PHYSICAL REVIEW B VOLUME 29, NUMBER 10 15 MAY 1984

Metal-nonmetal transitions in doped silicon

Key Taeck Park, Sung Ran Uhm, and Chul Koo Kim
Department of Physics, Yonsei University, Seoul 120, Korea

(Received 20 April 1983; revised manuscript received 13 December 1983)

The metal-nonmetal transitions in doped silicon are studied with the use of Hubbard's tight-binding model. We include in the calculation the variation in the dielectric constant and the wave function near the transition region. The results show that the change induced by the variation in the dielectric function approximately cancels the effect caused by the change in the effective wave function. This explains why all the previous calculations of the critical concentration for the metal-nonmetal transition, which neglected the variations in the dielectric constant and the wave function, could still yield good agreement with experiment.

From the time Mott[1] first derived a simple criterion for the metal-nonmetal (M-NM) transitions in condensed matter, a wide variety of theoretical and experimental studies have been undertaken to investigate M-NM transitions in heavily doped semiconductors.[2] The critical concentration can be calculated theoretically starting either from the metallic side[3] or from the insulating side.[4] When the M-NM transition is approached from the insulating side, the Hubbard tight-binding Hamiltonian[5] becomes a convenient means for obtaining the critical concentration n_c as demonstrated by Berggren[6] and by Edwards and Sienko.[7] The results obtained agree reasonably well with the experimental data and the Mott criterion calculated starting from the metallic side, even though a somewhat large deviation from the Mott criterion has been reported recently.[8]

In the process of calculation, the dielectric constant and the wave function are assumed to remain the same up to the M-NM transition. However, it is known that the dielectric constant increases sharply when the M-NM transition is approached from the insulating side.[9–11] The possibility of a connection between the M-NM transition and the dielectric divergence was early recognized by Mott and Davis[12] and studied by several investigators.[9–11,13] In this respect, it is not immediately apparent why the calculations of Refs. 6 and 7, which employ a constant dielectric function up to the M-NM transition, are generally successful for a wide range of materials. Indeed, the need for a calculation which used a more realistic dielectric function had been pointed out by Mott.[2] The dielectric function is not the only quantity which varies substantially near the M-NM transition. The ionization energy of a donor state in heavily doped semiconductors, which directly determines the effective Bohr radius, is expected to be quite different from that of a well isolated impurity state.[14,15] It is the aim of this paper to incorporate these variations of the dielectric function and the wave function in the calculation and investigate the possible physical consequences of the changes.

It it generally recognized that the Hubbard Hamiltonian formulation describes a transition from an insulating to a metallic state when the condition

$$W/U \sim 1 \tag{1}$$

is satisfied.[2,14] Here W is the unperturbed bandwidth of an array of one-electron states and U is the intradonor Coulomb repulsion energy between a pair of electrons on the same donor. For the critical value of W/U, we choose Hubbard's value of 1.15.[5,6] In order to apply the above condition to the M-NM transition in shallow donor states of germanium and silicon, Berggren[6] adopted the isotropic assumption that the ellipsoidal wave function may be replaced by a spherical isotropic charge distribution. Although, it is now known that the anisotropy has a non-negligible effect on the critical concentration,[16] we believe that the isotropic envelope function is still sufficient to determine the effect of the enhancement of the dielectric constant and of the wave-function radius on the critical concentration, which is the main purpose of this report. For the isotropic-conduction-band case, the ground-state wave function for the ith donor has the form[17]

$$\Psi_i(\vec{r}) = \sum_{p=1}^{N} \alpha_p F_p(\vec{r}) \phi_p(\vec{r}) \ , \tag{2}$$

where N is the number of conduction-band minima in $\vec{k}$ space and α_p the numerical coefficient of the appropriate linear combination of $F_p(\vec{r})\phi_p(\vec{r})$. In Si, the ground state is the symmetrical combination, $\alpha_p = N^{-1/2}(N=6)$. $\phi_p(\vec{r})$ is the Bloch wave at the pth minimum and $F_p(\vec{r})$ represents the simple hydrogenic envelope function:

$$F_p(\vec{r}) = (\pi a_H^3)^{-1/2} \exp\left(-\frac{r}{a_H}\right) \ , \tag{3}$$

with a_H a realistic (isotropic) Bohr radius for the donor ground state.

The hopping integral T between two adjacent orbitals is defined as

$$T = \int d\vec{r}\, \Psi_i^*(\vec{r}) H_1 \Psi_j(\vec{r}) \ , \tag{4}$$

where H_1 is the one-particle Hamiltonian including the kinetic energy operator and the electron-donor interactions. T is related to the unperturbed bandwidth via

$$W = 2Z|T| \ , \tag{5}$$

where Z is the coordination number. The intradonor Coulomb repulsion energy is given by

$$U = \int\int d\vec{r}_1 d\vec{r}_2 |\Psi_i(\vec{r}_1)|^2 \frac{e^2}{\epsilon_0 r_{12}} |\Psi_i(\vec{r}_2)|^2 \ , \tag{6}$$

where ϵ_0 is the background dielectric constant of the host material with no or negligible donor concentration.

Using the isotropic donor wave function and the hydrogenic evelope function given by Eqs. (1) and (3), Berggren obtained expressions for the intradonor repulsion energy U and the hopping integral T:

$$U=\int\int d\vec{r}_1 d\vec{r}_2\left(\frac{1}{N}\sum_{p=1}^{N}F_p^2(\vec{r}_1)\right)\frac{e^2}{\epsilon_0 r_{12}}\left(\frac{1}{N}\sum_{p=1}^{N}F_p^2(\vec{r}_2)\right)$$

$$=\int\int d\vec{r}_1 d\vec{r}_2|F_p(\vec{r}_1)|^2\frac{e^2}{\epsilon_0 r_{12}}|F_p(\vec{r}_2)|^2$$

$$=\frac{5}{8}\frac{e^2}{\epsilon_0 a_H}\ , \tag{7}$$

$$T=\int d\vec{r}\,\Psi_i^*(\vec{r})\left(-\frac{\hbar^2}{2m^*}\nabla^2-\frac{e^2}{\epsilon_0|\vec{r}-\vec{R}_i|}-\frac{e^2}{\epsilon_0|\vec{r}-\vec{R}_j|}\right)\Psi_j(\vec{r})$$

$$=-\frac{e^2}{2\epsilon_0 a_H}\langle\Psi_i|\Psi_j\rangle-\left\langle\Psi_i\left|\frac{e^2}{\epsilon_0|\vec{r}-\vec{R}_i|}\right|\Psi_j\right\rangle$$

$$=-\frac{e^2}{2\epsilon_0 a_H}S-L\ , \tag{8}$$

and

$$S=\frac{1}{N}\sum_{p=1}^{N}\exp(-i\vec{p}\cdot\vec{R})\left[1+\frac{R}{a_H}+\frac{1}{3}\left(\frac{R}{a_H}\right)^2\right]\exp\left(-\frac{R}{a_H}\right)\ , \tag{9}$$

$$L=\frac{e^2}{\epsilon_0 a_H}\frac{1}{N}\sum_{p=1}^{N}\exp(-i\vec{p}\cdot\vec{R})\left(1+\frac{R}{a_H}\right)\exp\left(-\frac{R}{a_H}\right)\ , \tag{10}$$

where R is the separation between donor states. In obtaining Eqs. (9) and (10), Berggren neglected terms for which the integrand is rapidly oscillating. For the evaluation of the $(1/N)\sum_p^N\exp(-i\vec{p}\cdot\vec{R})$ term, we follow the scheme used by Bhatt,[18] instead of using the method by Berggren who obtained the value of $1/\sqrt{N}$ by simply squaring it and neglecting all oscillating terms. Since $|T|$, and not T, determines W, the above term is reduced to

$$|\cos k_0X+\cos k_0Y+\cos k_0Z|/3\ ,$$

which is approximately 0.33.[18] Here k_0 is the magnitude of the wave vectors of the conduction valleys. This gives a reduction of the bandwidth by a factor of 0.82 relative to that obtained by Berggren.

Calculated values of the critical concentration n_c using the above expressions are generally in good agreement[7] with the experimental values despite several simplifying approximations made deriving the above results.

Here, we concern ourselves only with the fact that the dielectric constant and the wave function vary substantially near the M-NM transition, instead of remaining constant as assumed in the above calculations. In order to incorporate the change in the dielectric constant into the calculation, it is necessary to replace the constant dielectric function by spatially varying and concentration-dependent dielectric function. For this purpose, we employ a smoothly varying effective dielectric function $\epsilon(n,r)$, which has been proposed and applied successfully for the calculation of the donor polarization enhancement by Castner.[19] The effective dielectric function without the central cell correction has the form

$$\frac{1}{\epsilon(n,r)}=\frac{1}{\epsilon_0}\exp\left(-\frac{r}{r_s}\right)+\frac{1}{\epsilon(n)}\left[1-\exp\left(-\frac{r}{r_s}\right)\right]\ , \tag{11}$$

where $4\pi r_s^3n/3=1$, and $\epsilon(n)$ is the static dielectric constant with a donor concentration of n. The justification for using this effective dielectric function for a donor system has been amply discussed by Castner. $\epsilon(n)$ has been measured in silicon for various donors by Castner and his co-workers[9] and also for P by Hess *et al.*[11]

With the concentration-dependent spatial dielectric function properly defined, we now proceed to calculate U and T. Calculations of U and T can be carried out in a somewhat involved but quite similar way to the case with a constant dielectric function.[6,20] The intradonor repulsion energy is given by

$$U=\int\int d\vec{r}_1 d\vec{r}_2|\Psi_i(\vec{r}_1)|^2\frac{e^2}{\epsilon(n,r)r_{12}}|\Psi_i(\vec{r}_2)|^2$$

$$=\int\int d\vec{r}_1 d\vec{r}_2|F_p(\vec{r}_1)|^2\frac{e^2}{\epsilon(n,r)r_{12}}|F_p(\vec{r}_2)|^2$$

$$=\frac{e^2}{\epsilon(n)a_H}\left[\frac{5}{8}+8C\left(\frac{\epsilon(n)}{\epsilon_0}-1\right)\left(\frac{1}{16}-\frac{C}{2}+8CD\right)\right]\ , \tag{12}$$

where $C=[(a_H/r_s)^2-4]^{-1}$ and $D=(2+a_H/r_s)^{-2}$.

In order to obtain the hopping energy, we first calculate L, which is given as follows after some manipulation.

$$L=\int d\vec{r}\,\Psi_i^*(\vec{r})\frac{e^2}{\epsilon(n,r)|\vec{r}-\vec{R}_i|}\Psi_j(\vec{r})$$

$$=\frac{1}{N}\sum_{p=1}^{N}\exp(-i\vec{p}\cdot\vec{R})\int d\vec{r}\,F_i^*(\vec{r})\frac{e^2}{\epsilon(n,r)|\vec{r}-\vec{R}_j|}F_j(\vec{r})$$

$$=\frac{e^2}{\epsilon(n)a_H}\frac{1}{N}\sum_{p=1}^{N}\exp(-i\vec{p}\cdot\vec{R})\left\{2K_1\left(\frac{\epsilon(n)}{\epsilon_0}-1\right)\exp\left(-\frac{R}{r_s}\right)+\left[K_1\left(\frac{\epsilon(n)}{\epsilon_0}-1\right)\left(\frac{2R}{r_s}+\frac{Ra_H}{r_s^2}-2\right)+1+\frac{R}{a_H}\right]\right\}\exp\left(-\frac{R}{a_H}\right)\ , \tag{13}$$

where

$$K_1=\frac{4r_s^2}{Ra_H}\left(2+\frac{a_H}{r_s}\right)^{-2}\ .$$

TABLE I. Critical concentrations n_c and the constants used in the calculation.

Host donor	Sb	Si *p*	As
	11.4	11.4	11.4
$a_H(0)$[a]	16.6 Å	15.2 Å	14.1 Å
n_c[b] (theory)	3.0×10^{18} cm^{-3}	3.9×10^{18} cm^{-3}	4.9×10^{18} cm^{-3}
n_c[c] (theory)	2.8×10^{18} cm^{-3}	3.6×10^{18} cm^{-3}	4.8×10^{18} cm^{-3}
n_c (experiment)	3.0×10^{18} cm^{-3} [d]	3.5×10^{18} cm^{-3} [d]	6.4×10^{18} cm^{-3} [d]
		3.7×10^{18} cm^{-3} [e]	8.5×10^{18} cm^{-3} [f]

[a] From Ref. 7.
[b] Calculation with constant ϵ_0 and $a_H(0)$.
[c] Present calculation.
[d] From Ref. 9.
[e] From Ref. 8.
[f] From Ref. 22.

Now T is given by

$$T = -E(n)S - L \ . \tag{14}$$

Here $E(n)$ is the concentration-dependent donor binding energy to be determined variationally below, and S is defined in Eq. (9).

Before calculating the critical concentration n_c using the above expressions, we now consider the variation of the wave function near the M-NM transition.

The wave function of a donor state is known to vary substantially when the donor concentration approaches the critical value for the M-NM transition.[14,15] This variation can be conveniently represented in our formulation by replacing the constant a_H by a donor concentration-dependent $a_H(n)$. We obtain $a_H(n)$ using the variational method. For the variational calculation, we use the isotropic envelope function,

$$\Psi[r,a_H(n)] = [\pi a_H(n)^3]^{-1/2}\exp[-r/a_H(n)] \ ,$$

and the relation $E(n) = \langle\Psi|H_1|\Psi\rangle$ and minimize $E(n)$ with respect to $a_H(n)$.[19] Here H_1 is the one-particle Hamiltonian with the effective dielectric function $\epsilon(n,r)$. The $E(n)$ obtained in this method is a single-valley result. However, we do not try to include the many-valley effect in $E(n)$ and $a_H(n)$, because the many-valley correction obtained by Castner in Ref. 19 is rather small and does not change the essential physical picture of our calculations. For a more precise determination of $E(n)$ including the many-valley effect, we refer to Ref. 15.

We have calculated n_c for Sb, P, and As donors in silicon using the above expressions.[21] The results obtained and the constants used in the calculation are given and compared to experimental results[8,9,22] in Table I. As we can see from the table, the results from the present calculation are in reasonable agreement with those obtained using the constant dielectric constant and wave functions. In Fig. 1, we plot U and W as functions of the donor concentration for Sb doped silicon. We observe that, although the behavior of U and W in the present calculation is quite different from that with constant ϵ_0 and $a_H(0)$, the critical concentration n_c does not change much from the value obtained with ϵ_0 and $a_H(0)$. To identify the effects caused by variations of $\epsilon(n)$ and $a_H(n)$ more clearly, in Fig. 2 we have considered the changes in $\epsilon(n)$ and $a_H(n)$ separately. When we allow $\epsilon(n)$ to vary while keeping $a_H(n)$ equal to $a_H(0)$, we find that n_c shifts toward the higher concentration. This shift comes from the fact that the reduction in T due to the enhancement of the dielectric function will be larger than that of U. On the other hand, if we allow $a_H(n)$ to vary, while keeping $\epsilon(n,r)$ equal to ϵ_0, n_c moves toward a lower value. This is due to the fact that a larger Bohr radius causes a larger overlap and, consequently, rapid broadening of the bands. When these two effects are combined together, the resulting critical concentration does not differ much from the original value, as shown in Fig. 1. This fact explains clearly why all the previous calculations, which neglected the divergence and change in the wave function,

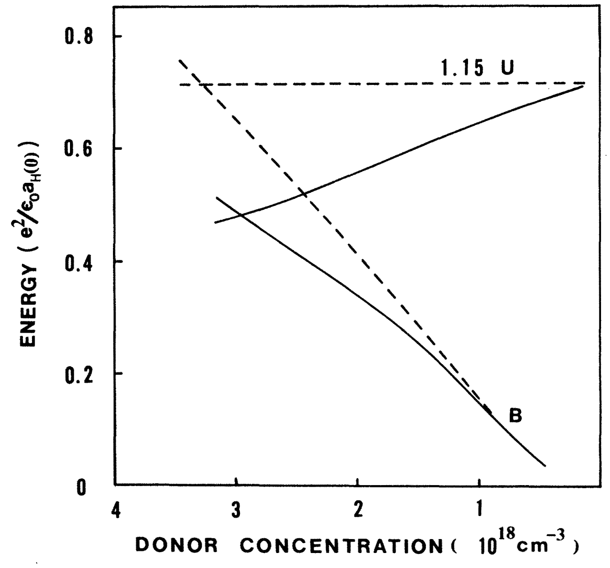

FIG. 1. Variations of the intradonor repulsion energy U and the bandwidth W as functions of the donor concentration when the donors are assumed to have a simple cubic structure. U has been multiplied by 1.15 so that the critical concentration n_c may be obtained directly from the graph. The concentration where the two curves cross corresponds to the critical concentration. The solid line was obtained using $\epsilon(n)$ and $a_H(n)$, while the dotted line relates to ϵ_0 and $a_H(0)$.

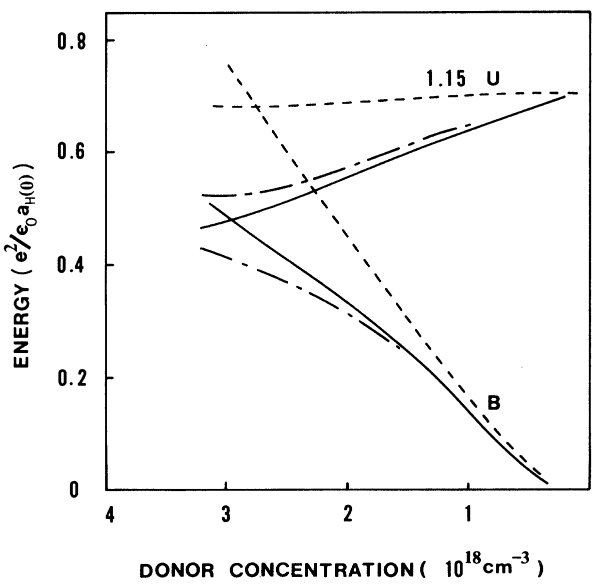

FIG. 2. Critical concentration as a function of the dielectric constant and the wave function considered separately. The solid line was obtained using $\epsilon(n)$ and $a_H(n)$, while the dotted line relates to ϵ_0 and $a_H(n)$. The dashed line is the result for $a_H(0)$ and $\epsilon(n)$.

yielded essentially correct results.[6,7]

Recently, a qualitative estimate of the change caused by the dielectric constant enhancement has been made by Mott and Davies.[23] By employing the Wigner-Seitz method and a simplified spatially varying dielectric function, they obtained a small shift of n_c toward a lower value upon inclusion of the dielectric enhancement, which is in agreement with our calculation. Another recent calculation on the critical concentration by Bhatt and Rice[15] also includes the dielectric and wave-function variations through a variational calculation of the Schrödinger equation. However, since the dielectric function does not appear explicitly in their calculation, their calculation may not be directly comparable to our calculation.

In conclusion, we have presented a theory for the M-NM transition in heavily doped semiconductors which incorporates the concentration dependency of the static dielectric constant and the wave function. Even though our theory does not include the donor randomness or the central cell correction, and employs an oversimplified wave function, it gives a plausible account of the physics involved. It is shown that the change induced by varying the dielectric constant approximately cancels the effect caused by the varying effective Bohr radius, thus explaining why a simple theory with constant Bohr radius and dielectric constant should give essentially correct result.

---

[1] N. F. Mott, Proc. Phys. Soc. London 62, 416 (1949).
[2] See, N. F. Mott, *Metal-Insulator Transitions* (Taylor and Francis, London, 1974).
[3] N. F. Mott, Can. J. Phys. 34, 1356 (1956).
[4] C. Kittel, *Introduction to Solid State Physics,* 4th ed. (Wiley, New York, 1971), p. 723.
[5] J. Hubbard, Proc. R. Soc. London Ser. A 276, 238 (1963); 281, 49 (1964).
[6] K. F. Berggren, Philos. Mag. 27, 1027 (1973).
[7] P. P. Edwards and M. J. Sienko, Phys. Rev. B 17, 2575 (1978).
[8] P. F. Newman and D. F. Holcomb, Phys. Rev. B 28, 638 (1983).
[9] T. G. Castner, Solid State Commun. 32, 121 (1979).
[10] M. Cappizzi, G. A. Thomas, F. Derosa, R. N. Bhatt, and T. M. Rice, Phys. Rev. Lett. 44, 1019 (1980).
[11] H. F. Hess, K. Deconde, T. F. Rosenbaum, and G. A. Thomas, Phys. Rev. B 25, 5573 (1982).
[12] N. F. Mott and E. A. Davis, Philos. Mag. 17, 1269 (1968).
[13] L. G. Caron and G. Kemeny, Phys. Rev. B 3, 3007 (1971).
[14] See, for example, *Proceedings of the Nineteenth Scottish Universities Summer School in Physics, St. Andrews, 1978,* edited by L. R. Friedman and D. P. Tunstall (SUSSP, Edinburgh, 1978).
[15] R. N. Bhatt and T. M. Rice, Phys. Rev. B 23, 1920 (1981).
[16] R. N. Bhatt, Phys. Rev. B 24, 3630 (1981).
[17] W. Kohn and J. M. Luttinger, Phys. Rev. 97, 883 (1955).
[18] R. N. Bhatt, Phys. Rev. B 26, 1082 (1982).
[19] T. G. Castner, Phys. Rev. B 21, 3523 (1980).
[20] J. C. Slater, *Quantum Theory of Molecules and Solids* (McGraw-Hill, New York, 1963), Vol. 1.
[21] The calculations have been carried out and averaged for different arrangements of the donors (simple cubic, body-centered cubic, and face-centered cubic).
[22] T. F. Rosenbaum, K. Andres, G. A. Thomas, and R. N. Bhatt, Phys. Rev. Lett. 45, 1723 (1980).
[23] N. F. Mott and J. H. Davies, Philos. Mag. B 42, 845 (1980).