f_{\sigma\sigma'}(\vec{k})$ is an envelope wave function peaked around $\vec{k}=0$. A Bose condensate of excitons means that $\langle A_{\vec{w}}^+ \rangle \neq 0$, which is identical to (3.2). Treating the crystal in the Hartree-Fock approximation is equivalent to treating the gas of excitons as a repulsive Bose gas which form a Bose condensate. The repulsion, in the Hartree-Fock approximation, arises from the Pauli exclusion principle on the electrons and the holes.

The interaction between two excitons is not purely repulsive. At large distances there will be a Van der Waals attraction between excitons. The short-range repulsion due to the Pauli exclusion principle will only apply to electrons or holes in a relative triplet state. Thus there is a strong attraction between two excitons if both the electrons and holes are in relative spin singlet states. Indeed it has been shown that a bound state, an excitonic molecule, sometimes called a biexciton can be formed in this case.¹³⁻¹⁵⁾ It is a very extended state, however, and the binding energy of an excitonic molecule is only a few per cent of E_x if the electron and hole masses do not differ by a

large amount. In the Hartree-Fock approximation the interaction is averaged over space and spin configurations and the molecular correlation ignored. However, in the true dilute limit spatial molecular correlations dominate.

For a nondegenerate band structure higher complexes are not bound because of the large repulsive core in the exciton-exciton interaction when either the holes or electrons are in a triplet state. Brinkman and Rice¹⁶⁾ have estimated the scattering length between two molecules by treating each molecule simply as the superposition of two individual excitons. They obtained a value of $7a_x$ for the scattering length (a_s) in the limit $m_e = m_h$ (a_x is the exciton Bohr radius). One question of interest is whether a molecular liquid or solid is stable. This is certainly true in the limit $m_e/m_h \rightarrow 0$, where the stable phase is a molecular solid (*i.e.*, solid H_2). In the opposite limit ($m_e = m_h$) Brinkman and Rice¹⁶⁾ have argued, using the de Boer theory¹⁷⁾ of the quantum corrections to equation of state, that no liquid-gas transition will occur and the energy increases monotonically with the density, n . They also propose that the molecular phase may be described as a weakly interacting gas of molecules when the interparticle spacing $r_s a_x > a_s$ [$n = 3/4\pi r_s^3 a_x^3$]. At higher density ($r_s \lesssim 7$ using the values of ref. 16) the energy gained by the formation of molecules is overcome by the repulsive interactions. In this density regime it is possible that the Hartree-Fock approximation, discussed above, could apply. However, as we shall see below, the energy of the metallic state will be very close by and it is doubtful that there will be any region of validity of the Hartree-Fock approximation, or equivalently of Bose condensed excitons.

It is interesting to study the limit of metastability of the low density insulating phase. [A first-order transition to a high density metallic phase can occur at much lower density.] Hanamura¹⁸⁾ considered the criterion for the limit of metastability to be the density at which the excitation energy (η) to create a spatially separate electron and

hole goes to zero. If $\eta < 0$ there is no gap in the spectrum of current carrying excitations and one has a metallic state. Hanamura¹⁸⁾ has used the Hartree-Fock approximation to estimate $\eta(n)$. In fact, the calculation is similar to those made in the context of the theory of the excitonic insulator. He found $\eta(n)$ increases initially linearly with n . This repulsive interaction arises from the Pauli exchange repulsion between the electron and the exciton. The long-range interaction between an electron and an exciton is dominated by the attraction between the electron and the polarization induced on the exciton. Hanamura¹⁸⁾ included a polarization correction in lowest order and found it was sufficiently attractive to overcome the repulsion found in Hartree-Fock. He estimated that the critical value of metastability of the insulating phase occurs at a value of $r_s^{M-1} \simeq 3$. This calculation, however, does not adequately treat the spatial and spin-dependent correlations in the low density limit. Rice¹⁹⁾ has recently made a crude attempt to estimate r_s^{M-1} based on an effective electron-exciton potential and obtained a value $r_s^{M-1} \lesssim 2$.

So far, we have discussed the simplest case $m_e \approx m_h$ and nonorbitally degenerate electrons and holes. If $m_h \gg m_e$ (or *vice versa*) new possibilities arise. The stability of the molecular state is greatly enhanced and, as the experience of H_2 tells us, molecular solids and liquids can form. However, the zero point motion gives corrections which goes as $(m_e/m_h)^{\frac{1}{2}}$ and such states are confined to the limit $m_h \gg m_e$.

IV. THE METALLIC STATE: ELECTRON-HOLE LIQUID

At low densities excitons may be viewed as in the last section as well-defined entities with only weak interactions between them. Keldysh²⁰⁾ first pointed out that a high density metallic state of electrons and holes could have a lower energy than the low density phase.

In such a metallic electron-hole liquid excitons cease to exist and any spatial correlation between a given electron and hole is extremely transitory. In this limit the exciton overlap is so strong that they lose their individuality and the e-h fluid must be viewed as two interpenetrating Fermi fluids of electrons and holes. We will consider first the case of the idealized band structure with nondegenerate electron and hole bands and later in this section review the complications introduced by the real band structure of Ge or Si.

In the high density limit the leading term in energy will be the kinetic energy,

$$E_K = \frac{3}{5} \left(\frac{k_F^2}{2m_e} + \frac{k_F^2}{2m_h} \right) = \frac{2.21}{r_s^2} \quad (4.1)$$

expressed as the energy per pair in units of E_x . The kinetic energy gives a large positive (*i.e.*, repulsive) contribution which increases with density as $n^{2/3}$.

The first correction is the exchange energy, E_{exch} , between the electrons and the holes separately. This is an attractive contribution arising from the reduction in Coulomb energy because of the spatial correlation imposed by the Pauli principle on electron (holes) in the same spin state.

$$E_{\text{exch}} = - \frac{3e^2 k_F}{2\pi\kappa} = - \frac{1.832}{r_s} \quad (4.2)$$

The remaining term is known as the correlation energy, E_{corr} , which arises from the correlations primarily between particles in unlike quantum states. In the e-h liquid these are of two types. One is electron-hole correlations, *i.e.*, between oppositely charged particles. The second is the electron-electron (or hole-hole) correlation between particles of like charge but opposite spin. These terms can be only approximately evaluated.

The first attack on the problem was by Hanamura¹⁸⁾ who used high density expansion to estimate this energy and obtained a local minimum in the total ground state energy-density curve $E_G(r_s)$ but with a minimum value at $r_s^0 = 1.7$ of $E_G^0 = -0.35E_x$. This value is substantially above the energy of separated excitons. It is known from experience with the single component electron gas that high density expansions in r_s are strictly limited to $r_s \leq 1$. In the single component electron gas a modified random phase approximation (RPA) is believed to give a good description of the energy even in the intermediate density regime ($r_s \sim 1-5$). This approximation has been applied to the electron-hole liquid by Brinkman, Rice, Anderson and Chui.^{21,16)} These authors find a much lower minimum value ($-0.86 E_x$) at a similar density.

The biggest question about the RPA is whether the correlation between electrons and holes are properly treated. Since it is just such correlations which give rise to the exciton bound state, they become increasingly important as the density is lowered. Within the RPA electron-hole correlations are included in lowest order and treated as opposite in sign to the electron-electron correlations. Recently Vashishta, Bhattacharyya and Singwi (VBS)²²⁾ have applied the Singwi-Tosi-Land-Sjolander (STLS) method to the electron-hole liquid. In this method a set of self-consistent equations are obtained by approximating the equations of motion. The results of VBS for $m_e = m_h$ give a lower minimum value of $E_G^0 = -0.99 E_x$ but still $> -E_x$. Their approximation, however, builds in substantial more electron-hole correlation. A measure of such correlations is the enhancement factor $g_{eh}(0)$ which is the ratio of the probability of finding an electron at a hole to the mean density. Whereas, Brinkman and Rice¹⁶⁾ find a value of 2 for $g_{eh}(0)$ at $r_s = 1.7$, Vashishta *et al.*²²⁾ report a value $g_{eh}(0) \approx 8$ at that density. A comparison of the two results for the ground state energy is shown in Fig. 1. Inoue and Hanamura²³⁾ have attempted a variational calculation with a correlated wave function. They are forced to make approx-

imations in their evaluation so that their answer is not truly variational. Their final answer is very close to that of VBS. It should be pointed out that all of the above calculations involve essentially uncontrolled approximations and it is not possible to assess their accuracy *a priori*. Nonetheless it appears that there is a metallic state which is a local minimum of the energy density curve but it is at an energy slightly higher than that of separated excitons. This conclusion must be regarded as tentative, however, since the energy differences are so small.

In the hydrogenic limit ($m_h \gg m_e$) Wigner and Huntington²⁴⁾ obtained a ground state energy of $-1.05E_x$ for metallic hydrogen, slightly below that of isolated hydrogen atoms but substantially above that of the molecular solid (-1.17). Their calculation using the Wigner-Seitz method included the electron-proton correlation to all orders. Several calculations^{16,25)} have been performed at intermediate values of m_e/m_h and it appears that for isotropic nondegenerate bands the low density molecular phase always has a lower energy than the high density state.

In general semiconductors have more complex band structures than the ideal model discussed above. This can lead to big changes in the energy. The cases of Ge and Si have been explored in detail by several groups. The electron bands in both Ge and Si are highly ellipsoidal and have a high orbital degeneracy (ν_e) with $\nu_e = 4$ (Ge) and $\nu_e = 6$ (Si). In addition there are coupled hole bands which are four-fold degenerate (including spin) at the zone center (Γ) and split away from Γ . The effects greatly reduce the kinetic energy cost and to a lesser extent the exchange energy. When the correlation energy is included a

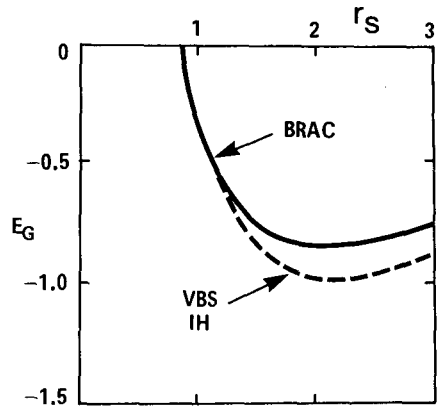


Fig.1 The ground state energy, E_G , of the isotropic nondegenerate e-h liquid in units of E_x as a function of the interparticle separation r_s in units of a_x .

greatly lowered ground state energy with a minimum at a higher density was found by several groups.^{21,26,16,22,27)} These results shown in Fig. 2 are based on the RPA. The results are fairly sensitive to the

details of the band structure but the corrections to the RPA are proportionally much less important than in the ideal band structure. The high degree of orbital degeneracy plays an important role both in stabilizing the metallic state and in minimizing the importance of the corrections to the RPA. This can be seen by examining a diagrammatic expansion of the ground state energy. In each order of an expansion in powers of e^2 the bubble diagrams, summed in the

RPA (see Fig. 3), are the leading term and all corrections are down by a factor ν^{-1} , where ν is the degeneracy factor. In Ge the degeneracy (including the holes and the spin) factor $\nu = 12$ and in Si, $\nu = 16$. The first principles calculations are in very good agreement with experiments on both Ge and Si.^{4,27,16)}

The implications of these results for the semiconductor-metal transition are clear. Consider reducing E_{gap} from a large positive value. Then before one reaches the condition $E_{\text{gap}} = E_x$ the condition $E_{\text{gap}} = -E_G^0$ will be satisfied and at that point a first-order transition

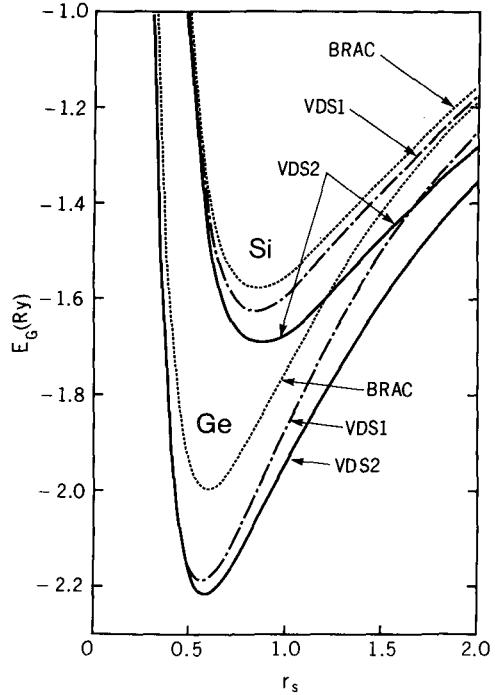


Fig.2 The ground state energy as function of r_s . The exciton rydberg and Bohr radius are defined with optical averaged effective masses [Ge; Ry = 2.65 meV; $a_x = 177 \text{ \AA}$; Si; Ry = 12.8 meV; $a_x = 49 \text{ \AA}$]. The curves BRAC²¹⁾ and VDS1²⁷⁾ were both calculated within RPA and a more accurate band structure was used in VDS1. The curves VDS2²⁷⁾ used the STLS approximation.

will take place directly from the semiconducting phase to a metallic phase with a finite density of electrons and holes. This statement is true only if $|E_G^0| > E_x$ but in general the energy bands are anisotropic and have orbital degeneracy both of which favor the metallic phase. Thus one expects in a first-order transition which bypasses the excitonic phases completely.

What of the excitonic instability in the metallic phase? In the earlier theory of the excitonic insulator, the excitonic instability arose as an instability of the particle hole channel, *i.e.*, from repeated electron-hole scattering [see Fig. 3(c)]. The self-energy effects on the electrons and holes were ignored and the chemical potential taken as E_F , *i.e.*, as noninteracting electrons and holes. The chemical potential μ is given by

$$\mu = E_G(n) + n\partial E_G(n)/\partial n \tag{4.3}$$

so that at the minimum energy value, $\mu = E_G^0$. Thus the chemical potential is much lower than E_F . This lowering of the chemical potential is due to the attractive potential seen by an electron or hole in the metallic liquid. The Pauli exclusion principle keeps apart electrons of the same orbital and spin quantum numbers and thereby causes a net positive charge in the vicinity of an electron. Similarly correlations between electrons keep them apart and correlations between electrons and holes enhance their overlap. Explicit calculations of the self-energy give a rigid shift of the band edges in the metallic

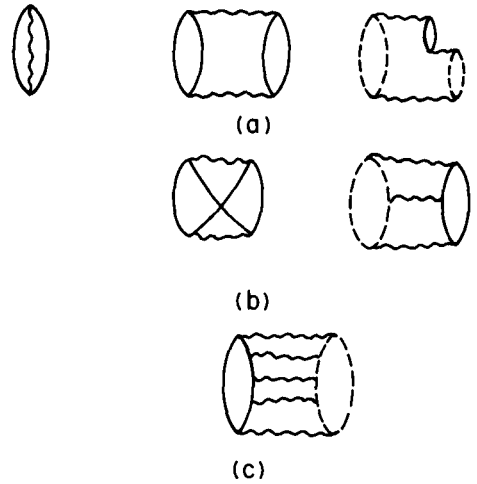


Fig.3 (a) The graphs included in the RPA; (b) graphs omitted in the RPA; (c) electron-hole ladder graphs. Solid and dashed lines denote electron and hole propagators.

phase.²⁸⁾

If one reduces the density, n , below the equilibrium density, n_0 , the electron-hole liquid is mechanically unstable towards a collapse. Nonetheless it is interesting to explore the question of the limit of metastability of the metallic phase ignoring the mechanical instability. This limit of metastability is set by the Mott criterion²⁾ which is the condition for the occurrence of an electron-hole or excitonic bound state. The metallic phase is stable if $q_{F.T.} > a_x^{-1}$ where $q_{F.T.}$ is the Fermi-Thomas screening wave vector. In a simple band structure this yields a condition $r_s < 10$,¹⁶⁾ which is four times larger than for a collection of hydrogen atoms. This increased metastability of the metallic phase is due to the increased screening and increased kinetic energy cost of localization when both positive and negative charges are light. The metastability is further enhanced in more complex band structures by orbital degeneracy, *etc.*

To summarize the phase diagram at zero temperature of the semiconductor-metal transition is a first-order transition between the semiconducting phase for $E_{\text{gap}} > |E_G^0|$ to a metallic phase when $E_{\text{gap}} < |E_G^0|$ for a general phase diagram. For a simple band structure where the lowest electron-hole state is a molecular phase, this phase diagram may not change since the molecular phase does not require a change of crystal symmetry and will not be a distinguishable semiconducting phase.¹¹⁾ If, however, the metallic phase has a band structure with 'nesting', *i.e.*, electron and hole Fermi surfaces which 'nest' when translated by some wave vector, then the distortions characteristic of the excitonic insulator, such as charge density waves or spin density waves can occur. An example of 'nesting' is the isotropic nondegenerate band structure where the electron and hole Fermi surfaces are identical spheres. Under these circumstances the occurrence of charge or spin density wave states is more properly viewed as instability of a metal with a special band structure, as for example in Cr.

V. THE SEMICONDUCTOR-METAL TRANSITION AT FINITE TEMPERATURE

In the previous sections we discussed the high density and low density limits and the transition between them at zero temperature. We turn now to finite temperatures and the interesting possibilities raised by the large region of metastability of the metallic and insulating phases.

At low temperatures, the calculation of the phase boundaries of the first-order transition discussed in the previous section is straightforward. Taking temperature and density of electron-hole pairs as thermodynamic variables, at a density $n < n_0$ the system will break up into a high density liquid and a low density gas. The density in the gas phase is very small at low temperatures and it will exert a negligible pressure on the liquid. The Landau theory of Fermi liquids can be used to calculate the linear thermal expansion.

At low temperatures the gas outside will be composed of excitons, excitonic molecules and ionized electrons and holes. The overall density will be very low and we can therefore ignore interaction effects and treat the equilibrium between the various species by classical statistical mechanics. The chemical potential of each species is equal to that of the liquid. The density of each species, i , is controlled by its work function, ζ_i . This is the energy required to take each subspecies from the liquid to infinity at zero temperature. Thus for $i = \text{exciton}$ then $\zeta_{\text{ex}} = |E_G^0| - E_x$, the binding energy of the e-h liquid relative to the exciton. For excitonic molecules

$$\zeta_{\text{ex.mol}} = 2\zeta_{\text{ex}} - E_{\beta\text{mol}} \quad (5.1)$$

where $E_{\beta\text{mol}}$ is the binding of the excitonic molecule with respect to dissociation into two excitons. Since $E_{\beta\text{mol}}/\zeta_{\text{ex}}$ is estimated to be very small¹³⁻¹⁵ ($\ll 1$) in Ge or Si it follows that there will be very few molecules in the gas outside.

The energies ζ for electrons and holes can be estimated as follows. First, the sum

$$\zeta_e + \zeta_h = \zeta_{ex} + E_x . \quad (5.2)$$

This follows at once from the fact that the energy to take out an electron and hole separately is equal to the energy cost of separating them from an exciton in the gas. Secondly, overall charge neutrality will set $\zeta_e = \zeta_h$ at finite temperature. Thus $\zeta_e = \zeta_h = \frac{1}{2} (\zeta_{ex} + E_x) > \zeta_{ex}$ in Ge or Si. The number of ionized e-h pairs will therefore be small relative to excitons at low temperatures. For simpler band structures, where $E_x \gg \zeta_{ex}$, this ionized fraction will be even smaller. The gas at low temperatures is predominantly composed of excitons.

As the temperature is raised the difference in densities between the liquid and the gas decreases. The simplest phase diagram possible would be to extrapolate the densities on both sides to form a single gas-liquid critical point (n_c, T_c) in the manner shown in Fig.4. This single critical point gives a single line for the phase diagram for the semiconductor-metal transition when we translate to the thermodynamic variables (μ, T) .

In Ge and Si it has been found possible to make a reasonably accurate calculation of the gas-liquid critical point. This was first pointed out by Combescot²⁹⁾ who used essentially the RPA for the free energy. As discussed in the previous section, the corrections to the RPA are greatly reduced by the high degree of orbital degeneracy in Ge and Si. The free energy per pair F can be written as

$$F(n, T) = F_0(n, T) + F_{xc}(n, T) \quad (5.3)$$

where F_0 is the free energy of a noninteracting fluid of electrons and holes. The second term F_{xc} represents the exchange and correlation contributions

Within the RPA the effective masses are only slightly renormalized.²⁸⁾ Thus there is very little change in the low lying excitations caused by the interactions. The principle effect of the interactions is to introduce the plasmon collective excitations at high energy. For densities $\sim 10^{17} \text{ cm}^{-3}$ in Ge the plasmon energy $\hbar\omega_p \approx 14 \text{ meV}$ is very high and there will be no appreciable occupation of these modes until very high temperature. It is safe to assume that $\hbar\omega_p \gg k_B T$ near the critical point. This suggests replacing F_{xc} by $E_{\text{exch}}(n) + E_{\text{corr}}(n)$.

The conditions for a critical point at temperature T_c and density n_c can be expressed simply in terms of the chemical potential μ as

$$\left. \frac{\partial \mu}{\partial n} \right|_{n_c, T_c} = \left. \frac{\partial^2 \mu}{\partial n^2} \right|_{n_c, T_c} = 0 \tag{5.4}$$

where $\mu (= F + n \partial F / \partial n)$ is the chemical potential for an e-h pair. These conditions involve a knowledge of the third derivative of F with respect to n . This leads to some calculational difficulties. The

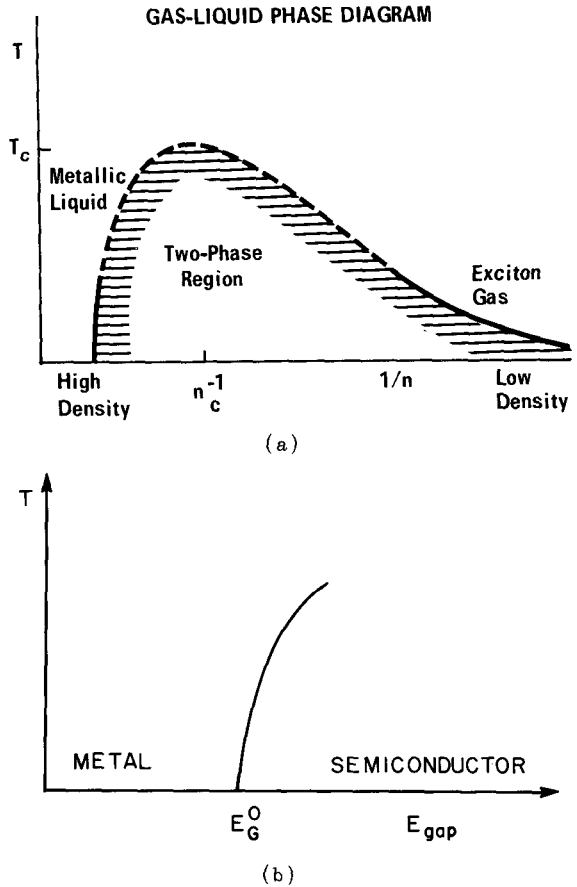


Fig.4 The simplest phase diagram of the e-h liquid obtained by extrapolating the low temperature behavior (a) temperature and density plane, (b) temperature and chemical potential plane.

calculations have been performed by several groups now ^{29,19,27)} and the answers are in reasonable agreement. They predict a critical point with $T_c \approx 7$ K and $n_c \approx 10^{17} \text{ cm}^{-3}$ in Ge. Indeed a very different approach by Reinecke and Ying ³⁰⁾ using the Fisher droplet model gave fairly similar values. Further, in the experiments of Thomas *et al.* ³¹⁾ at critical point was found at $T_c = 6.5$ K and $n_c = 0.8 \times 10^{17} \text{ cm}^{-3}$ in Ge. These values are still well within the metallic regime by any criterion. This assertion is also supported by the luminescence lineshape observed near the critical point by Thomas *et al.* ³¹⁾ We can conclude, with some certainty, that in Ge (and presumably Si too) that there is a liquid-gas critical point within the metallic regime.

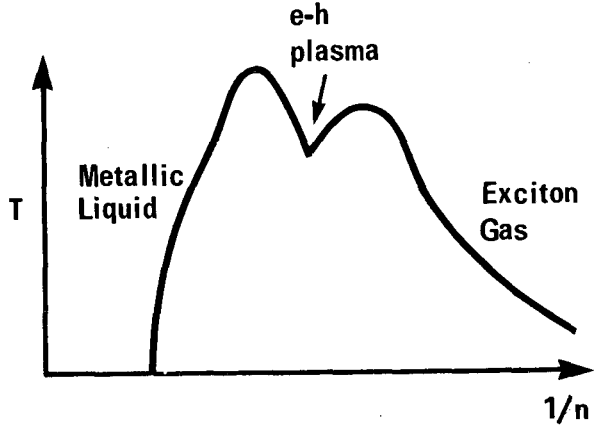
As the temperature is lowered, there is a transition on the low density side of liquid-gas curve between a metallic or strongly ionized plasma in the vicinity of the liquid-gas critical point to a low density weakly ionized exciton gas. This transition is a form of metal-to-nonmetal transition. In previous sections we discussed briefly some of the criteria that can be used at zero temperature to determine the boundary between metallic and insulating behavior. At zero temperature the limits of metastability of the metallic and insulating phases were very different implying a large first-order transition if one could envisage a uniform expansion of the e-h liquid. Such a first-order transition could persist to a high enough temperature to be visible along the coexistence curve. This possibility was suggested by the present author ¹⁹⁾ who revived the early suggestion by Landau and Zeldovich ¹⁾ for fluid Hg. This would lead to a phase diagram as shown in Fig. 5.

The phase transition between weakly ionized and strongly ionized plasma has been studied by several authors. Reviews have been written recently by Norman and Starostin ³²⁾ and by Ebeling, Kraeft and Kremp. ³³⁾ In their approach the ionization equilibrium is determined by an equation of the form

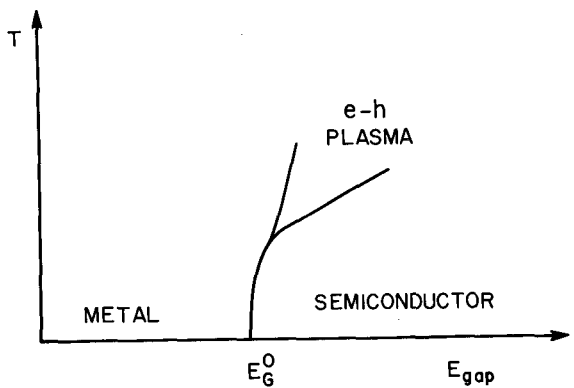
$$n_{ex} = n_e^2 \left(\frac{2\pi\hbar^2 M_{ex}}{kT m_e m_h} \right)^{3/2} \frac{v_{ex} e^{\beta E_x} - \beta \delta E_x}{v_e v_h} \tag{5.5}$$

where n_e and n_{ex} are the density of electrons (or holes) and excitons, respectively, M_{ex} is the translational mass of an exciton and v_{ex} , v_e and v_h are the degeneracy factors of the excitons, electrons and holes including spin and orbital degrees of freedom. (In principle excited states of the exciton should also be included but in practice they are ignored.) The correction δE_x to the exciton binding energy includes the Debye-Huckel screening by the electrons and holes. So far the expression is purely classical but then the lowest order quantum correction to δE_x is added in, giving

$$\delta E_x = e^2 / \kappa r_D [1 - C\Lambda / r_D] \tag{5.6}$$



(a)



(b)

Fig.5 Phase diagram expected if a second critical point associated with the metal-nonmetal transition occurs (a) temperature and density plane and (b) temperature and chemical potential plane.

where Λ is the thermal de Broglie wavelength for electrons (Λ

$= h(2\pi m_e kT)^{-\frac{1}{2}}$ and r_D is the Debye-Huckel screening radius ($r_D = (8\pi\beta n_e e^2/\kappa)^{-\frac{1}{2}}$ and $\beta^{-1} = k_B T$). The constant C is a numerical constant ≈ 0.1 . A further approximation can be made which has similar accuracy to namely

$$\delta E_x = \frac{e^2}{\kappa r_D} \frac{1}{1 + C\Lambda/r_D} \quad (5.7)$$

Norman and Starostin point out that the structure of (5.7) is typical for the theory of strongly nonideal electrolytes. Thus it will hopefully be more valid than (5.6) at higher densities.

The condition for mechanical stability can be expressed as $(\partial n_{ex}/\partial n_e)_T > 0$. If this condition is violated, then as one adds more electrons and holes to the system the number of excitons would decrease, clearly leading to an ionization catastrophe. The critical point is determined by the conditions

$$\left. \frac{\partial n_{ex}}{\partial n_e} \right|_{T_c} = \left. \frac{\partial^2 n_{ex}}{\partial n_e^2} \right|_{T_c} = 0 \quad (5.8)$$

Kremp, Ebeling and Kraeft³⁴⁾ have applied this theory (using 5-7 with $C=1/8$) to the electron-hole fluid in Ge and estimate a critical point temperature $T = 9$ K and density $n \approx 10^{15} \text{ cm}^{-3}$. This temperature and density are still within the classical regime. These authors point out that their values are very different to that obtained above for the liquid-gas critical point and support the possibility of two separate critical points.

At present, the question of two critical points in Ge is an open one and this represents one of the outstanding experimental questions which are unresolved at the present time. In the few metals that have been studied near their critical point, Hg and Cs, no evidence of two critical points has been found. However, in Ge and Si the orbital degeneracies and anisotropy in the band structure greatly stabilize

metal and separate the liquid-gas critical point from the region of the metal-to-nonmetal transition, thus affording an opportunity to see two critical points. The observation of the second critical point would be a striking confirmation of the early ideas discussed in the introduction on the metal-insulator transition.

REFERENCES

- 1) L. D. Landau and G. Zeldovich: Acta Phys.-Chim. USSR 18, (1943) 194 [Eng. Trans. in Collected Works of L. D. Landau, ed. by D. TerHaar (Pergamon Press) 1965].
- 2) N. F. Mott: Proc. Phys. Soc. (London) A62, (1949) 416.
- 3) N. F. Mott: *Metal-Insulator Transitions* (Taylor and Francis, Ltd., 1974).
- 4) Ya. E. Pokrovskii: Phys. Stat. sol. (a) 11, (1972) 385.
- 5) L. V. Keldysh: in "*Eksitony v Poluprovodnikah*" (Excitons in Semiconductors) (Nauka, Moscow) 1971, p. 5.
- 6) T. M. Rice, J. C. Hensel, G. A. Thomas and T. G. Phillips (to be published).
- 7) B. I. Halperin and T. M. Rice: in *Solid State Physics*, ed. by F. Seitz, D. Turnbull and H. Ehrenreich (Acad. Press, NY, 1968) 21, 115.
- 8) R. S. Knox: Solid State Phys. Suppl. 5, (1963) 100.
- 9) J. des Cloizeaux: J. Phys. Chem. Solids 26, (1965) 259.
- 10) L. V. Keldysh and Yu. V. Kopaev: Fiz. Tverd. Tela 6, (1964) 2791. [Eng. Trans., Sov. Phys.-Solid State 6, (1965) 2219].
- 11) B. I. Halperin and T. M. Rice: Rev. Mod. Phys. 40, (1968) 755.
- 12) W. Kohn: in *Many-Body Physics*, ed. by C. deWitt and R. Balian (Gordon and Breach, NY, 1968), p. 351.
- 13) E. A. Hylleraas and A. Ore: Phys. Rev. 71, (1947) 493.
- 14) A. Akimoto and E. Hanamura: Solid State Comm. 10, (1972) 253.

- 15) W. F. Brinkman, T. M. Rice and B. J. Bell: Phys. Rev. B8, (1973) 1590.
- 16) W. F. Brinkman and T. M. Rice: Phys. Rev. E7, (1973) 1508.
- 17) J. deBoer, in *Prog. in Low Temp. Phys.*, ed. by C. J. Gorter (North-Holland, Amsterdam 1957), II.1.
- 18) E. Hanamura: in *Proc. of the Tenth Int. Conf. on the Physics of Semiconductors*, Cambridge, Mass., 1970, ed. by S. P. Keller, J. C. Hensel and F. Stern (USAEC, Div. of Tech. Info., 1970), p. 504.
- 19) T. M. Rice: in *Proc. of the Twelfth Int. Conf. on the Physics of Semiconductors*, Stuttgart, 1974, ed. by M. H. Pilkuhn, (Teubner, Stuttgart), p. 23.
- 20) L. V. Keldysh: in *Proc. of the Ninth Int. Conf. on the Physics of Semiconductors*, Moscow, 1968, ed. S.M. Ryvkin (Nauka, Leningrad, 1968) p.1303.
- 21) W. F. Brinkman, T. M. Rice, P. W. Anderson and S.-T. Chui: Phys. Rev. Lett. 28, (1972) 961.
- 22) P. Vashishta, P. Bhattacharyya and K. S. Singwi: Phys. Rev. Lett. 30, (1973) 1248.
- 23) M. Inoue and E. Hanamura: J. Phys. Soc. Japan 35, (1973) 643.
- 24) E. Wigner and H. B. Huntington: J. Chem. Phys. 3, (1935) 764.
- 25) P. Vashishta, P. Bhattacharyya and K. S. Singwi: Phys. Rev. B10, (1974) 5108.
- 26) M. Combescot and P. Nozières: J. Phys. C 5, (1972) 2369.
- 27) P. Vashishta, S. G. Das and K. S. Singwi: Phys. Rev. Lett. 31, (1974) 911.
- 28) T. M. Rice: Il. Nuo. Cim. B23, (1974) 226.
- 29) M. Combescot: Phys. Rev. Lett. 32, (1974) 15.
- 30) T. L. Reinecke and S. C. Ying: Phys. Rev. Lett. 35, (1975) 311.
- 31) G. A. Thomas, T. M. Rice and J. C. Hensel: Phys. Rev. Lett. 33, (1974) 219.

- 32) G. E. Norman and A. N. Starostin: *Tep. Vys. Temp.* 8, (1970) 413.
[Eng. Trans. *High Temperature* 8, (1970) 381].
- 33) W. Ebeling, W. D. Kraeft and D. Kremp: *Theory of Bound States and Ionization Equilibrium in Plasmas and Solids* (Akademie-Verlag, Berlin, 1975).
- 34) D. Kremp, W. Ebeling and W. D. Kraeft: *Phys. Stat. sol. (b)* 69, (1975) K59.