Multi-ion Screening in Uncompensated Semiconductors

J. R. Meyer and F. J. Bartoli
U.S. Naval Research Laboratory, Washington, D.C. 20375
(Received 19 May 1986)

We have developed a self-consistent multi-ion screening formalism which, unlike conventional treatments, explicitly accounts for the fact that free electrons cannot screen a given donor in a multi-ion system as effectively as they can screen the same donor in a single-ion system. In uncompensated, metallic Si:P at low temperatures, employment of the "multi-ion" screening length significantly improves the agreement between theoretical and experimental electron mobilities.

PACS numbers: 71.55.Fr, 72.20.Jv

In this paper, we present a self-consistent formalism for the screening of ionized impurities in an uncompensated semiconductor. In contrast to previous treatments of the problem, we explicitly account for the fact that free electrons cannot screen a given donor in a multi-ion system as effectively as they can screen the same donor in a single-ion system. On the basis of the corrected scattering potentials, we calculate improved theoretical electron mobilities as a function of N_D for low-temperature n-type Si.

The inadequacy of previous treatments is most easily understood if we first consider the "tight-screening" regime, where the screening length calculated by the conventional theory is much shorter than the distance between impurities and neighboring potentials do not overlap significantly. This occurs, for example, in heavily doped n-type Si at low temperatures. (At $N_D = 10^{19}$ cm^{-3}, the average distance between impurities is nearly seven times the Thomas-Fermi screening length.) The final results of this paper, however, will be in a form which applies to any uncompensated semiconductor.

Consider a large sphere of volume V containing a single donor as well as a free-electron gas of density n. The excess charge density at position $\mathbf{r}$ due to redistribution of the wave function for state i is

$$\bar{\rho}_i(\mathbf{r}) = -e[|\Psi_i(\mathbf{r})|^2 - |\Psi_i^F(\mathbf{r})|^2], \tag{1}$$

where Ψ_i is the perturbed wave function for the state and Ψ_i^F is the free-electron wave function. Equation (1) may be summed over electron states to give the total excess charge at position $\mathbf{r}$:

$$\rho(\mathbf{r}) = \sum_i \bar{\rho}_i(\mathbf{r}) f_0(E_i), \tag{2}$$

where f_0 is the Fermi distribution function and E_i is the energy of state i. We obtain $\bar{\rho}_i(\mathbf{r})$ from a lowest-order Kohn-Sham calculation.[1] In this method one solves Poisson's equation and the Schrödinger wave equations for the various electron states simultaneously. Although the general solution requires a considerable amount of numerical computation,[2,3] it is useful to consider an analytic form which is valid to terms linear in the potential. It will be shown in a separate paper that if exchange and correlation effects are ignored, one obtains[4]

$$\bar{\rho}_i(\mathbf{r}) \approx e\phi(\mathbf{r})/2VE_i, \tag{3}$$

where $\phi(\mathbf{r})$ is the screened potential. It is easily demonstrated that when this form is employed in the sum over electron states in Eq. (2), the charge density $\rho(\mathbf{r})$ reverts to the conventional linearized Thomas-Fermi (LTF) result.[5] The solution to Poisson's equation therefore yields a screened Coulomb potential,

$$\phi(\mathbf{r}) = -e^2 e^{-|\mathbf{r}-\mathbf{r}_j|/\lambda}/\kappa_0|\mathbf{r}-\mathbf{r}_j|, \tag{4}$$

where λ is the screening length, κ_0 is the static dielectric constant, and $\mathbf{r}_j$ is the position of the impurity. In the LTF approximation, λ is given by[5]

$$\lambda_0^{-2} = \frac{4\pi n e^2}{\kappa_0 k_B T} \frac{\mathcal{F}_{-1/2}(\eta)}{\mathcal{F}_{1/2}(\eta)}, \tag{5}$$

where $\mathcal{F}_p$ is the Fermi integral of order p and $\eta = E_F/k_BT$ is the reduced Fermi energy. While LTF provides a useful lowest-order approximation for the single-ion potential, we now show that if one follows the usual procedure for generalizing to a multi-ion system, physically unreasonable results are obtained.

For a single donor at position $\mathbf{r}_j$, the total screening charge contributed by state i is given by

$$q_{ij} = \int d^3r\, \bar{\rho}_i(\mathbf{r}) \rightarrow -2\pi e^3\lambda^2/V\kappa_0 E_i, \tag{6}$$

where the latter result is obtained following evaluation of the integral with use of Eqs. (3) and (4). In LTF, the net multi-ion potential is simply a linear superposition of single-ion potentials:

$$\phi_T(\mathbf{r}) = -\sum_{j=1}^{\mathcal{N}_D} \frac{e^2 e^{-|\mathbf{r}-\mathbf{r}_j|/\lambda}}{\kappa_0|\mathbf{r}-\mathbf{r}_j|}, \tag{7}$$

where $\mathcal{N}_D$ is the number of donors in the system. The total screening charge contributed by a given electron is then

$$Q_i = \sum_{j=1}^{\mathcal{N}_D} q_{ij} \rightarrow -\frac{2\pi N_D e^3\lambda^2}{\kappa_0 E_i}, \tag{8}$$

where $N_D = \mathcal{N}_D/V$ is the donor density, which is equal to

n for an uncompensated semiconductor.

In LTF, the weighted sum of Q_i over electron states gives a net screening charge which exactly neutralizes the donor charges,

$$\sum_i (-Q_i/e) f_0(E_i) = \mathcal{N}_D. \tag{9}$$

However, a closer examination of Eq. (8) shows that this leads to an inconsistency. From Eq. (8) one finds that $-Q_i/e < 1$ for high-energy electrons; that is, a given electron's total charge $-e$ cannot be completely devoted to screening. On the other hand, for low-energy electrons $-Q_i/e > 1$, which indicates that the assumed contribution to the screening is greater than the available charge. This is physically unreasonable since we are specifically considering the regime where LTF theory predicts no significant overlap of neighboring donor potentials (the same electron may not contribute to the screening of more than one donor).

In a more realistic calculation of the screening, we must require that $-Q_i/e \leq 1$ for all electrons. The simplest way of accomplishing this is to multiply the charge density for state i in Eq. (3) by an additional factor S_i:

$$\bar{\rho}_i'(\mathbf{r}) \rightarrow S_i \bar{\rho}_i(\mathbf{r}), \tag{10}$$

where

$$S_i(E_i) = \begin{cases} \mathcal{E}/E_i, & E_i < \mathcal{E}, \\ 1, & E_i \geq \mathcal{E}, \end{cases} \tag{11}$$

and

$$\mathcal{E} = E_q' = 2\pi N_D e^2 \lambda^2/\kappa_0. \tag{12}$$

Using Eqs. (6) and (8), one immediately obtains for the corrected screening charge $Q_i' = S_i Q_i$.

Since $-Q_i/e$ for low-energy electrons has been effectively decreased (it may no longer exceed unity), we find that Eq. (9), which assures the complete screening of all donors, is no longer satisfied. While each low-energy electron ($E < E_q'$) now contributes a charge of exactly $-e$ to the screening of the donors, the high-energy electrons ($E > E_q'$) do not screen effectively and each contributes less than its total charge.[6] Since $n = N_D$, this means that the donors will not be fully screened. Simply changing the screening length from λ_0 to some higher value λ' does not remove the apparent difficulty, since no matter how much the screening length (and hence E_q') is increased, at finite temperatures there will always be electrons in the high-energy tail of the Fermi distribution which are not fully devoted to screening. We conclude that Eq. (9) can never be satisfied. However, the failure to satisfy Eq. (9) is not a drawback as long as neighboring potentials overlap somewhat. At distances greater than half the average interdonor separation, $D = (4\pi N_D/3)^{-1/3}$, it is unnecessary to insist that each donor be fully screened, only that overall charge neutrality be preserved.[7] This considerably relaxes the magnitude of the total screening charge required, and makes it unnecessary to satisfy Eq. (9).

Although we do not require that each donor be "fully" screened independently of the others, there must still be enough electron charge available to provide screening out to the overlap region. "Complete" screening of a given donor would require that the excess electron screening charge integrated over all space add up to exactly $-e$. Here we define g to be that fraction of the total which would be contained within a sphere of radius D surrounding the donor. Using the fact that $\rho(\mathbf{r}) \propto \phi(\mathbf{r})$, we obtain

$$g = 4\pi \int_0^D \phi(r) r^2 dr / 4\pi \int_0^\infty \phi(r) r^2 dr \rightarrow 1 - (1 + D/\lambda_M) e^{-D/\lambda_M}, \tag{13}$$

where the "multi-ion" screening length λ_M will be defined below. We now rewrite Eq. (9) to reflect the relaxed requirement that the total donor charge to be screened has been reduced by a factor of g:

$$\sum_i (-Q_i''/e) f_o(E_i) = g \mathcal{N}_D. \tag{14}$$

The restricted region of integration also causes the electron charge per donor q_{ij} to be smaller by the same factor. By use of Eqs. (6), (8), and (10), the total screening charge for a given state i becomes

$$Q_i'' = -2\pi g N_D e^3 \lambda^2 S_i/\kappa_0 E_i. \tag{15}$$

Again, we must not allow this charge to exceed $-e$. In Eq. (11) for S_i, we therefore set

$$\mathcal{E} = E_q'' = 2\pi g N_D e^2 \lambda_M^2/\kappa_0. \tag{16}$$

The screening charge requirements may now be fulfilled by our adjusting the screening length λ_M until Eq. (14) is satisfied. This condition may be written in the compact form

$$\lambda_M^{-2} = \Gamma \lambda_0^{-2}, \tag{17}$$

where

$$\Gamma = \frac{\int_0^{z_q} (z^{1/2}/z_q) f_0 dz + \int_{z_q}^\infty z^{-1/2} f_0 dz}{\int_0^\infty z^{-1/2} f_0 dz}, \tag{18}$$

and

$$z_q \equiv \frac{E_q''}{k_B T} = \frac{2\pi g N_D e^2 \lambda_M^2}{\kappa_0 k_B T}. \tag{19}$$

We refer to λ_M as the "multi-ion" screening length because its value is quite sensitive to the inter-donor spacing (through the factor g).

It should be observed here that we have shown the "tight-screening" limit to be unphysical, since Eq. (14) can never be satisfied unless neighboring impurity potentials have at least a modest amount of overlap. We conclude that ionized-impurity scattering in a semiconductor is intrinsically a multi-ion process.

Free-carrier transport properties may now be calculated by the usual methods, except that the corrected screening length λ_M should be used instead of λ_0. We consider the specific case of electron mobilities in heavily doped Si:P in the low-temperature limit. Ionized-impurity scattering has been treated by the partial-wave phase-shift method.[8,9] Silicon parameters were the same as those employed in Ref. 9, where temperature-dependent mobilities were calculated for samples having varying degrees of compensation. For electron densities between the critical density for the metal-insulator transition, n_c ($\approx 3.8\times10^{18}$ cm^{-3}), and 3×10^{20} cm^{-3}, the λ_M which satisfies Eq. (17) is found to be 1.2 to 1.4 times λ_0. Furthermore, the parameter $b=4k^2\lambda^2$ is on the order of unity in this regime, which means that the calculated mobility is quite sensitive to changes in the screening length.[8] This is illustrated in Fig. 1, which compares mobilities obtained with use of the "multi-ion" screening length λ_M (solid curve) with results of the conventional theory (dashed curve) and with experiment.[10-12] The figure shows that use of the conventional screening length leads to mobilities which are too high by as much as a factor of 4 near n_c. The mobility correction is largest at the lower densities where the small electron Fermi wave vector leads to smaller b and hence scattering cross sections which are much more sensitive to changes in the screening length.

The agreement with experiment is much better when the more consistent treatment of the screening is employed, although it should be remembered that there are additional higher-order effects which can lead to theoretical uncertainties.[13] For example, Krieger *et al.*[12,14] have suggested that the large discrepancy between the conventional theory and experiment may be due to the effect of conduction-band anisotropy on the screening. However, the more recent results of Saso and Kasuya for germanium,[3] which has an electron effective mass even more anisotropic than that of silicon, seem to indicate that screening anisotropy has a relatively small effect on the calculated mobility. Since we have assumed the electrons to be free and have ignored localization effects, the theoretical curves in Fig. 1 naturally do not reproduce the precipitous drop of the mobility with decreasing electron density starting at n slightly above n_c.

Summarizing, we have shown that conventional treatments, which assume that the screening length calculated for a single-ion system will be appropriate when applied to a multi-ion system, overestimate the ability of the free electrons to screen all ions at the same time. To correct this shortcoming, we have developed a simple, self-consistent screening formalism based on the requirement that no electron state should contribute more than $-e$ to the net screening charge. In some regimes, the correction to the screening length has a significant effect on calculated electron mobilities. The altered screening length also has implications for the nature of the metal-insulator transition. In a future paper, we will discuss the generalization to compensated semiconductors.

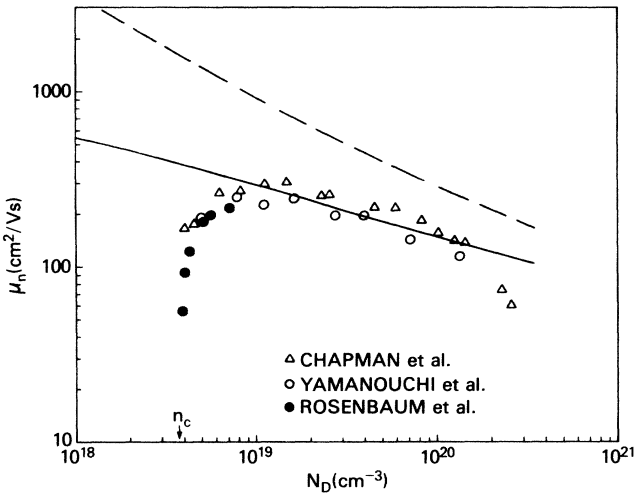

FIG. 1. Theoretical (curves) and experimental (Refs. 10–12) (points) electron mobilities for uncompensated Si:P extrapolated to the low-temperature limit. The dashed curve was obtained with use of the conventional screening length λ_0 while the solid curve was obtained with use of the multi-ion screening length λ_M. The theoretical curves were corrected for the Hall factor r_H (including anisotropy).

[1]W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[2]P. Jena and K. S. Singwi, Phys. Rev. B **17**, 3518 (1978).

[3]T. Saso and T. Kasuya, J. Phys. Soc. Jpn. **48**, 1566 (1980), and **49**, 578 (1980).

[4]The Kohn-Sham approximation (see Ref. 3) for V_{xc} contributes an additional term to the Schrödinger equation, which in lowest order has the same functional form as the screened Coulomb potential. For an attractive interaction, V_{xc} adds to the effective strength of the potential, making the multi-ion phenomena discussed below even more important. Similarly, errors introduced into Eq. (3) by the retention of only linear terms will be of second order, since they affect the functional form of $\bar{\rho}_i(\mathbf{r})$ but not the basic relation that low-energy electrons contribute much more to the screening than high-energy electrons.

[5]R. B. Dingle, Philos. Mag. **46**, 831 (1955).

[6]The "leftover" charge may be thought of as providing a uniform background which preserves the overall charge neutrality of the system.

[7]J. R. Meyer and F. J. Bartoli, Phys. Rev. B **31**, 2353 (1985). In the severe overlap regime ($\lambda_s \gg D$), the scattering by individual impurities is unimportant and the principal mechanism is scattering by random potential fluctuations. While the electrons adjust their positions so as to minimize the magnitude of the fluctuations, it is not meaningful to specify that a given impurity with many overlapping neighbors is "fully screened" at large distances by some particular distribution of electron charge. The same reasoning implies that even if the

impurity potentials are only slightly overlapping, it is unnecessary to require "complete" screening of a given donor at distances beyond the overlap region.

[8]J. R. Meyer and F. J. Bartoli, Phys. Rev. B **23**, 5413 (1981).

[9]J. R. Meyer and F. J. Bartoli, Phys. Rev. B **24**, 2089 (1981).

[10]P. W. Chapman, O. N. Tufte, J. D. Zook, and D. Long, J. Appl. Phys. **34**, 3291 (1963).

[11]C. Yamanouchi, K. Mizuguchi, and W. Sasaki, J. Phys. Soc. Jpn. **22**, 859 (1967).

[12]T. F. Rosenbaum, R. F. Milligan, M. A. Paalanen, G. A. Thomas, R. N. Bhatt, and W. Lin, Phys. Rev. B **27**, 7509 (1983).

[13]D. Chattopadhyay and H. J. Queisser, Rev. Mod. Phys. **53**, 745 (1981).

[14]J. B. Krieger, J. Gruenebaum, and T. Meeks, Phys. Rev. B **9**, 3627 (1974).