PHYSICAL REVIEW B VOLUME 21, NUMBER 8 15 APRIL 1980

Donor-polarizability enhancement as the insulator-metal transition is approached from the insulating side

T. G. Castner
University of Rochester, Rochester, New York 14627

(Received 20 July 1978; revised manuscript received 9 October 1979)

The low-temperature static dielectric constant for n-type semiconductors shows a significant upward deviation from Clausius-Mossotti behavior as the donor concentration N approaches the critical value N_c for the onset of metallic behavior. After reviewing the possible reasons for this upward deviation and the various types of screening applied to the insulator-metal problem, a new continuum-model calculation of the donor-polarizability enhancement $\alpha_D(N)/\alpha_D(0)$ is presented which is based on an electrostatic donor-donor interaction potential and the Clausius-Mossotti relationship for the long-wavelength static dielectric constant $\epsilon(N)$. The Hassé variational approach is employed to calculate $\alpha_D(N)$. The calculated enhancement contains two contributions: (1) a size effect associated with an enlargement of the donor wave function with N, and (2) an increase in the average effective dielectric constant $\langle \epsilon(r,N) \rangle$ with N seen by the localized donor wave function. The results show $\alpha_D(N_c)/\alpha_D(0)$ ranging from about 2 for Si:Sb to about 3 for Si:As. Comparison with the experimental data for Si:As shows the calculated upward deviation from Clausius-Mossotti behavior remains smaller than the data. The values of $N_{c\text{-calc}}$ [based on a polarization catastrophe, a new self-consistent Clausius-Mossotti relationship, and a self-consistent Herzfeld criterion for N_c from $\epsilon(N \rightarrow N_c) \rightarrow \infty$] are between 60% and 100% larger than the experimental values for the silicon shallow donors. Possible corrections to this calculation resulting from the donor randomness, charge transfer, and corrections to the Lorentz-Lorenz local field are also discussed.

I. INTRODUCTION

The metal-nonmetal (MNM) transition in monovalent impurity atom-insulator host systems such as n-type semiconductors, alkali or noble-metal atoms in the solid rare gases, the tungsten bronzes, metal ammonia solutions, and others have been the subject of intensive experimental and theoretical investigation.[1] The principal experimental approaches employed have involved transport-properties studies,[2] magnetic-properties studies,[3] and more recently optical investigations.[4-6] The dielectric properties, as $N \rightarrow N_c$ from the insulating side of the transition, have received less attention,[7,8] but nonetheless might be expected to play an important role in advancing our understanding of the nature of the MNM transition. For the case of an isolated shallow-donor impurity atom having a binding energy, neglecting central-cell corrections, of magnitude $m^*e^4/2\hbar^2\epsilon_h^2$ (m^* is an isotropic effective mass and ϵ_h the static dielectric constant of the host semiconductor) the host-dielectric constant ϵ_h plays the familiar role of screening the attractive Coulomb potential and thereby substantially reduces the binding energy when $\epsilon_h \gg 1$. Less well established is the question of how this screening changes with doping as $N \rightarrow N_c$. Recent experimental evidence[9] has suggested that the low-temperature, low-frequency dielectric constant $\epsilon(N, T \rightarrow 0)$ rises more rapidly as $N \rightarrow N_c$ than predicted by the Clausius-Mossotti relationship and has led to the idea of a polarizability enhancement as a possible explanation for the "upward" deviation from Clausius-Mossotti behavior. The present work is concerned therefore primarily with how the screening changes with concentration on the insulator side of the MNM transition as $N \rightarrow N_c$. This work also suggests that screening plays the dominant role in reducing the activation energies for transport and might properly be viewed as the driving force for the insulator-metal transition.

Since screening (Thomas-Fermi) was the mechanism originally employed by Mott[10] to obtain his remarkably accurate estimate $N_c^{1/3}a \approx \frac{1}{4}$, where a is the Bohr radius of the isolated impurity-electron wave function, many authors have employed various types of screening to calculate $N_c^{1/3}a$ and other quantities relating to the MNM transition (Edwards and Sienko[11] have recently shown for the monovalent MNM systems studied experimentally that $N_c^{1/3}a = 0.26 \pm 0.05$). The three basic types of screening applied to the MNM transition case have been the following: (1) metallic screening such as Thomas-Fermi, Lindhard, Hubbard-Sham, and related cases resulting from strongly degenerate free carriers,[10,12-14] (2) classical semiconductor screening of the Debye-Hückel type resulting from thermally activated free carriers in a semiconductor,[15-19] and (3) classical insulator screening based on the venerable Clausius-Mossotti expression[20] incorporating the Lorentz-Lorenz local-field correction for localized electrons.[21-25] Recently Takeshima[26] has attempted to apply a combination of metallic screening (1) (resulting from a small

© 1980 The American Physical Society

number of free carriers) and insulator screening (3) to calculate the dielectric response on the insulating ($N < N_c$) side of the transition. While it will not be possible to fully discuss the broad range of ideas embodied in the various approaches mentioned above, an attempt will be made to compare the present approach with these earlier approaches and with the experimental data.

The physical situation to be explored herein is the case of a random (probably not fully random) distribution of monovalent impurities in an insulating host which itself can be considered as a continuum dielectric. By considering only a single type of impurity we specifically exclude compensation effects such as random internal electric fields. Within the framework of the Anderson model for electron localization the randomness implies predominantly lateral or off-diagonal disorder. The temperature range will be restricted to the ultralow temperature limit ($T \to 0$ K), or effectively to temperatures such that $kT \ll \epsilon_2$ or any hopping activation energy. This situation is that originally envisaged by Mott in which the average atomic spacing is reduced at $T = 0$ K until N exceeds the critical value N_c for metallic behavior. An expression for the classical screening length developed by Stern[27] diverges as $T \to 0$ K as the compensation becomes negligible. Thus in the present case the Debye-Hückel screening makes no contribution and our attention will be restricted to the third type of screening and to the dielectric behavior of insulators.

The modern approach to the self-consistent dielectric response of solids has been formulated by Ehrenreich and Cohen[28] and extended by others. This approach amounts to a perturbation theory involving a summation over matrix elements between occupied states and low-lying empty states. This approach has been employed by Wiser[29] to demonstrate that the Lorentz-Lorenz local-field correction $4\pi p/3\epsilon_h$ is valid for localized electrons (in a cubic lattice), but that no local-field correction is required for itinerant electrons. However, rather than use this modern perturbation-theory approach, we will utilize the Clausius-Mossotti expression and extend it to include the donor-polarizability enhancement. It is to be emphasized that, although the Clausius-Mossotti expression is frequently viewed as a classical result which relates the macroscopic dielectric constant to the microscopic polarizability, the polarizability must be calculated with quantum mechanics and must include a self-consistent potential which incorporates interactions between the impurities. The donor polarizability $\alpha_D(N)$ is calculated using the variational method of Hassé.[30] The results yield a self-consistent Clausius-Mossotti equation which qualitatively explains the upward deviations from standard Clausius-Mossotti behavior observed in the data[9] on n-type Si. The effects of the randomness on the polarization and N_c will also be briefly considered.

For the case of induced dipoles the Clausius-Mossotti relationship predicts a divergence of the static dielectric constant $\epsilon(N)$ as $4\pi N\alpha/3\epsilon_h \to 1$ with a resulting polarization catastrophe (PC). This feature has resulted in substantial criticism of the Clausius-Mossotti expression and it is frequently asserted[20,31,32] that this divergence cannot occur for induced dipoles. It is perhaps surprising that the early application of the Clausius-Mossotti expression, using measured gaseous molar polarizabilities and known solid densities, provided Herzfeld[21] with a reasonably accurate prediction of which elements were metallic and which were insulating in the solid state. The present work extends Herzfeld's application of the Clausius-Mossotti result to the MNM transition by making it self-consistent and also examines some crucial questions concerning the applicability of the Clausius-Mossotti relationship as $N \to N_c$.

II. BACKGROUND

For uncompensated n-type semiconductors containing N_D (we shall henceforth drop the D subscript) neutral donors per unit volume at zero temperature ($T = 0$ K) the polarization per unit volume $\vec{P}$ will be given by

$$\vec{P}(N) = \sum_{i=1}^{N} \overleftrightarrow{\alpha}_i(\vec{r}_i, N) \cdot \vec{E}_{i\text{-loc}}(\vec{r}_i, N) , \tag{1}$$

where $\overleftrightarrow{\alpha}_i(\vec{r}_i)$ is the polarizability tensor of the ith donor at position $\vec{r}_i$ and $\vec{E}_{i\text{-loc}}(\vec{r}_i)$ is the local field at the ith donor. By averaging over a large number of neutral donors in a unit volume the average values for $\overleftrightarrow{\alpha}_i(\vec{r}_i, N)$ and $\vec{E}_{i\text{-loc}}(\vec{r}_i, N)$ are given by

$$\langle \alpha_D(N) \rangle \overleftrightarrow{1} = \frac{1}{N} \sum_{i=1}^{N} \overleftrightarrow{\alpha}_i(\vec{r}_i, N) , \tag{2a}$$

$$\begin{aligned} \langle \vec{E}_{\text{loc}}(N) \rangle &= \frac{1}{N} \sum_{i=1}^{N} \vec{E}_{i\text{-loc}}(\vec{r}_i, N) \\ &= \vec{E}_x + 4\pi \vec{P}(N)/3\epsilon_h , \end{aligned} \tag{2b}$$

where $\langle \alpha_D(N) \rangle$ is a density-dependent average scalar polarizability and the density-dependent average local field is the normal Lorentz local field expected for a cubic array of donors in a continuous-host medium of the static dielectric constant ϵ_h. Even though $\vec{P}_i$ and $\vec{E}_{i\text{-loc}}$ are not parallel to the external field $\vec{E}_x$, both $\vec{P}$ and $\langle \vec{E}_{\text{loc}}(N) \rangle$ will be parallel to $\vec{E}_x$. By considering the fluctuations of the local polarizability $\overleftrightarrow{\alpha}_i(\vec{r}_i, N)$ and $\vec{E}_{i\text{-loc}}(\vec{r}_i, N)$ from the average values $\langle \alpha_D(N) \rangle$ and $\langle \vec{E}_{\text{loc}}(N) \rangle$, $\vec{P}(N)$

from Eq. (1) can be written as

$$\begin{aligned}\vec{P}(N) = {} & N\langle\alpha_D(N)\rangle\langle\vec{E}_{\text{loc}}(N)\rangle \\ & + \sum_{i=1}^{N} \overleftrightarrow{\alpha}_i(\vec{r}_i, N) - \langle\alpha_D(N)\rangle\overleftrightarrow{1}] \cdot \langle\vec{E}_{\text{loc}}(N)\rangle \\ & + \langle\alpha_D(N)\rangle\overleftrightarrow{1} \cdot \sum_{i=1}^{N} \delta\vec{E}_{i\text{-loc}}(\vec{r}_i, N) \\ & + \sum_{i=1}^{N} [\overleftrightarrow{\alpha}_i(\vec{r}_i, N) - \langle\alpha_D(N)\rangle\overleftrightarrow{1}] \cdot \delta\vec{E}_{i\text{-loc}}(\vec{r}_i, N) \,,\end{aligned} \tag{3}$$

where $\delta\vec{E}_{i\text{-loc}}(\vec{r}_i, N) = \vec{E}_{i\text{-loc}}(\vec{r}_i, N) - \langle\vec{E}_{\text{loc}}(N)\rangle$. Because of Eqs. (2a) and (2b) the second and third terms vanish. The last term represents the correlation between polarizability and local-field fluctuations. $\vec{P}(N)$ can be written as

$$\begin{aligned}\vec{P}(N) = {} & N\{\langle\alpha_D(N)\rangle\langle\vec{E}_{\text{loc}}(N)\rangle \\ & + \langle[\overleftrightarrow{\alpha}(\vec{r}, N) - \langle\alpha_D(N)\rangle\overleftrightarrow{1}] \cdot \delta\vec{E}_{\text{loc}}(\vec{r}, N)\rangle_c\} \,,\end{aligned} \tag{4}$$

where $\langle[\overleftrightarrow{\alpha}(r, N) - \alpha_D(N)\overleftrightarrow{1}] \cdot \delta\vec{E}_{\text{loc}}(\vec{r}, N)\rangle_c$ is the average correlation per donor of the two fluctuating quantities. The determination of the magnitude of the second term in Eq. (4) is a difficult problem and will not be considered in this work. This term incorporates the effects of the spatial randomness on the dielectric properties of doped semiconductors on the insulating side of the transition ($N < N_c$). The relative magnitude of this term compared to the first term $N\langle\alpha_D(N)\rangle\langle\vec{E}_{\text{loc}}(N)\rangle$ [which we designate the regular array polarization $\vec{P}_r(N)$] has not been established and will depend on the nature of the randomness. For a totally random (Poisson) distribution the second term will be far more important than for a more restricted type of randomness. Although $\vec{P}_r(N) = \vec{P}(N)$ for a cubic array of donors we expect $\vec{P}(N)$ to be larger than $\vec{P}_r(N)$ for a random array. While not calculating $\vec{P}(N) - \vec{P}_r(N)$ here, it is worth making some qualitative comments about the fluctuation correlation term in Eq. (4).

The largest contribution to $\langle[\overleftrightarrow{\alpha}(\vec{r}_i, N) - \langle\alpha_D(N)\rangle\overleftrightarrow{1}] \cdot \delta E_{\text{loc}}(\vec{r}_i)\rangle_c$ will come from the immediate neighbors of a particular ith donor. For a particular coordination number z, for example, 6, one could consider an XY^6 octahedral complex and expand the random positions of seven donors making up the XY^6 complex in terms of the various normal modes of this complex as developed by Van Vleck.[33] The Q_1 breathing mode (donor density fluctuation) will contribute to $\langle\overleftrightarrow{\alpha}(r_x, N) - \alpha_D(N)\overleftrightarrow{1}\rangle$ by changing the donor-donor interactions and the magnitude of the screening [$V_{\text{int}}(r, N)$ in Eq. (10)], but will not contribute to $\delta\vec{E}_{\text{loc}}(\vec{r}_x, N)$ because there is no change from cubic symmetry. One can also demonstrate that rotations and translations of Y^6 relative to X make only zero or small contributions to $\delta\vec{E}_{\text{loc}}(\vec{r}_x, N)$ and $\overleftrightarrow{\alpha}(\vec{r}_x, N) - \langle\alpha_D(N)\rangle\overleftrightarrow{1}$. However, tetragonal distortions (modes Q_2 and Q_3 in Ref. 33) will make first-order contributions to both $\delta\vec{E}_{\text{loc}}(\vec{r}_x, N)$ and $\overleftrightarrow{\alpha}(\vec{r}_x, N) - \langle\alpha_D(N)\rangle\overleftrightarrow{1}$ and hence to $\langle[\overleftrightarrow{\alpha}(r, N) - \alpha_D(N)\overleftrightarrow{1}] \cdot \delta E_{\text{loc}}(\vec{r}, N)\rangle$. Although this XY^6 complex approach might allow one to estimate the effects of randomness, it is not going to lead to simple analytical formulas which transparently demonstrate the consequence of randomness. One can more easily calculate $\langle[\delta\vec{E}_{\text{loc}}(\vec{r}_i, N)]^2\rangle$ for positional deviations from a cubic array. Employing a Poisson distribution one finds $\{\langle[\delta\vec{E}(\vec{r}_i, N)]^2\rangle\}^{1/2} \approx 0.45\, 4\pi\vec{P}(N)/3\epsilon_n$. This result shows that local-field fluctuations can be comparable to the average Lorentz local field $4\pi\vec{P}(N)/3\epsilon_n$ with the rms local-field fluctuation nearly one-half of $4\pi\vec{P}(N)/3\epsilon_n$. The difficult problem is to determine how $\overleftrightarrow{\alpha}(\vec{r}_i, N) - \alpha_D(N)\overleftrightarrow{1}$ is correlated with $\delta\vec{E}_{\text{loc}}(\vec{r}_i, N)$. From the above discussion this correlation is clearly less than unity and this might suggest that the fluctuation-correlation term in Eq. (4) is less important than the first term $\vec{P}_r(N)$. The viewpoint may appear opposed to the generally accepted idea that spatial fluctuations represent an important or even dominant feature of the MNM transition for systems with random distributions of impurities. However, it is worth emphasizing that the static long-wavelength dielectric constant is a physical quantity that is averaged over distances much larger than the average distance between donors. As $N \to N_c$ the Lorentz contribution to $\vec{E}_{\text{loc}}(\vec{r}_i, N)$ becomes increasingly important and is dominated by the long-range $4\pi\vec{P}(N)/3\epsilon_n$, where $\vec{P}(N)$ is the total average polarization. Other physical properties, which are also averaged over the random donor distribution, may feature predominantly short-range interactions and therefore may be much more affected by randomness than the long-wavelength dielectric constant. The magnetic susceptibility should have important corrections from the randomness because of the short-range nature of the exchange interaction. The randomness plays a key role for the ESR spectra[34] of pairs and larger clusters. One might argue that if randomness does not greatly affect the long-wavelength dielectric constant, which helps determine the average binding-energy $E_{1s\text{-}A_1}(N)$ of the shallow donors, then randomness might not have a large effect on the MNM transition criterion $N_c^{1/3}a \sim \frac{1}{4}$. This viewpoint is supported by calculations, such as the Hubbard model,[11] which apparently yield results for $N_c^{1/3}a$ close to the experimental values.

The experimental results[9] for Si:As demonstrated an upward deviation from Clausius-Mossotti behavior given by

$$\frac{\epsilon(N) - \epsilon_h}{\epsilon(N) + 2\epsilon_h} = \frac{4\pi N\alpha_D}{3\epsilon_h} \,, \tag{5}$$

where α_D is a constant-donor polarizability. Three possible reasons for this upward deviation were offered, namely: (i) an increase in $\alpha_D(N)$ resulting from a weakening attractive potential and binding energy as $N \to N_c$, (ii) fluctuations in $\vec{E}_{\text{loc}}(\vec{r}_i)$ resulting from the random distribution of donors, and (iii) cluster effects as observed by ESR.[34] Let us review the various possibilities for deviations from Clausius-Mossotti behavior before proceeding with the calculation of $\langle\alpha_D(N)\rangle$.

Corrections to Eq. (5) have been treated by many authors and have been discussed by Böttcher.[32] Here we consider only nonpolar impurities and thereby neglect the correction to Eq. (5) from the Onsager reaction field for permanent dipoles. Even for the nonpolar case several types of corrections have been considered (see Ref. 32) and these will not be detailed here. Perhaps the most relevant correction to consider is that correction resulting from "translational fluctuations" as applied to gases and liquids. This correction was derived by Böttcher,[35] employing a continuum model, and also from a virial expansion $(\epsilon - 1)/(\epsilon + 2) = (4\pi N\alpha/3)(1 + S)$, where

$$S = \alpha^2(b_1 N + b_2 N^2 + \cdots) + O(\alpha^3)\,, \tag{6}$$

as given by de Boer, van der Maesen, and Ten Seldam.[36] The coefficients b_1, b_2, etc., can be calculated for a particular intermolecular interaction and the above authors gave results for Lennard-Jones and square-well potentials. Kirkwood[37] evaluated these coefficients for a hard-sphere intermolecular interaction and obtained $b_1 = 8\pi/3\sigma^3$ and $b_2 = -5\pi^2/3$ where σ is the hard-sphere diameter. The correction to the Clausius-Mossotti equation from Böttcher's continuum approach and from Kirkwood's coefficients leads to an expression correct to order α^2 in S given by

$$\frac{\epsilon - 1}{\epsilon + 2} = z\left(1 + \frac{2\alpha z}{\sigma^3} - \tfrac{15}{16} z^2\right), \tag{7}$$

where $z = 4\pi N\alpha/3$. This equation with the left-hand side replaced by $(\epsilon - \epsilon_h)/(\epsilon + 2\epsilon_h)$ and with $z = 4\pi N\alpha/3\epsilon_h$ was adapted by Van Metter[38] to attempt to fit the corrections to Clausius-Mossotti behavior for the Si:As data.[9] A reasonable fit was obtained for $\sigma = 15$ Å. However, the same σ did not appear to give a good fit for the more restricted Si:P and Si:Sb data. Regardless of this reasonable fit for the Si:As data there remains the serious question of whether a modified equation (7) is correct for a random distribution of fixed impurities in a solid at 0 K.

Employing ESR, the Zeeman spectra of donor pairs, triplets, and perhaps larger clusters have been observed and identified.[34] However, as a precautionary warning, these discrete clusters have only been observed for $N \ll N_c$ ($N < 6 \times 10^{17}$ for Si:P) and as $N \to N_c$ a single exchange-narrowed line is observed and the experimental information on separate clusters of a given size is lost. If one attempted to write $\vec{P}(N)$ in terms of a cluster expansion, one would have

$$\vec{P}(N) = \sum_{m=1} N^m \langle\alpha^m\rangle\langle\vec{E}^m_{\text{loc}}\rangle\,, \tag{8}$$

where $N = \sum_{m=1} N^m$, $\langle\alpha^m\rangle$ is the average polarizability of a cluster of m donors, N^m is the number of such m-donor clusters per unit volume, and $\langle\vec{E}^m_{\text{loc}}\rangle$ is the mean local field for the m-donor cluster. Even if one could argue that $\langle\vec{E}^m_{\text{loc}}\rangle$ was independent of the cluster size one would still face a formidable problem in determining the $\langle\alpha^m\rangle$, even though the N^m are readily calculated for a Poisson distribution. There is no convincing way of extracting the $\langle\alpha^m\rangle$ from the data of $\epsilon(N)$ vs N. In fact, if the first reason for the deviation from Clausius-Mossotti behavior is important, then the $\langle\alpha^m\rangle$ themselves would be functions of N and Eq. (8) would represent a very complex virial expansion. Any attempt to fit the data with Eq. (8) would appear to be speculative at best. Moreover, since the largest contribution to the local field $\vec{E}_{\text{loc}}(\vec{r}_i, N)$ (besides $\vec{E}_x$) comes from the long-range $4\pi\vec{P}(N)/3\epsilon_h$ term rather than from the much shorter-range $\delta E_{\text{loc}}(\vec{r}_i, N)$ from spatial fluctuations, one might well question whether the concept of independently behaving clusters (albeit large clusters) is sound as $N \to N_c$. Despite this shortcoming several comments should be made concerning $\langle\alpha^2\rangle$ and $\langle\alpha^3\rangle$. The polarizability of a hydrogenic donor molecule $\alpha^2(R)$ has been considered by Takeshima[26] as a function of the internuclear distance R ($\alpha^2(R) = \frac{1}{3}[\alpha^2_{\parallel}(R) + 2\alpha^2_{\perp}(R)]$). Employing a Poission distribution one finds $\langle\alpha^2(R)\rangle_{\text{Poisson}}$ always less than $2\langle\alpha^1\rangle$ and thereby concludes that pairs cannot account for an upward deviation from Clausius-Mossotti behavior. [In the limit as $R \to 0$, $\alpha^2(R) \to \alpha_{\text{"He"}} \ll 2\alpha_{\text{"H"}}$.] A triad of donors in its ground state will involve higher orbitals than just the $1s$-A_1 orbitals for small R and this orbital promotion may produce the desired result $\langle\alpha^3(R_{12}, R_{23}, R_{31})\rangle > 3\langle\alpha^1\rangle$. (Note that in the limit R_{12}, R_{23}, and R_{31} all go to zero $\alpha^3 \to \alpha_{\text{"Li"}} \gg 3\alpha_{\text{"H"}}$, since $\alpha_{\text{H}} = 0.666$ Å^3 and $\alpha_{\text{Li}} \approx 22$ Å^3.) Presumably larger clusters, because of the polarization admixture of higher orbitals, may also have $\langle\alpha^m\rangle > m\langle\alpha^1\rangle$ and would thereby contribute to upward deviations from Clausius-Mossotti behavior. Takeshima[26] has expressed the view that interactions between donors will reduce the polarizability per donor to a smaller value than that for isolated (noninteracting) donors. The above discussion suggests this is probably only correct for donor pairs but not for triads and larger clusters in general. Moreover, since the consequence of donor-donor interactions is to decrease donor binding energies (and increase effective Bohr radii) as $N \rightarrow N_c$, it is difficult to see how the polarizability per donor can decrease with increasing N. This would imply downward deviations from Clausius-Mossotti behavior contrary to the experimental results.[9]

Whatever the merits of the "translational" corrections to E_{1oc} embodied in Eq. (7) or of the cluster approach represented by Eq. (8), there is the correction [reason (i)] which is present even for a regular lattic of donors. As the lattice parameter decreases (N increases), the donor-donor interactions begin to screen the large-r part of the attractive Coulombic potentials and thereby weaken the binding energy per donor electron. This in turn enhances the effective Bohr radius, and the donor polarizability $\alpha_D(N)$ increases with N. It is therefore of interest and importance to reliably determine the donor-polarizability enhancement $\alpha_D(N)/\alpha_D(0)$ as $N \rightarrow N_c$. Only after this contribution to the upward deviation from Clausius-Mossotti behavior has been reliably ascertained will it be possible to assess the relative importance of the randomness to the dielectric anomaly as $N \rightarrow N_c$. In the next section a continuum-model approach for calculating $\alpha_D(N)$ will be formulated.

III. THE HAMILTONIAN AND EFFECTIVE ONE-ELECTRON POTENTIAL

Because of difficulties[39] associated with using the formal perturbation expression for the wavenumber-dependent polarizability $\alpha(\vec{q},N)$, as employed by Leroux Hugon and Ghazali[25] in their calculation of the dielectric enhancement $\epsilon(N)/\epsilon_h$, we employ a variational approach originally used by Hassé.[30] This approach is known to give an exact result for an isolated hydrogenic atom. It avoids the problem of how the various excited states shift with donor density, but does yield the density dependence of the ground-state energy and an isotropic Bohr radius $a(N)$. The present work also utilizes a different interaction potential between neutral donors than that employed by Leroux Hugon and Ghazali.

The many-body problem of N interacting neutral donors per unit volume is replaced by a single-donor problem with an averaged effective-interaction term incorporating the interactions with all the donors. The single-donor effective Hamiltonian will be of the form $H_0 = H_{\mathrm{id}} + V_{\mathrm{int}}(r,N,a(N))$, where H_{id} is the isolated donor Hamiltonian and $V_{\mathrm{int}}(r,N,a(N))$ is an averaged spherically symmetric self-consistent interaction potential that is a function of the average donor density and also a function of the self-consistent Bohr radius $a(N)$ of the central and neighboring donors. The physical situation is pictured in Fig. 1. The attractive Coulomb interaction of an electron occupying a state with wave function $\psi_0(r,a(N))$ centered on a particular donor is screened by the host dielectric constant ϵ_h for $r < r_s$, where $4\pi r_s^3 N/3 = 1$. For $r \gg r_s$ the attractive Coulomb potential is screened not only by the host atoms but also by the intervening donor atoms, $V(r) = -e^2/\epsilon(N)r$ and the dielectric constant to be used is given by the Clausius-Mossotti expression [Eq. (5)], but with the additional caveat that the donor polarizability is a function of concentration. Since the host is being treated as a dielectric continuum and the interaction with other neutral donors is essentially averaged over the discrete donor positions to yield a spherical one-electron effective potential $V_{\mathrm{eff}}(r,N,\alpha_D(N))$, the spatially dependent dielectric function $\epsilon(r,N)$ varies from ϵ_h for $r < r_s$ to $\epsilon(N)$ [Eq. (5)] for $r \gg r_s$. If one considers a two-dielectric medium with $\epsilon = \epsilon_h$ for $r < r_s$ and $\epsilon(N)$ for $r > r_s$, elementary electrostatics suggests a $V_{\mathrm{int}}(r,N,\alpha_D(N))$ of the form $(e^2/r)[1/\epsilon_h - 1/\epsilon(N)]$. On general physical grounds one expects a smoothly varying $\epsilon(r,N)$ and a simple, convenient form, for the radial dependence involves an exponential e^{-r/r_s} which results in simple integrals in the matrix elements required for the calculation of $\alpha_D(N)$. With this type of r dependence $V_{\mathrm{eff}}(r,N,\alpha_D(N))$ takes the form

$$V_{\mathrm{eff}}(r,N,\alpha_D(N)) = -\frac{e^2}{r}\left(\lambda_D e^{-k_D r} + \frac{1}{\epsilon_h} e^{-r/r_s} + \frac{1}{\epsilon(N)}\left(1 - e^{-r/r_s}\right)\right) + V_\infty , \tag{9}$$

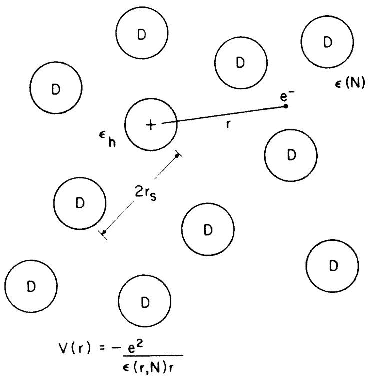

FIG. 1. Random group of neutral donors. The attractive Coulomb interaction of the central-donor electron is screened by the host-dielectric constant ϵ_h for $r < r_s$ [$r_s = (3/4\pi N)^{1/3}$] and by the long-wavelength dielectric constant $\epsilon(N)$ for $r \gg r_s$, $\epsilon(N)$ includes the screening resulting from the highly polarizable intervening neutral donors.

where $\lambda_D = (Z_D - Z_h - 1/\epsilon_h)$ and Z_D is the atomic number of the donor and Z_h that of the host. The first term in the large parentheses represents a central-cell correction with k_D a donor-dependent screening parameter (note that $k_D \gg 1/r_s$). The term $V_\infty\,[V_\infty \approx -(e^2/\epsilon_h a)(2\pi N a^3)]$ represents a constant potential as $r \to \infty$ representing the lowering of the total potential resulting from the continuum-donor background. Ignoring the V_∞ term, one observes, as $r \to 0$, $V_{\text{eff}}(r) \to -(e^2/r)(Z_D - Z_h)$ (a central-cell pseudopotential), for $1/k_D \ll r < r_s$ $V_{\text{eff}}(r) \sim -e^2/\epsilon_h r$, while for $r \gg r_s$ $V_{\text{eff}} \sim -e^2/\epsilon(N)r$. It should be noted that as $N \to N_c$ and $\epsilon(N)$ becomes very large the long-range Coulomb interaction (LRCI) is cut off for $r \gg r_s$. It is this weakening of the attractive LRCI that is responsible for the donor-polarizability enhancement.

When $V_{\text{eff}}(r, N, \alpha_D(N))$ is written in terms of an isolated-donor potential plus a donor-donor interaction term, $V_{\text{int}}(r, N, \alpha_D(N))$ takes the form

$$V_{\text{int}}(r, N, \alpha_D(N)) = \frac{e^2}{r}\left(\frac{1}{\epsilon_h} - \frac{1}{\epsilon(N)}\right) \times (1 - e^{-r/r_s}) + V_\infty\,. \tag{10}$$

This interaction form (excluding V_∞) is that employed by Haken[40] in treating the interaction of excitons with optical phonons and is also similar to the interaction term employed for the polaron problem in ionic crystals because of the $1/\epsilon_h - 1/\epsilon(N)$ factor. This factor is proportional to $N\alpha_D(N)$ for $N \ll N_c$ thereby showing that the donor-donor interactions are proportional to the polarization per unit volume. Whereas in ionic crystals the polarization effects from the optical phonon modes result from the factor $1/\epsilon_\infty - 1/\epsilon_0$, here the $1/\epsilon_h - 1/\epsilon(N)$ factor accounts for the polarization contribution of the donor electrons to the Coulombic interactions between neutral donors. Ordinarily when considering a many-body problem the Hamiltonian will include separate interaction terms of an electron with the other electrons and with the other ion cores. The present formulation combines these two types of terms and considers the interaction as that between neutral polarizable donor atoms, which might be viewed as the natural approach on the insulating side of the MNM transition.

An alternative approach for calculating $V_{\text{int}}(r, N, a(N))$ has been employed by Leroux Hugon and Ghazali.[25] They use the Coulomb interaction for a donor molecule (D_2), for which the molecular wave function is of the Heitler-London form valid for large internuclear separations. They then integrate this D_2 Coulomb interaction over a continuum-donor distribution beyond a concentration-dependent cutoff radius. Their approach is briefly outlined in Appendix A. Several curves of their $V_{\text{int}} - V_\infty$ for different concentrations are shown in Fig. 2 along with the corresponding $V_{\text{int}} - V_\infty$ from Eq. (10) as a function of the dimensionless radius $\rho = r/a$. Although these two types of interactions are qualitatively similar, two distinct differences between them should be emphasized. At large values of r ($\rho \gg 1$) the electrostatic model $V_{\text{int}} - V_\infty$ falls off as $1/r$ while their $V_{\text{int}} - V_\infty$ falls off exponentially as $e^{-2r/a}$. It is our viewpoint that for an insulator ($T = 0$ K) a screened point charge should produce a potential falling off as $1/\epsilon(N)r$ at large r, where $\epsilon(N)$ is the total static dielectric constant of the solid, including the impurity contribution. The $V_{\text{int}} - V_\infty$ term of Leroux Hugon and Ghazali includes the contribution of the sum over all pairs of donors, but neglects all higher-order (third-order, etc.) contributions to the one-electron $V_{\text{int}}(r)$. The electrostatics model $V_{\text{int}} - V_\infty$ implicitly includes these higher-order contributions. The second difference worth noting is the concentration dependence at small r ($\rho \ll 1$). The electrostatics model potential $[V_{\text{int}}(\rho \ll 1) - V_\infty] \propto N^{4/3}$ for $N \ll N_c$ while Leroux Hugon and Ghazali find $[V_{\text{int}}(\rho \ll 1) - V_\infty] \propto N$. This different N dependence at small ρ is also thought to arise from the effect of higher-order contributions to the one-electron $V_{\text{int}}(r)$. Their treatment, based on pairs, leads to a linear dependence on N while the electrostatics model $V_{\text{int}} - V_\infty$ leads to a nonlinear N dependence. Figure 2 shows that the model potential is smaller at $z = 0.1$, but is already substantially larger at $z = 0.5$. The curves based on the D_2 molecular treatment also show the surpris-

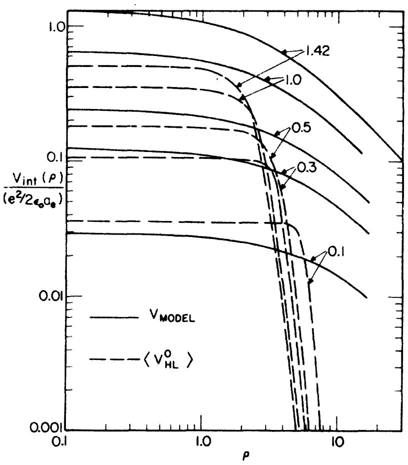

FIG. 2. Comparison of the model interaction potential [Eq. (10)] based on electrostatics with the Heitler-London-type Coulomb interaction utilized by Leroux Hugon and Ghazali (Ref. 25). $\rho = r/a$ and curves are shown for various values of the quantity $z = [4\pi\alpha_D(0)/3\epsilon_h]\,N$. Note that at large ρ the model potential falls off as $1/\rho$, while the Heitler-London $V_{\text{int}}(\rho)$ falls off as $e^{-2\rho}$.

ing feature of crossing in the range $\rho \sim 2.3$ to 4.2 leading to the result that at large ρ $[V_{\rm int}(\rho \gg 1) - V_\infty]$ actually decreases with increasing concentration. The model $V_{\rm int}(r,N,\alpha_D(N))$ does not show this crossing feature and would appear to be more physically reasonable. On the other hand, the large ρ region, where the ρ dependence of the two different $V_{\rm int} - V_\infty$ forms are markedly different, may not give large differences in $\alpha_D(N)$ because the wave function $\psi(r,N)$ will be small in this region.

Although the indirect-gap semiconductors have multiple conduction-band valleys, each characterized by a mass tensor with components m_l and m_t, we shall employ a single isotropic mass m^* chosen to produce the correct effective mass approximation (EMA) binding energy for a single valley, namely, $m^*/m = |E_{1s\text{-EMA}}|\epsilon_h^2/\text{Ry}$ ($m^*/m = 0.299$ for Si and 0.17 for Ge). Dexter[41] has calculated by a variational technique the EMA polarizability values $[\alpha_{\rm EMA} = \frac{1}{3}(\alpha_\parallel + 2\alpha_\perp)]$ for isolated donors. His results show $\alpha_{\rm EMA} = 4.328 \times 10^{-19}$ cm^3 for Si and $\alpha_{\rm EMA} = 9.329 \times 10^{-18}$ cm^3 for Ge. Dexter's results are, respectively, 2.5% and 24% larger than the isotropic $\alpha_{\rm EMA}$ values $[\alpha_{\text{EMA-iso}} = \frac{9}{2}\epsilon_h a_e^3$, where $a_e = a_B \epsilon_h (m/m^*)]$. The small difference, at least for Si, suggests it is a good approximation to neglect the mass anisotropy. This use of an isotropic mass m^* in the kinetic energy term of H_0 allows the use of a spherical ground-state wave function $\psi_0(r,a(N))$ for each conduction-band valley—an important simplification in the calculation of $\alpha_D(N)/\alpha_D(0)$ to follow.

IV. VARIATIONAL CALCULATION OF THE DONOR-POLARIZABILITY ENHANCEMENT

In the presence of an electric field each donor will become polarized and will have an induced dipole moment $\vec{p}(\vec{r}_i) = \overleftrightarrow{\alpha}_D(\vec{r}_i, N) \cdot \vec{E}_{\rm loc}(\vec{r}_i, N)$. Considering a regular donor array there will be no contribution from the fluctuation-correlation term in Eq. (4). For this regular array the correct orthogonalized localized wave functions are, of course, Wannier functions. Since the polarization operator is the sum of one-electron operators, the polarization $\vec{P}(N)$ will be given by $N\alpha_D(N)\vec{E}_{\rm loc}(N)$ (dropping the $\langle\ \rangle$ symbols), where $\alpha_D(N)$ is the polarizability of a donor electron occupying a Wannier function localized about a specific donor site. The Wannier functions will have the symmetry of the lattice, will of necessity have nodes, and will be very difficult to determine accurately, particularly as $N \to N_c$ and the overlap of atomic wave functions becomes large. As an approximation we will employ spherically symmetric valley-envelope functions of the form $\psi_0(r,a(N)) = [\pi a(N)^3]^{-1/2} e^{-r/a(N)}$, where $a(N)$ is a concentration-dependent Bohr radius. The use of a nonorthogonal $\psi_0(r, a(N))$ will cause a monotonically increasing error as $N \to N_c$, but it should not be assumed that the results will be very different than those for a calculation of $\alpha_D(N)$ employing a reliable Wannier function. This question will receive additional consideration in the Discussion.

The Hamiltonian for the donor electron in a uniform electric field along the z axis will be $H = H_0 - eFz$, where F is the magnitude of the electric field and H_0 is the total donor Hamiltonian including $V_{\rm int}(r,N,\alpha_D(N))$ in the absence of the applied electric field. The Hassé[30] trial variational wave function is

$$\psi_z = \psi_0(r,a(N))(1 + bz + crz)\,. \tag{11}$$

Minimization of the energy $E = \langle\psi_z|H|\psi_z\rangle/\langle\psi_z|\psi_z\rangle$ with respect to the variational parameters b and c yields after some manipulation

$$E = E_0(N) - e^2F^2\left(\frac{\lambda_1^2(B - \lambda_2 E_0) + \lambda_3^2(A - \lambda_1 E_0) - 2\lambda_1\lambda_3(C - \lambda_3 E_0)}{(A - \lambda_1 E_0)(B - \lambda_2 E_0) - (C - \lambda_3 E_0)^2}\right), \tag{12}$$

where $E_0(N) = \langle\psi_0|H_0|\psi_0\rangle$, $\lambda_1 = \langle z\psi_0|z\psi_0\rangle$, $\lambda_2 = \langle rz\psi_0|rz\psi_0\rangle$, $\lambda_3 = \langle rz\psi_0|z\psi_0\rangle$, $A = \langle z\psi_0|H_0|z\psi_0\rangle$, $B = \langle rz\psi_0|H_0|rz\psi_0\rangle$, and $C = \langle rz\psi_0|H_0|z\psi_0\rangle$. Clearly the coefficient of F^2 is just $\alpha_D(N)/2$. Although Eq. (12) is rigorous for the trial wave function ψ_z in Eq. (11), this expression only gives an exact result for an isolated hydrogenic atom. Since the correct donor $\psi_0(r)$ is well approximated by a hydrogenic wave function, the above ψ_z should be expected to yield reliable results for shallow donors. For the particular hydrogenic $\psi_0(r,a(N))$ given above, $\alpha_D(N)$ is given by

$$\alpha_D(N) = 2e^2a^4\left(\frac{25a^2(A - a^2E_0) + 4(B - \frac{15}{2}a^4E_0) - 10a(C - 5a^3E_0)}{4(A - a^2E_0)(B - \frac{15}{2}a^4E_0) - (C - 5a^3E_0)^2}\right), \tag{13}$$

while the single-valley energy $E_0(N)$ is given by

$$\frac{E_0(N)}{e^2/2\epsilon_h a_e} = \frac{a_e}{a(N)}\left[2 - \frac{a_e}{a(N)} - 2\left(1 - \frac{\epsilon_h}{\epsilon(N)}\right)(1 - y^2) + 2x(N)\right], \tag{14}$$

where a_e is the EMA isotropic Bohr radius, $y(N) = 1/[1+a(N)/2r_s]$, and $x(N) = \lambda_D \epsilon_h/[1+k_D a(N)/2]^2$. The $E_0(N)$ given in Eq. (14) has been obtained by minimizing $E_0(N)$ with respect to $a(N)$, which yields a concentration-dependent Bohr radius $a(N)$ given by

$$a(N) = a_e/\{1 + x(N)\beta(N) - [1-\epsilon_h/\epsilon(N)](1-y)^2(1+2y)\}\,, \qquad (15)$$

where $\beta(N) = 3 - 2/[1+k_D a(N)/2]$. In the dilute limit $\epsilon(N) \to \epsilon_h$, $y(N) \to 1$ and the Bohr radius is given by $a(0) = a_e/[1+x(0)\beta(0)]$. As the central-cell correction increases, $x(0)$ increases and the Bohr radius $a(0)$ decreases [$\beta(N)$ is slightly less than three and does not vary much with N and varies only slightly with the donor]. In minimizing $E_0(N)$ with respect to $a(N)$ in obtaining Eq. (15) the spatially constant V_∞ was excluded even though V_∞ is a function of $a(N)$. The reason is that a constant potential (spatially constant) can make no contribution to the extent [of order $a(N)$] of a localized wave function even though it does contribute to the energy. [If one includes V_∞ in the minimization of $E_0(N)$, one obtains unphysical behavior for $a(N)$.] The matrix elements A, B, and C as a function of x, y, $1-\epsilon_h/\epsilon(N)$, β, and $a(N)$ are given in Appendix B.

Substituting A, B, and C from Appendix B and $E_0(N)$ into $\alpha_D(N)$ in Eq. (13), one obtains after considerable algebraic manipulation the donor-polarizability enhancement $W(N) = \alpha_D(N)/\alpha_D(0)$ given by

$$W(N) \equiv \frac{\alpha_D(N)}{\alpha_D(0)} = \left(\frac{a(N)}{a(0)}\right)^3 \left(\frac{I(N)}{I(0)}\right), \qquad (16a)$$

where

$$I(N) = \frac{9 + (10+4\beta)x - f(N,\alpha_D,y)(9+18y+5y^2+20y^4)}{2 + 7x + (5+2\beta-\beta^2)x^2 - f(N,\alpha_D,y)g(y,x,\beta) + f^2(N,\alpha_D,y)(1+y)^4(2+y^2)}\,, \qquad (16b)$$

with

$$g(y,x,\beta) = 4 + 8y + 5y^2 + 2y^3 + 5y^4 + 4x\beta y^2(1+y) + x(7+14y+5y^2+10y^4)\,, \qquad (16c)$$

and $f(N,\alpha_D,y) = [1-\epsilon_h/\epsilon(N)](1-y)^2$. The polarizability in the dilute limit is given by $\alpha_D(0) = \epsilon_h a(0)^3 I(0)$. With no central-cell correction ($\lambda_D = 0$), $\alpha_D(0) \equiv \alpha_{\text{EMA iso}} = \frac{9}{2}\epsilon_h a_e^3$. With the central-cell correction $\alpha_D(0)$ is given by

$$\alpha_D(0) = \epsilon_h a(0)^3 \left(\frac{9 + [10+4\beta(0)]x(0)}{2 + 7x(0) + [5+2\beta(0)-\beta(0)^2]x(0)^2}\right). \qquad (17)$$

For a given donor λ_D is determined, and the only remaining undetermined parameter in the Hamiltonian is the central-cell correction screening parameter k_D. Once k_D is determined $x(0)$, $\beta(0)$, $a(0)$, and thus $\alpha_D(0)$ are all determined. Our procedure has been to determine k_D and these other quantities by force-fitting $\alpha_D(0)$ to the experimentally determined values[42]; however, there are certain problems associated with the accurate experimental determination of $\alpha_D(0)$ as will be discussed later.

The polarizability enhancement $W(N)$ given in Eq. (16) is a single-valley result. It is possible to obtain an approximate many-valley result employing the wave function

$$\psi_{1s\text{-}A_1} = (1/N_v)\sum_{i=1}^{N_v} e^{i\vec{k}_i\cdot\vec{r}}\psi_0(\vec{r},N)\,,$$

where N_v is the number of valleys and $\vec{k}_i$ is the k vector to the ith-valley minimum. Because of the extra r^2, r^3, and r^4 dependences of the matrix elements λ_1, λ_2, λ_3, A, B, and C one can demonstrate that, although the central-cell potential makes an important correction through the valley-valley coupling terms to the ground-state energy $E_{1s\text{-}A_1}(N)$, the valley-valley corrections to λ_1, λ_2, λ_3, A, B, and C are negligible. The largest correction to $\alpha_D(N)$ then comes from replacing $E_0(N)$ in Eq. (13) by $E_{1s\text{-}A_1}(N)$, which is given approximately by

$$E_{1s\text{-}A_1}(N) = E_0(N) + \eta E_e[a(0)/a(N)]^3\,, \qquad (18)$$

where ηE_e ($E_e = e^2/2\epsilon_h a_e$) is the valley-valley coupling contribution to the ground-state energy in the dilute limit. The correction to $W(N)$ comes through the $I(N)$ term, which is modified by adding the following: (1) the quantity $5p$ to the numerator, (2) the quantity $(p/2)(7+10x+2x\beta) + 5p^2/4$ to the denominator, and (3) replacing the last x in $g(y,x,\beta)$ in the denominator by $x + p/2$. The quantity $p[\eta, a(N)] = \eta[a(N)/a_e][a(0)/a(N)]^3$. One should emphasize, however, that $a(N)$ in Eq. (15) has not been determined by a full many-valley treatment.

The $W(N)$ given in Eq. (16a) with the quantity $I(N)/I(0)$ can readily be solved iteratively to high accuracy arbitrarily close to the divergence of $\epsilon(N)$ as $N \to N_c$. The procedure we employed was to obtain a polynomial in $W(N)$ by eliminating $[a(N)/a(0)]^3$ from Eq. (16a) using Eq. (15); however, the complex form of $I(N)/I(0)$ precludes eliminating it from Eq. (16a) in any simple manner. The polynomial in $W(N)$ is

$$2z\left(1-\frac{3(1-y)^2(1+2y)}{2[1+x(N)\beta(N)]}\right)\frac{W^{4/3}}{[I(N)/I(0)]}+\frac{W^{1/3}}{[I(N)/I(0)]}-\left(\frac{1+x(0)\beta(0)}{1+x(N)\beta(N)}\right)(2zW+1)=0\;, \quad (19)$$

where $z=4\pi N\alpha_D(0)/3\epsilon_h$. This equation can be solved to the desired accuracy by iteration for the quantities $W(N)$, $I(N)/I(0)$, $1-\epsilon_h/\epsilon(N)$, $y(N)$, $x(N)$, and $\beta(N)$ for each value of N. The numerical calculations were conveniently handled by a TI59 calculator, utilizing one of the module programs for finding the zeros of a polynomial as a subroutine. The iteration was continued until successive values of $W(N)$ differed by less than one part in 10^5 typically and in special situations to much smaller differences. The convergence of the iterations becomes much slower as $N \to N_c$, presumably because of the highly nonlinear behavior of $W(N)$ as $N \to N_c$. For the worst case [as with the largest $W(N_c)$ value] the iteration could take more than an hour for N close to N_c. The computational accuracy of the numerical results for the various quantities is thought to be much better than 1%.

Before proceeding to discuss the results, two final points should be emphasized. The main idea of the present calculation is to ascertain how the static screening $[\epsilon(N)]$ changes with concentration as $N \to N_c$, approaching from the insulating side. Although various approximations have been made, namely, (1) a spherical $\psi_0(r,a(N))$ of hydrogenic form, (2) the neglect of the orthogonality requirement for discrete neighbors through the use of a continuum model, and (3) the neglect of randomness, spin, and exchange—for the particular $V_{\text{int}}(r,N,_D(N))$ employed, no screening terms have been neglected. The quantity $y(N)$ given by $1/[1+a(N)/2r_s]$ is the screening parameter (in fact $1-y$ is directly proportional to the screening) and varies from 1 to 0.76 as N increases from zero to N_c. No terms in y have been neglected and it is noted that $I(N)$ contains various polynomials in y, including y^n terms up to $n=10$. The point to be stressed is, that for the given model $V_{\text{int}}(r,N,\alpha_D(N))$, the screening has been treated with no approximations. The second point to be noted is that the calculation is fully self-consistent (SC), by which we mean that all quantities that should vary with N, even if only by small amounts [for example $x(N)$ and $\beta(N)$], have been allowed to change with N. One could more easily do a simpler screening calculation with fixed values of x and β [$x(0)$ and $\beta(0)$] and obtain somewhat different numerical results, but this calculation would not be fully SC.

V. RESULTS AND DISCUSSION

The donor-polarizability enhancements $[W(N)=\alpha_D(N)/\alpha_D(0)]$ and the factors $I(N)/I(0)$ are shown in Fig. 3 for Si:As, Si:P, Si:Sb, and the Si EMA case. The results show that $W(N)$ starts increasing, at first slowly, above $N \sim 10^{18}/\text{cm}^3$ with the increase becoming monotonically steeper as $N \to N_c$. For the more tightly bound donors $W(N)$ increases more slowly for a given N, but ultimately becomes larger as $N \to N_c$. The values of $W(N_c)$ and the respective N_c values are given in Table I along with some of the parameters employed in the calculation such as $a(0)$, $k_D a(0)$, $x(0)$, and $\beta(0)$. Based on the values of $\alpha_D(0)$ from Ref. 42, $W(N_c)$ increases from the EMA value of 1.883 to 2.84 for Si:As. New smaller values of $\alpha_D(0)$ have recently been obtained[43] for Si:P and Si:Sb. These smaller values of $\alpha_D(0)$ lead to somewhat larger enhancements at N_c and to larger values of $N_{c\text{-calc}}$. These new values are also shown in Table I (in parentheses). The comparison of the calculated and experimental N_c values will be discussed below.

Figure 3 also shows the quantity $I(N)/I(0)$ rising proportionately to $W(N)$ and growing at approximately $\frac{1}{3}$ the rate of $W(N)$. Recalling Eq. (16a), the total polarizability enhancement $\alpha_D(N)/\alpha_D(0)$ can be broken into two contributions, namely, (1) the wave-function volume effect proportional to $[a(N)/a(0)]^3$ and (2) the $I(N)/I(0)$ factor. The physical origin of this second contribution arises be-

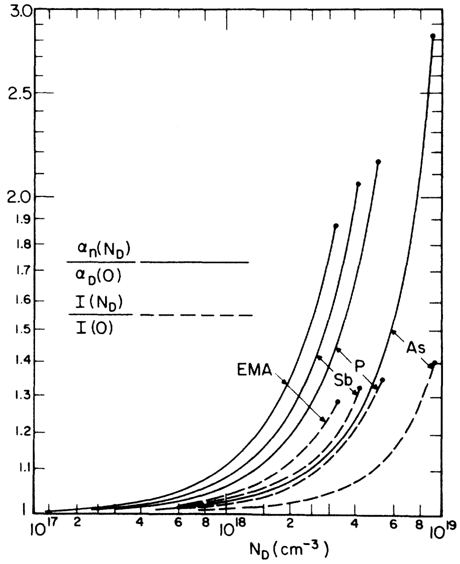

FIG. 3. Polarizability enhancement $W(N)=\alpha_D(N)/\alpha_D(0)$ and the quantity $I(N)/I(0)$ vs N_D for Si:As, Si:P, Si:Sb, and the EMA case for Si. The dots terminating the curves are the values of the quantitites at $N_{c\text{-calc}}$ where $\epsilon(N_c)=\infty$ [based on the $\alpha_D(0)$ values from Ref. 42].

TABLE I. Polarizability-enhancement parameters for silicon shallow donors.

	EMA	Sb	P	As
$\alpha_D(N_c)/\alpha_D(0)$	1.883	2.06 $(2.26)^c$	2.17 $(2.34)^c$	2.84 $(3.13)^c$
$I(N_c)/I(0)$	1.274	1.31 (1.34)	1.33 (1.36)	1.43 (1.46)
$a(N_c)/a(0)$	1.139	1.163 (1.190)	1.175 (1.199)	1.258 1.291
$E_{1s}(N_c)E_{1s}(0)$	0.30	0.33 (0.32)	0.31 (0.31)	0.31 (0.30)
$10^{-18}N_{c\text{-calc}}$ (cm^{-3})	3.34	4.25 (6.02)	5.23 (7.27)	9.57 (10.90)
$10^{-18}N_{c\text{-expt}}$ (cm^{-3})		3.0 ± 0.2^d	3.5 ± 0.4^d	6.4 ± 0.5^d
$\alpha_D(0)$ (Å^3)	4.328×10^5	$3.1\times10^{5\,a}$ $(2.0\times10^5)^b$	$2.4\times10^{5\,a}$ $(1.6\times10^5)^b$	$1.0\pm10^{5\,a}$ $(0.78\pm10^5)^b$
$a(0)$ (Å)	20.18	19.90 (17.25)	18.49 (16.23)	14.57 (13.51)
$\frac{1}{2}k_D a(0)$		298.888 (84.898)	17.189 (9.977)	39.694 (34.858)
$x(0)$	0	0.0047 (0.0570)	0.0314 (0.0863)	0.1302 (0.1676)
$\beta(0)$		2.993 (2.977)	2.890 (2.818)	2.951 (2.944)
$\lambda_D\epsilon_h$	0	420.8	10.4	215.6
η	0	0.2856	0.3239	0.5740

[a] After Bethin *et al.*, Ref. 42.

[b] See Ref. 43 for Si:Sb and Si:P and Ref. 9 for Si:As.

[c] The lower values based on the smaller values of $\alpha_{D\text{-dilute}}\sim\alpha_D(0)$ are thought to be the more reliable values.

[d] After Castner *et al.*, Ref. 9.

cause the outer portion of the donor electron's wave function is screened not only by the host but also by the intervening neutral donors. Employing the radial dependence of $\epsilon(r,N)$ implied by Eq. (9) numerical analysis shows that $I(N)/I(0)$ is well approximated (except very close to N_c) by the relation

$$\frac{I(N)}{I(0)}\sim\frac{\langle\epsilon(r,N)\rangle}{\epsilon_h}=\frac{1}{\epsilon_h}\int\epsilon(r,N)\,|\psi_D(r,N)|^2dV\,. \tag{20}$$

As N increases, the donors interact more strongly and the average dielectric constant $\langle\epsilon(r,N)\rangle$ seen by the donor electron increases. As $N\rightarrow N_c$ the long-wavelength dielectric constant $\epsilon(N)$ diverges, but $\langle\epsilon(r,N_c)\rangle$ remains finite because of the strong spatial dispersion of $\epsilon(r,N_c)$ and the limited extent of $\psi_D(r,N_c)$.

An alternative method of plotting $\alpha_D(N)/\alpha_D(0)$ is given in Fig. 4 showing $\ln[\alpha_D(N)/\alpha_D(0)]$ vs N. The figure shows that $\ln[\alpha_D(N)/\alpha_D(0)]$ varies as N^{m_D}, except very close to N_c (the EMA case appears to show no deviation close to N_c). The polarizability enhancement is well fit by the empirical relationship $\alpha_D(N)/\alpha_D(0)=\exp(\beta_D N^{m_D})$. The deviation from this behavior is largest for As, smallest for Sb, and apparently zero for the EMA case, thereby suggesting the deviation is directly related to the magnitude of the central-cell correction to the energy $E_{1s\text{-}A_1}$ of the donors. In Table II are given the values of β_D and m_D for the three donors Sb, P, As, and for the EMA case for Si. The results show $m_D\sim1.5$ and values of β_D which rapidly decrease with the donor binding energy. This exponential functional dependence could not have been obtained from a virial expansion containing only a few terms with increasing powers of the density N, as has typically been employed to discuss deviations for Clausius-Mossotti behavior for highly dense gases or liquids. The most widely discussed mechanism for the deviation from Clausius-Mossotti behavior for the dense gases is the translational fluctuation mechanism. The mechanism

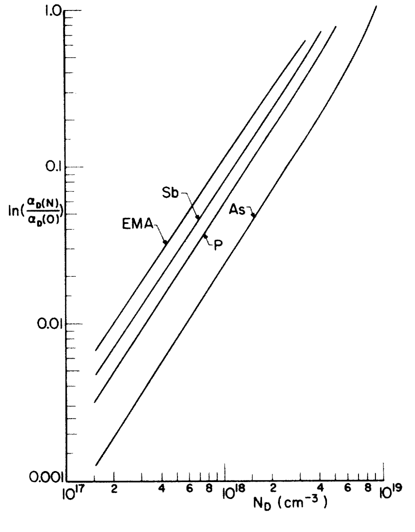

FIG. 4. $\ln[\alpha_D(N)/\alpha_D(0)]$ vs N_D for Si:As, Si:P, Si:Sb, and the Si EMA case. $\alpha_D(N_D)/\alpha_D(0)$ is well fitted, except close to N_c by $\alpha_D(N_D)/\alpha_D(0)=\exp(\beta_D N_D^{m_D})$ where $m_D \sim 1.5$. The upward correction from this result near N_c is negligible for the EMA case and increases with the magnitude of the central-cell correction to the donor binding energy $E_{1s\text{-}A_1}(N)$. These curves are based on the $\alpha_D(0)$ values from Ref. 42.

employed here is the nonlinear screening resulting from donor-donor interactions represented by $V_{\rm int}(r,N,\alpha_D(N))$.

In Fig. 5 the energies for the donor ground state $E_{1s\text{-}A_1}(N)$ and the $2p$ excited state [calculated with the same variational approach and the same $V_{\rm int}(r,N,{}_D(N))$] are shown versus donor concentration ($E_{1s\text{-}A_1}(N)=E_0(N)+\eta E_e[a(0)/a(N)]^3$, where $E_0(N)$ [see Eq. (14)] is the single-valley energy and $\eta E_e[a(0)/a(N)]^3$ is the valley-valley coupling contribution which has been approximated by finding the change in $|\psi(0,N)|^2$, since the valley-valley coupling terms are dominated by the central-cell small-r part of the wave function). The single-valley energy $E_0(N)$ changes more rapidly with N than the valley-valley contributions, which only change significantly for $N > 0.4N_c$. The $E_{1s\text{-}A_1}(N)$ curves drop by approximately a factor of 3 as $N \rightarrow N_c$ (the values at N_c are given in Table I). The $2p$ levels (no central-cell correction required) drop smoothly to zero, reaching zero for values of N slightly less than N_c. This behavior appears reasonable since the more extended $2p$ levels are screened more effectively by $V_{\rm int}(r,N,\alpha_D(N))$ than the more localized ground state which is "held in" partially by the stronger central-cell potential. This behavior of the $2p$ level with N is different from that reported by Leroux Hugon and Ghazali,[25] who found virtually no change of E_{2p} with N as $N \rightarrow N_c$. As a result they found a larger reduction in $|E_{1s}-E_{2p}|$ than found here for the EMA case. The solid line in Fig. 5 shows $E_{2p}-E_{1s\text{-}A_1}$ for the three donors and the EMA case. For the EMA case Fig. 5 shows $(E_{2p}-E_{1s})/E_e$ reduced from 0.75 at $N=0$ to about 0.3 at N_c ($N_c=3.34\times10^{18}/\text{cm}^3$ for the EMA case), however, at $N\approx1.25\times10^{18}/\text{cm}^3$ the reduction in $E_{2p}-E_{1s}$ is only about 10%. At this same concentration (which is the N_c in Ref. 25) Leroux Hugon and Ghazali found $E_{2p}-E_{1s}$ reduced by a factor of 3, which in turn leads to their overestimate of the polarizability enhancement at this concentration. Figure 3 shows $\alpha_{\rm EMA}(N)/\alpha_{\rm EMA}(0)\approx1.2$ at this concentration. The infrared absorption data of Townsend[4] on Si:P also suggest no significant change with N of the $E_{1s\text{-}A_1}\rightarrow E_{2p_0}$ transition until $N > 0.3N_c$. Figure 5 shows only a small reduction in $\Delta E=E_{2p}-E_{1s\text{-}A_1}$ at $N\sim0.3N_c$. On the other hand, the interesting new results of Capizzi *et al.*[44] demonstrate the formation of donor exciton bands at lower energies than the isolated donor transitions. This suggests additional oscillator strength below the transition energies shown in Fig. 5 resulting from charge-transfer excitonlike absorption below the upper Hubbard band. The present calculation does not include any charge-transfer excitations; however, the effect of this extra low-energy absorption corresponds to an extra enhancement of the static dielectric constant when viewed within the framework of the Kramers-Kronig relationship

$$\epsilon(\omega=0,N)-\epsilon_h=8\int_0^\infty \frac{\sigma(\omega',N)d\omega'}{\omega'^2}\,, \qquad (21)$$

where $\sigma(\omega',N)$ is the frequency-dependent conductivity resulting from the neutral donors ($4\pi\sigma(\omega)=\omega\,\text{Im}[\epsilon(\omega)]$). Because of the strong weighting of low frequencies in Eq. (21), extra absorption from the lower-energy charge-transfer excitations significantly enhance the $\epsilon(\omega=0,N)$ val-

TABLE II. Parameters for the empirical expression $\alpha_D(N)\alpha_D(0)=\exp(\beta_D N_D^{m_D})$ for the polarizability enhancement.[a]

	EMA	Sb	P	As
m_D	1.476	1.506	1.518	1.557
β_D (cm^{3m_D})	2.86×10^{-28}	6.11×10^{-29}	2.75×10^{-29}	2.17×10^{-30}

[a] These values are based on the $\alpha_D(0)$ values from Ref. 42.

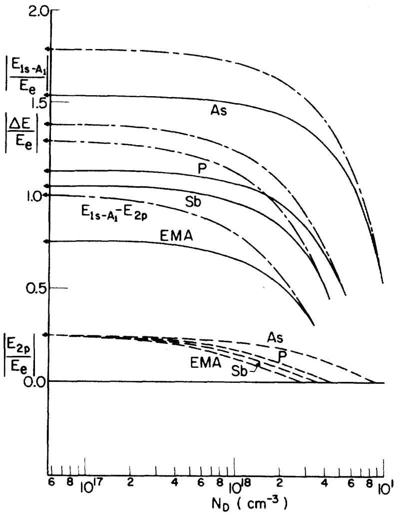

FIG. 5. The magnitudes of the energies (relative to $E_e = e^2/2\epsilon_h a_e$) $E_{1s\text{-}A_1}(N_D)$ is denoted by the long-short dash curve; $E_{2p}(N_D)$, by the dashed curve, and the energy difference $\Delta E(N_D) = E_{2p}(N_D) - E_{1s\text{-}A_1}(N_D)$ by the solid line curve vs N_D for Si:As, Si:P, Si:Sb, and the Si EMA case. These curves are based on parameters obtained by using the $\alpha_{D\text{-dilute}}$ value from Ref. 42 to fit to the theoretical expression for $\alpha_D(0)$ [Eq. (17)].

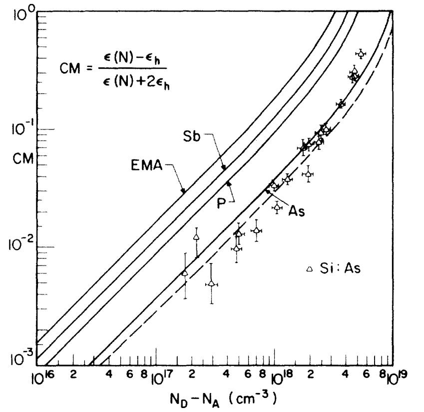

FIG. 6. The calculated Clausius-Mossotti factor CM $= [\epsilon(N) - \epsilon_h]/[\epsilon(N) + 2\epsilon_h]$ vs N_D ($N_D - N_A$ for the experimental data) for Si:As, Si:P, Si:Sb, and the Si EMA case. Also shown is the Si:As data from Ref. 9. The calculated curves are based on parameters obtained by using the $\alpha_{D\text{-dilute}}$ values from Ref. 42 fit to the theoretical expression for $\alpha_D(0)$ [Eq. (17)]. Smaller experimental values of $\alpha_{D\text{-dilute}}$ will lower these curves somewhat and result in slightly higher values of $N_{c\text{-calc}}$. The dashed curved is for $\alpha_{\text{As}} = 0.78 \times 10^5$ Å^3.

ues over those based on single-site excitations alone. An additional source of low-energy absorption shown in the data of Townsend[4] is the growth of the forbidden transition $1s\text{-}A_1 \rightarrow 1s\text{-}T_2$ of "isolated" donors with increasing concentration. The clustering of the donors and the growth of random internal electric fields makes this low-energy absorption band (100 cm^{-1} in Si:P) increasingly allowed as $N \rightarrow N_c$. This transition is also not accounted for in the continuum-model calculation of $\alpha_D(N)/\alpha_D(0)$ given above.

Figure 6 shows the calculated Clausius-Mossotti factor $[\epsilon(N) - \epsilon_h]/[\epsilon(N) + 2\epsilon_h]$ for the Si EMA case, Si:Sb, Si:P, and Si:As. Also shown in Fig. 6 are the Si:As data from Ref. 9. The calculated curves are based on the dilute polarizabilities given in Ref. 42. [Recent, more accurate measurements[43] and the recognition of the polarizability enhancement in the 0.5 to 2×10^{18} cm^3 range have led to smaller values of $\alpha_{D\text{-dilute}} \sim \alpha_D(0)$.] The values of $N_{c\text{-calc}}$ are given in Table I, including new values based on the smaller values of $\alpha_D(0)$. For the smaller, more reliable values of $\alpha_D(0)$ the values of $N_{c\text{-calc}}$ range from 60% to 100% larger than the values of $N_{c\text{-expt}}$ for the three donors. In Fig. 6 the Si:As data clearly rise more rapidly than the calculated curve above 3×10^{18}/cm^3, demonstrating that the $\alpha_{\text{eff}}(N)$ from the data rises more rapidly than the $\alpha_D(N)$ calculated with the above continuum model. The polarizability enhancement as calculated is insufficient to explain the experimental results; however, it does lead to a substantial improvement over the $N_{c\text{-CM}}$ values calculated using the uncorrected Herzfeld criterion for N_c [$N_{c\text{-CM}} = 3\epsilon_h/4\pi\alpha_D(0)$] based on the dilute values of α_D. The calculation demonstrates that the polarizability enhancement is a substantial effect ranging from about two for Si:Sb to about three for Si:As. Although one might, because of some of the approximations made, expect some difference in $\alpha_D(N)/\alpha_D(0)$ for a regular array of donors characterized by Wannier functions associated with the lower Hubbard band, it is not at all clear the results would be changed significantly. Wannier functions will have nodes (before reaching the nearest neighbors), but will also extend further than the hydrogenic wave function $\psi_0(r, a(N))$ and, in addition, reflect the symmetry of the regular array of donors. Recalling that the matrix elements λ_1, λ_2, λ_3, A, B, and C which enter the calculation of $\alpha_D(N)$ through Eqs. (12) and (13) emphasize the more distant portions of the wave function because of the r^2, r^3, and r^4 factors in these matrix elements, it is plausible that the polarizability enhancement for an accurate Wannier function might be larger than calculated with the continuum model and $\psi_0(r, a(N))$. However, obtaining a reliable trial wave function ψ_ϵ in an applied electric field E_ϵ for a reliable Wannier function may be difficult. Moreover, because of the inadequacy of the dipole-dipole approximation leading to $\vec{E}_{\text{loc}} = \vec{E}_{\text{ext}} + 4\pi\vec{P}/3\epsilon_h$ (as $N \rightarrow N_c$ and the donor wave functions overlap more and more) the calculation should take account of the spatial dependence of the local field $E_{\text{loc}}(\vec{r})$. The wave-number-dependent local field $\vec{E}_{\text{loc}}(\vec{q})$ has been calculated by Nagel and Witten[45] for simple cubic (sc), fcc, and bcc lattices for the [100], [110], and [111] crystal axes. A suitable average of the three directions for the latter two cases might yield an estimate of $\langle \vec{E}_{\text{loc}}(\vec{r}) \rangle$ for a random distribution of donors. Their results suggest the coefficient of $\vec{P}$ ($4\pi/3\epsilon_h$ for the point-dipole case) may increase somewhat as overlap increases as $N \to N_c$. No detailed analysis of this correction can be given at this time, but their results suggest the Lorentz-Lorenz correction, although modified, will remain important as $N \to N_c$ and may even be enhanced. This is a crucial point, since it is the Lorentz-Lorenz local-field correction that produces the polarizability catastrophe. Thus additional corrections beyond the polarizability enhancement $\alpha_D(N)/\alpha_D(0)$ calculated herein resulting from donor randomness, the $\vec{r}$ dependence of $\vec{E}_{\text{loc}}(\vec{r})$, and charge transfer contributions to $\alpha_{\text{eff}}(N)$ may actually further improve the agreement between the Clausius-Mossotti approach and the experimental data. None of these corrections seem to alter the basic prediction of a smooth second-order polarization catastrophe as $N \to N_c$. Note that if one accepts the quantitative validity of the continuum-model $\alpha_D(N)/\alpha_D(0)$ calculation, Eq. (4) suggests the fluctuation contribution to the total polarization $\vec{P}(N)$ is comparable in magnitude to the $N\langle \alpha_D(N) \rangle \langle \vec{E}_{\text{loc}}(N) \rangle$ term.

On the other hand, the Kramers-Kronig relation [Eq. (21)] predicts a specific result depending on the frequency dependence of the conductivity. The heavy weighting of the low-frequency ($\omega \to 0$) portion of the spectrum implies that, if $\sigma(\omega) \propto \omega^s$ as $\omega \to 0$, one obtains a divergence of $\epsilon(\omega = 0, N \to N_c)$ if $s < 1$ and no divergence if $s > 1$. For the case of localized states at the Fermi level, no compensation (no empty impurity sites), and $T = 0$ K, Mott[46] has shown, employing the Kubo-Greenwood approach, that the low-frequency conductivity $\sigma(\omega)$ has a frequency dependence $\omega^2[\ln(1/\omega)]^4$. This result does not yield a divergence of $\epsilon(\omega = 0, N \to N_c)$ and therefore supports the idea of a discontinuity in $\epsilon(N)$ at N_c characteristic of a first-order abrupt transition. However, the role of electron-electron interactions and self-consistent screening have apparently not yet been incorporated into Mott's calculation. From the experimental viewpoint the question of a smooth PC as $N \to N_c$ at $T = 0$ K is an extremely difficult one to resolve, even with dilution refrigerator experiments. Even the smallest amount of compensation for $N < N_c$ can lead to a characteristic hopping conductivity $\sigma(\omega) \propto \omega^s$ with $s < 1$, thus leading to a divergence of $\epsilon(0, N \to N_c)$. Although this hopping contribution is supposed to vanish as $T \to 0$, this depends critically on the distribution $p(\Delta E)$ of hopping energies ΔE. If $p(\Delta E)$ remains finite as $\Delta E \to 0$, one has the possibility of $\sigma(\omega, T \to 0)$ remaining finite. This point requires further experimental and theoretical investigation close to N_c. At present the question of a smooth [or a very large rise, e.g., $\epsilon(N \to N_c) > 100\epsilon_h$] or abrupt polarization catastrophe as $N \to N_c$ has not been resolved.

The Mott criterion $N_c^{1/3}a \approx \frac{1}{4}$ has led to many calculations using a variety of types of screening. Table III shows the calculated and experimental values of $N_c^{1/3}a(0)$ obtained in this work. The values of $a(0)$ were obtained by force-fitting the theoretical expression for $\alpha_D(0)$ to the dilute experimental values $\alpha_{D\text{-dilute}} \sim \alpha_D(0)$. The values of $N_{c\text{-calc}}$ correspond to the concentration for which $\epsilon(N \to N_c) \to \infty$ (the values in parentheses based on the smaller values of $\alpha_{D\text{-dilute}}$ are considered the more reliable values). These calculated values average about 24% larger than the experimental values. The experimental values are in excellent agreement with the Edwards and Sienko tabulation yielding $N_c^{1/3}a = 0.26 \pm 0.05$. Although the comparison between theory and experiment is not as good as for the Hubbard model [$N_c^{1/3}a = 0.25$ for six conduction-band valleys, Ref. 11], the additional contributions from donor randomness and charge

TABLE III. Mott criterion $N_c^{1/3}a$ values for silicon.

	EMA	Sb	P	As
$a(0)$	20.18	19.90[a] (17.25)[b]	18.49[a] (16.23)[b]	14.57[a] (13.51)[b]
$N_{c\text{-calc}}{}^{1/3}a(0)$	0.3016	0.322[a] (0.314)[b]	0.321[a] (0.314)[b]	0.309[a] (0.299)[b]
$N_{c\text{-expt}}{}^{1/3}a(0)$		0.287 (0.249)	0.281 (0.246)	0.270 (0.250)

[a] The values are based on the $\alpha_D(0)$ values from Ref. 42.

[b] These values are based on new smaller values of $\alpha_D(0)$; see Ref. 43 and Ref. 9 for Si:As.

transfer might bring down the $N_{c\text{-calc}}^{1/3}a(0)$ values another 20% and make the agreement with experiment look much better. It should be noted that the Hubbard model is oversimplified and its application by Berggren[24] and Edwards and Sienko[11] to n-type semiconductors neglects the donor randomness and self-consistency (the Bohr radius is not varied with N). The good agreement of the Hubbard model with experiment might well be viewed as fortuitous, considering the many approximations employed in the Hubbard model.

VI. CONCLUSIONS

The donor-polarizability enhancement resulting from screening based on a continuum-model calculation is shown to be a large effect, ranging from 1.88 for the EMA case to about 3 for Si:As. The calculation shows the enhancement results from (1) an increase in the Bohr radius $a(N)$ with N, and (2) an increase in the effective dielectric constant $\langle\epsilon(r,N)\rangle$ seen by the localized donor electron. Although the polarizability enhancement gives a significant upward deviation from Clausius-Mossotti behavior, the calculated result is still too small to explain the upward deviation for the Si:As data. The corrected self-consistent Herzfeld criterion still yields N_c values, a factor of 1.6 to 2, too large for the Si shallow donors. The addition of donor randomness and charge transfer is expected to improve the agreement between theory and experiment; however, the question of a smooth polarization catastrophe $[\epsilon(N\rightarrow N_c)\rightarrow\infty$, no discontinuity$]$ is not yet resolved.

ACKNOWLEDGMENTS

The author is grateful to R. Bitler for assistance on the early numerical calculations on the polarizability enhancement, to H. S. Tan for useful discussions, and to Professor D. L. Dexter for bringing to his attention the Hassê variational approach. Stimulating discussions with Dr. D. Agassi concerning the many-body aspects of dielectric theory are gratefully acknowledged. This work was supported in part by NSF Grant No. DMR77-20450.

APPENDIX A

The Coulomb interaction energy between a pair of neutral hydrogenic atoms, one with an electron in a state $\psi_a(\vec{r}_1)$ centered on atom a and a second with an electron in state $\psi_b(\vec{r}_2)$ centered on atom b, is given by

$$E_c^P=\frac{e^2}{\epsilon_h}\int\int|\psi_a(\vec{r}_1)|^2\,|\psi_b(\vec{r}_2)|^2 d\vec{r}_1\,d\vec{r}_2 \times\left(\frac{1}{r_{12}}-\frac{1}{r_{1b}}-\frac{1}{r_{2a}}+\frac{1}{R}\right), \tag{A1}$$

where $|\vec{R}|$ is the distance between the two atoms at a and b, $r_{12}=|\vec{r}_1-\vec{r}_2|$, $r_{1b}=|\vec{r}_1-\vec{R}|$, and $r_{2a}=|\vec{r}_2-\vec{R}|$. The effective Coulomb interaction potential seen by electron 1 at $\vec{r}_1$ due to the second hydrogenic atom is obtained by integrating over the coordinates of electron 2, namely,

$$V_{c\text{-int}}^P(\vec{r}_1)=\frac{e^2}{\epsilon_h}\int|\psi_b(\vec{r}_2)|^2 d\vec{r}_2 \times\left(\frac{1}{r_{12}}-\frac{1}{r_{1b}}-\frac{1}{r_{2a}}+\frac{1}{R}\right), \tag{A2}$$

where only the first and third terms in the square brackets are functions of $\vec{r}_2$. Using $\psi_b(\vec{r}_2)=(\pi a^3)^{-1/2}e^{-r_2/a}$ the integration yields

$$V_{c\text{-int}}^P(\vec{r}_1)=\frac{e^2}{\epsilon_h}\left[e^{-2R/a}\left(\frac{1}{R}+\frac{1}{a}\right)-e^{-2|\vec{r}_1-\vec{R}|/a}\left(\frac{1}{|\vec{r}_1-\vec{R}|}+\frac{1}{a}\right)\right]. \tag{A3}$$

The latter two terms containing the r_1 dependence were employed by Leroux Hugon and Ghazali[25] in their calculation. The first two terms represent a constant interaction term. Note also that $V_c^P(\vec{r}_1)\rightarrow 0$ as $\vec{r}_1\rightarrow 0$. To find the total pairwise interaction of the electron on atom a at $\vec{r}_1$ for a continuum distribution of neutral donors $V_{c\text{-int}}^P(\vec{r}_1,\vec{R})$ is integrated over this continuum distribution beyond a concentration-dependent cutoff radius d, thus obtaining

$$V_{c\text{-int}}(r_1,d)=\int_d^\infty V_{c\text{-int}}^P(\vec{r}_1,\vec{R})N d\vec{R}, \tag{A4}$$

where d is determined by $4\pi d^3N/3=1$. Performing the integration for the two cases $|\vec{r}_1|<d$ and $|\vec{r}_1|>d$ one finds

$$V_{c\text{-int}}(r_1,d)=(e^2/\epsilon_h)2\pi Na^2e^{-2x}[2(1+x)^2-(5+10x+4x^2)(e^{2\rho}-e^{-2\rho})/16\rho+\tfrac{1}{8}(2x+1)(e^{2\rho}+e^{-2\rho})] \tag{A5a}$$

for $r_1<d$ and

$$V_{c\text{-int}}(r_1,d)=(e^2/\epsilon_h)2\pi Na^2[2e^{-2x}(1+x)^2-1-(e^{-2\rho}/4\rho)(e^{2x}-e^{2x})(x^2+\tfrac{1}{2}\rho+\tfrac{5}{4}) +(e^{-2\rho}/4\rho)(e^{2x}+e^{-2x})(\tfrac{5}{2}x+\rho x)] \tag{A5b}$$

for $r_1 > d$, where $\rho = r_1/a$ and $x = d/a$. For $N < N_c$ one has $x > 2.5$. As $r_1 \to \infty$, $V_{c\text{-int}}(r_1)$ approaches the value

$$V_{c\text{-int}}(r_1 \to \infty) = (e^2/\epsilon_h)2\pi N a^2 (2e^{-2x}(1+x)^2 - 1]$$
$$\approx -(e^2/\epsilon_h)2\pi N a^2 , \qquad \text{(A6)}$$

which is just the negative constant background potential [V_∞ in Eq. (10)] from the continuum of donors. For $r_1 \gg d$, $[V_{c\text{-int}}(\vec{r}_1) - V_{c\text{-int}}(r_1 \to \infty)] \propto e^{-2\rho}$ so that this interaction potential minus V_∞ falls off exponentially with r_1. As $r_1 \to 0$,

$$V_{c\text{-int}}(r_1 \to 0) = (e^2/\epsilon_h)(2\pi N a^2) e^{-2x}(1+x)^2 ,$$

which is slightly positive.

APPENDIX B

For the hydrogenic trial wave function $\psi_0(r) = [\pi a(N)^3]^{-1/2} e^{-r/a(N)}$ the matrix elements $\lambda_1 = \langle z\psi_0 | z\psi_0 \rangle$, $\lambda_2 = \langle zr\psi_0 | zr\psi_0 \rangle$, and $\lambda_3 = \langle zr\psi_0 | z\psi_0 \rangle$ are given, respectively, by $\lambda_1 = a(N)^2$, $\lambda_2 = \frac{15}{2} a(N)^4$, and $\lambda_3 = \frac{5}{2} a(N)^3$. The matrix elements $A = \langle z\psi_0 | H_t | z\psi_0 \rangle$, $B = \langle zr\psi_0 | H_t | zr\psi_0 \rangle$, and $C = \langle zr\psi_0 | H_t | z\psi_0 \rangle$ are given, respectively, by

$$A = \frac{\hbar^2}{2m^*} - \frac{e^2 a}{2\epsilon_h}\left[\frac{\epsilon_h}{\epsilon(N)} + \left(1 - \frac{\epsilon_h}{\epsilon(N)}\right) y^4 + \epsilon_h \lambda_D z^4\right] , \qquad \text{(B1a)}$$

$$B = \frac{7}{4}\frac{\hbar^2}{m^*}a^2 - \frac{5}{2}e^2\frac{a^3}{\epsilon_h}\left[\frac{\epsilon_h}{\epsilon(N)} + \left(1 - \frac{\epsilon_h}{\epsilon(N)}\right) y^6 + \epsilon_h \lambda_D z^6\right] , \qquad \text{(B1b)}$$

and

$$C = \frac{3}{4}\frac{\hbar^2}{m^*}a - \frac{e^2 a^2}{2\epsilon_h}\left[\frac{\epsilon_h}{\epsilon(N)} + \left(1 - \frac{\epsilon_h}{\epsilon(N)}\right) y^5 + \epsilon_h \lambda_D z^5\right] , \qquad \text{(B1c)}$$

where $y = 1/[1 + a(N)/2r_s]$ and $z = 1/[1 + k_D a(N)/2]$. The $\lambda_D z^n$ ($n = 4$, 5, and 6) contributions represent the central-cell corrections to A, B, and C which turn out to be totally negligible since $z \ll 1$. In fact $x(N) = \lambda_D z(N)^2$ is a small quantity [see Table I for values of $x(0)$] and $x(N)\,z(N)^2$ will be less than 10^{-3} for all donors. The central-cell correction is important for the single-valley energy $E_0(N)$ [see Eq. (14)] but not for the matrix elements A, B, and C which enter $\alpha_D(N)$. Neglecting the central-cell terms and using $\hbar^2/m^* = e^2 a_e/\epsilon_h$, the matrix elements A, B, and C reduce to the expressions

$$A = \left(\frac{e^2}{2\epsilon_h a_e}\right) a_e a \left[\frac{a_e}{a} - 1 + \left(1 - \frac{\epsilon_h}{\epsilon(N)}\right)(1 - y^4)\right] , \qquad \text{(B2a)}$$

$$B = \left(\frac{e^2}{2\epsilon_h a_e}\right) a_e a^3 \left[\frac{7}{2}\frac{a_e}{a} - 5 + 5\left(1 - \frac{\epsilon_h}{\epsilon(N)}\right)(1 - y^6)\right] , \qquad \text{(B2b)}$$

and

$$C = 2\left(\frac{e^2}{2\epsilon_h a_e}\right) a_e a^2 \left[\frac{3}{2}\frac{a_e}{a} - 2 + 2\left(1 - \frac{\epsilon_h}{\epsilon(N)}\right)(1 - y^5)\right] . \qquad \text{(B2c)}$$

---

[1] For recent reviews, see *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); N. F. Mott, *Metal Insulator Transitions* (Taylor and Francis, London, 1974).

[2] See, for example, the chapters by H. Fritzsche, M. Pollak, and L. F. Friedman, in *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).

[3] See the chapters by D. F. Holcomb and H. Kamimura, in *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).

[4] P. Townsend, J. Phys. C 11, 1481 (1978).

[5] M. Capizzi, G. A. Thomas, G. Devlin, F. DeRosa, J. Mock, and S. J. Allen, Jr., in *Proceedings of the Fourteenth International Conference on the Physics of Semiconductors, Edinburgh, 1978*, edited by B. L. H. Wilson (Institute of Physics, London, 1978), p. 957.

[6] J. Golka, E. Pavlidou, and L. Piela, in *Proceedings of the Fourteenth International Conference on Physics of Semiconductors, Edinburgh, 1978*, edited by B. L. H. Wilson (Institute of Physics, London, 1978), p. 953.

[7] F. A. D'Altroy and H. Y. Fan, Phys. Rev. 103, 1671 (1956).

[8] D. W. Mahaffey and D. A. Jerde, Rev Mod. Phys. 40, 710 (1968).

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

[10] N. F. Mott, Proc. Phys. Soc. London A62, 416 (1949).

[11] P. P. Edwards and M. J. Sienko, Phys. Rev. B 17, 2575 (1978).

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

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

[14] S. P. Li, W. F. Love, and S. C. Miller, Phys. Rev. 162, 728 (1967).

[15] See H. Brooks, in *Advances in Electronics and Electron Physics*, edited by L. Marton (Academic, New York, 1955), Vol. 7, p. 85.

[16] K. Colbow and D. Dunn, Philos. Mag. 22, 237 (1970).

[17] K. Colbow and D. W. Nyberg, Phys. Lett. A25, 250 (1967).

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

[19] G. F. Neumark, Phys. Rev B 5, 408 (1972); 20, 1519

(1979).

[20] For a review of Clausius-Mossotti relationship and the early references, see R. Landauer, in *Proceedings of the First Conference on Electrical Properties of Inhomogeneous Media, Ohio State University, 1977,* edited by J. C. Garland and D. B Tanner (American Institute of Physics, New York, 1978), No. 40.

[21] K. F. Herzfeld, Phys. Rev. 29, 701 (1927).

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

[23] D. G. H. Frood, Proc. Phys. Soc. London 75, 185 (1960).

[24] K. F. Berggren, J. Chem. Phys. 60, 3399 (1974).

[25] P. Leroux Hugon and A. Chazali, Phys. Rev. B 14, 602 (1976); J. Phys. (Paris) Colloq. 4, 323 (1976).

[26] M. Takeshima, Phys. Rev. B 17, 3996 (1978).

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

[28] H. Ehrenreich and M. H. Cohen, Phys. Rev. 115, 786 (1959).

[29] N. Wiser, Phys. Rev. 129, 62 (1963).

[30] H. R. Hassé, Proc. Cambridge Philos. Soc. 26, 542 (1930).

[31] J. C. Slater, *Quantum Theory of Molecules and Solids* (McGraw-Hill, New York, 1969), Vol. 3, pp. 111 and 112.

[32] C. J. F. Böttcher, *Theory of Electric Polarization*, 2nd ed. (Elsevier, Amsterdam, 1973), pp. 170–172.

[33] J. H. Van Vleck, J. Chem. Phys. 7, 72 (1939).

[34] See, for example, P. R. Cullis and J. R. Marko, Phys. Rev. B 1, 632 (1970). This reference includes the early important references on this subject.

[35] C. J. F. Böttcher, Physica 9, 945 (1942); see also pp. 228 and 239 of Ref. 32.

[36] J. de Boer, F. van der Maesen, and C. A. Ten Seldam, Physica 19, 265 (1953).

[37] J. G. Kirkwood, J. Chem. Phys. 4, 592 (1936).

[38] R. Van Metter (unpublished).

[39] A critique of the perturbation approach as applied by Leroux Hugon and Ghazali (Ref. 25) has been given by T. G. Castner, Phys. Rev. B 20, 3505 (1979).

[40] H. Haken, Z. Naturforsh. 9a, 228 (1954); Nuovo Cimento [10], 3, 1230 (1956).

[41] D. L. Dexter, in *Proceedings of the Thirteenth International Conference on the Physics of Semiconductors, Rome, 1976,* edited by F. G. Fumi (North-Holland, Amsterdam, 1976), p. 1137. (The Ge result is unpublished).

[42] J. Bethin, T. G. Castner, and N. K. Lee, Solid State Commun. 14, 1321 (1974).

[43] H. S. Tan and T. G. Castner (unpublished).

[44] M. Capizzi, G. A. Thomas, F. DeRosa, R. N. Bhatt, and T. M. Rice, Solid State Commun. 31, 611 (1979).

[45] S. R Nagel and T. A. Witten, Jr., Phys. Rev. B 11, 1623 (1975).

[46] N. F. Mott, Philos. Mag. 22, 7 (1970).