Self-energy shifts in heavily doped, polar semiconductors

Bo E. Sernelius*

Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831
and Department of Physics, University of Tennessee, Knoxville, Tennessee 37996
(Received 24 February 1987)

We present theoretical results for the self-energy shifts due to the interactions in a heavily doped, polar semiconductor. For the purpose of numerical demonstration we apply the results to n-type CdS. The density of states for both the majority and minority carriers are calculated in the region near the band edges. The polar coupling modifies the self-energy shifts, and new structures appear in the density of states, both above and below the Fermi level. At least some traces of these structures should be observed in experiments. Furthermore, the doping-induced band-gap narrowing is determined and we compare the obtained results with those from the much simpler so-called ϵ_0 approximation.

I. INTRODUCTION

In a heavily doped semiconductor the dopant carriers occupy the states near the band edges of the host semiconductor. For n-type doping the majority carriers are electrons at the bottom of the conduction band (CB) and for p-type doping they are holes at the top of the valence band (VB). The correlation between the carriers and also with the dopant ions gives rise to energy shifts for the states, and causes, e.g., band-gap narrowing (BGN). The energy dispersions of the bands, in particular near the band edges, are modified, which produces deformations of the density of states (DOS).

The BGN has attracted an increasing experimental and theoretical interest in the recent past. We have earlier, in Refs. 1–5, addressed the BGN in the technologically very important semiconductors Si, Ge, and GaAs. The reader is referred to these references and references therein for a more general background. Reference 6 gives a review of the important experimental and theoretical results from before 1978 and Refs. 7–15 contain some of the key results of more recent dates. Si and Ge are nonpolar and the polar coupling in GaAs is so weak that it can be neglected.

In this work we study the effects from the polar coupling. The theory that we have used before, for nonpolar semiconductors, is here extended to take the polar coupling into account. We have chosen to present numerical results for CdS for two reasons. Firstly, this semiconductor has a strong enough polar coupling to produce detectable effects and, secondly, it has a band structure which is as simple as possible. CdS is a direct-gap semiconductor and its VB edge is nondegenerate near the Γ point, where the valence bands are split by spin-orbit interactions and crystal-field effects. This simple band structure makes the calculations easier and the physics more transparent. The results are not complicated by the coupling between degenerate valence bands as in, e.g., GaAs.[4] The polar coupling is expected to give rise to structures in the DOS

which might be important enough to be detectable experimentally. The purpose of this work is to investigate this point and also to find out how good the so-called ϵ_0 approximation is in the calculation of the BGN for polar semiconductors. In that approximation the semiconductor is treated as nonpolar and the static dielectric function is used as a background screening constant.

The details of the effects from the polar coupling on the self-energy shifts, apart from being of theoretical interest, are of great technological importance. In one field of semiconductor applications, heavy doping is a way to tailor the optical properties of semiconductor films. The reason is as follows. A large band-gap semiconductor is transparent for photon energies in the visible region and below. Heavy doping has two main effects. One is that the optical band gap is changed. This change depends on two competing effects. Many-body effects tend to reduce the gap, where this reduction is counteracted by a band-gap widening due to the blocking of the lowest states in the conduction band (for n-type doping). This widening is known as the Burstein-moss effect. In the extreme high doping limit, the Burstein-Moss shift wins over the BGN from many-body effects and the gap increases. This moves the upper boundary for the optical window upwards. The second effect from the doping is to introduce the free-carrier absorption at the low-energy side, moving the lower boundary of the window upwards. Thus these heavily doped semiconductor films have interesting optical properties. They have the properties of "dirty" metals at the low-energy side and that of semiconductors on the high-energy side. They can be used, e.g., as coatings on highly energy efficient windows, giving good solar transmittance and low thermal emittance.[16-18] The large band-gap semiconductors are usually polar; hence it is important to get the effects from the polar coupling correct. In a recent work,[19] we studied the effects from this coupling on the optical properties near the lower boundary of the optical window. The treatment here is related more to the properties near the upper boundary.

In Sec. II we give a brief description of the energetics of the dopant system, show how the polar coupling enters the problem, and derive the expressions for the self-energy shifts. Numerical results are presented for states in both the conduction and valence bands of n-type CdS. We show, further, the obtained DOS for both the majority and minority carriers and demonstrate the modifications caused by the polar coupling.

Section III is devoted to the BGN. We introduce three different BGN's and present numerical results for CdS from the full calculation and from the calculation in the ϵ_0 approximation. Finally, in Sec. IV we present a brief summary and conclusions.

II. INTERACTIONS AND SELF-ENERGY SHIFTS IN A HEAVILY DOPED, POLAR SEMICONDUCTOR

In this section we study the energy of the dopant system in a heavily n-type doped, polar semiconductor and derive from this the quasiparticle energies and self-energies. We use the same formalism as in our earlier publications in this field.[1-5] What is new here is the inclusion of the coupling to the optical phonons.

We include in the system the released donor electrons, occupying the states at the bottom of the host CB, the donor ions, assumed to be randomly distributed, and the longitudinal optical phonons. We also include a small number of holes in the VB. These are included for the purpose of obtaining the self-energy shifts for the VB states. We refer to these holes as the "few" holes in what follows. The interactions with the host valence electrons (and to a minor extent with the core electrons) have the effect that all potentials are screened by the background dielectric constant ϵ_∞. That the dopant system, in this way, can be decoupled from the host is because the regions in the ωq plane giving contributions to the interaction energies are separated and not overlapping for the two systems. For the dopant system the region is limited to small ω ($\lesssim 4E_F/\hbar$) and small q ($\lesssim 2k_F$) while for the host system, because of the band gap, it is limited to large ω ($\gtrsim E_g/\hbar$). This is discussed in more detail in Ref. 4.

The interaction energies are due to particle correlations in the system. Each electron in the CB is effectively surrounded by a depletion in the electron density, a so-called exchange and correlation hole. This effect is partly caused by the Fermi statistics and partly by the electrostatic repulsion between the electrons, and results in a negative energy contribution. In a similar way, each of the few holes in the VB is surrounded by an enhancement in the CB electron density caused by the electrostatic attraction between the hole and the electrons. This too produces a negative energy contribution. The ions are surrounded by enhancements in the electron density for exactly the same reason as are the holes. Furthermore, they are surrounded by a reduction of the hole density. The ion positions are fixed. The ions are assumed to be rigidly connected to the crystal and cannot move. The change in particle density near the ions comes from the adjustment of the particle motion to the ionic potentials. These correlations with the ions further reduce the energy.

In a polar semiconductor there is a slight displacement of the charge from one of the atom types to the other. This means that the host atoms are weakly charged. A moving, charged particle causes a displacement polarization, which follows the particle in the crystal. An electron attracts the positively charged atoms and repels the negatively charged ones. Effectively the electron is surrounded by a charge cloud, of positive charge. In the same way, a VB hole is surrounded by a charge cloud of negative charge. The particle and its charge cloud can be treated as a new particle, a polaron. Also, the ions polarize the crystal. All these correlations between the charged particles and the host atoms cause negative energy contributions. We mentioned, above, that we could decouple the dopant system from the host valence electrons, as the processes describing the interactions occur in nonoverlapping regions in the ωq plane. The polarization of the host atoms involves processes where valence electrons are excited across the band gap. The displacement polarization, on the other hand, involves much lower energies, viz., of the order of tens of meV's. The decoupling cannot be utilized here except, possibly, at the low doping limit. In the so-called ϵ_0 approximation one replaces ϵ_∞ by ϵ_0 for the screening of all potentials. This is done to replace the polar coupling. This approximation becomes increasingly good towards lower doping levels where E_F is smaller than the optical phonon energies. It is quite correct if E_F is much smaller than the phonon energies, in which case a decoupling can be performed, and the combined effect from the interactions with the valence electrons and the phonons is that all potentials are screened by the static dielectric constant, ϵ_0. For most densities this decoupling cannot be done and the two contributions interfere. Both have to be calculated as one entity.

We will now derive all the energy contributions mentioned above, and we start by discussing the Hamiltonian for the dopant system. It can be written as

$$H = H_e + H_{\text{ion}} + H_{\text{ph}} + H_{e\text{-ion}} + H_{e\text{-ph}} + H_{\text{ion-ph}} \, . \quad (2.1)$$

We have here divided our system into the three subsystems: carriers, ions, and phonons. The three first terms in the Hamiltonian describe these subsystems and the remaining three terms represent the interactions between them. We have collected the CB electrons and the few VB holes into one subsystem and could also have included the ions in this. However, because of the absence of a kinetic energy term for the ions the ion-phonon interaction can be taken care of in a particularly simple way and it is hence better to keep the ions in a separate subsystem.

The six terms in Eq. (2.1) can be expressed as

$$H_e = \sum_{j,\mathbf{k},\sigma} \frac{\hbar^2 k^2}{2m_j} \hat{n}_{j,\mathbf{k},\sigma} + \frac{1}{2\Omega} \sum_{\mathbf{q}} \frac{v(q)}{\epsilon_\infty} [\rho_{\mathbf{q}}^\dagger \rho_{\mathbf{q}} - (N + N_h)] \, ,$$

$$(2.2)$$

$$H_{\text{ion}} = H_{\text{ion-ion}} = \frac{1}{2\Omega} \sum_{\mathbf{q}} \frac{v(q)}{\epsilon_\infty} (\rho_{\text{ion},\mathbf{q}}^\dagger \rho_{\text{ion},\mathbf{q}} - N) \, , \quad (2.3)$$

$$H_{\text{ph}} = \sum_{\mathbf{q}} \hbar \omega_{\mathbf{q}} C_{\mathbf{q}}^\dagger C_{\mathbf{q}} \, , \quad (2.4)$$

$$H_{e\text{-ion}} = \frac{1}{\Omega} \sum_q \frac{v(q)}{\epsilon_\infty} \rho_q^\dagger \rho_{\text{ion},q} \ , \tag{2.5}$$

$$H_{e\text{-ph}} = -\frac{1}{\sqrt{\Omega}} \sum_q g(q)(C_q + C_{-q}^\dagger)\rho_q \ , \tag{2.6}$$

and

$$H_{\text{ion-ph}} = -\frac{1}{\sqrt{\Omega}} \sum_q g(q)(C_q + C_{-q}^\dagger)\rho_{\text{ion},q} \ . \tag{2.7}$$

Equation (2.2) represents the subsystem consisting of the donor electrons and the few VB holes. The first term gives the kinetic energy. The index j runs over the two types of carriers and takes on the values e or h, denoting electrons and holes, respectively. The masses m_e and m_h are the host CB and VB masses, respectively. We have here assumed isotropic energy dispersions near the band edges. The operator $\hat{n}$ is the particle number operator. The second term represents the interactions within this subsystem. The constants N and N_h are the number of donor electrons (which is equal to the number of donor ions) and the number of few VB holes, respectively. Ω denotes the total volume of the system and $v(q)$ is the Fourier transform of the Coulomb potential. The density operators are here defined as

$$\rho_q = \sum_{j,k,\sigma} Z_j a_{j,k,\sigma}^\dagger a_{j,k+q,\sigma} \tag{2.8}$$

and

$$\rho_q^\dagger = \sum_{j,k,\sigma} Z_j a_{j,k+q,\sigma}^\dagger a_{j,k,\sigma} \ , \tag{2.9}$$

where the Z_j is the charge of particle type j, in units of e (e being a positive number). Here Z_e equals -1 and Z_h equals $+1$. The operators $a^\dagger$ and a are creation and destruction operators for the particles, respectively.

Equation (2.3) gives the Hamiltonian for the ion subsystem. It is similar to that in Eq. (2.2), but does not contain a kinetic energy term. Besides, the density operators for the ions are not q numbers, but c numbers. They are defined as

$$\rho_{\text{ion},q} = Z_{\text{ion}} \sum_{j=1}^N e^{-i q \cdot R_j} \tag{2.10}$$

and

$$\rho_{\text{ion},q}^\dagger = \rho_{\text{ion},q}^* = Z_{\text{ion}} \sum_{j=1}^N e^{i q \cdot R_j} \ , \tag{2.11}$$

where R_j is the position vector of ion number j. Z_{ion} is the charge of each ion and is here equal to $+1$.

The Hamiltonian for the longitudinal, optical phonons is given in Eq. (2.4). We will later neglect the phonon dispersion and replace the phonon frequencies ω_q by a constant, ω_L. The operators $C^\dagger$ and C are creation and destruction operators for the phonons, respectively.

The interaction between the particles and ions is given in Eq. (2.5), where we have approximated the particle-ion interaction potential by a pure Coulomb potential.

The interactions with the phonons are shown in Eqs. (2.6) and (2.7), where the coupling constant $g(q)$ is given by

$$g(q) = \left[\frac{v(q)}{2} \frac{\epsilon_0 - \epsilon_\infty}{\epsilon_\infty \epsilon_0} \hbar \omega_L \right]^{1/2} . \tag{2.12}$$

The fact that the Hamiltonian does not contain a kinetic energy term for the ions makes it possible to take care of the ion-phonon interaction by the following unitary transformation:[20]

$$U = \exp\left[\sum_q f(q)(C_q - C_{-q}^\dagger) \right], \quad f(q) = \frac{g(q)}{\hbar \omega_q} \rho_{\text{ion},q}^\dagger \ . \tag{2.13}$$

This transformation modifies only terms containing phonon operators. The phonon operators are transformed in the following way:

$$U C_q U^\dagger = C_q + f(q) \ ,$$
$$U C_q^\dagger U^\dagger = C_q^\dagger + f^\dagger(q) \ , \tag{2.14}$$

and the transformed Hamiltonian takes on the form

$$\begin{aligned}
H = &\sum_{j,k,\sigma} \frac{\hbar^2 k^2}{2m_j} \hat{n}_{j,k,\sigma} + \frac{1}{2\Omega} \sum_q \frac{v(q)}{\epsilon_\infty} [\rho_q^\dagger \rho_q - (N + N_h)] \\
&- \frac{1}{\sqrt{\Omega}} \sum_q g(q)(C_q^\dagger + C_{-q})\rho_q + \frac{1}{\Omega} \sum_q \frac{v(q)}{\epsilon_0} \rho_q^\dagger \rho_{\text{ion},q} \\
&+ \frac{1}{2\Omega} \sum_q \frac{v(q)}{\epsilon_0} (\rho_{\text{ion},q}^\dagger \rho_{\text{ion},q} - N) + \sum_q \hbar \omega_q C_q^\dagger C_q \\
&+ \frac{1}{2\Omega} \sum_q v(q) \left[\frac{1}{\epsilon_0} - \frac{1}{\epsilon_\infty} \right] N \ . \tag{2.15}
\end{aligned}$$

The transformation has had the following effects on the Hamiltonian. The impurity-phonon interaction has been eliminated, and all impurity potentials are now screened by ϵ_0 instead of by ϵ_∞ as they were before the transformation. Apart from these changes a new term, the last in Eq. (2.15), has appeared. This term gives the change in the infinite self-energy of the ions from the interaction with the phonons. It has no effect on the energy of the electronic states, which is our concern here. This term also comes from the adjustment of the host atoms to the ion potentials. It is interesting to note that the phonon energies are not changed by this adjustment. In obtaining Eq. (2.15) we have made use of the identity

$$\frac{v(q)}{\epsilon_\infty} - \frac{2g(q)^2}{\hbar \omega_L} = \frac{v(q)}{\epsilon_0} \ . \tag{2.16}$$

When one calculates the energy of a system one always has to decide which energy reference to choose. Here we have chosen, as the reference energy for each particle, the energy for each respective particle at the band edge in the semiconductor in the absence of the polar coupling. Similarly for the ions, the reference energy has been chosen as the energy of a single ion present in the semiconductor in the absence of the polar coupling. Neglecting the last term in Eq. (2.15), which we do henceforth, corresponds to defining the reference energy for the ions in the presence of the polar coupling. In the next section when we treat the BGN we are interested in the shifts of the electronic states due to the doping. As

the polar coupling is present also in the absence of the doping, the reference energy for the particles should be defined in the presence of the polar coupling. We return to this point in the next section.

The term containing particle-phonon interactions in Eq. (2.15) describes processes where a particle is scattered and a phonon is absorbed or emitted. It leads to an extra particle-particle interaction

$$V_{ph}(\mathbf{q},\omega) = g(q)^2 D^0(\mathbf{q},\omega) , \qquad (2.17)$$

where $D^0(\mathbf{q},\omega)$ is the phonon propagator

$$\hbar D^0(\mathbf{q},\omega) = \frac{1}{\omega - \omega_q + i\eta} - \frac{1}{\omega + \omega_q - i\eta} . \qquad (2.18)$$

The Coulomb interaction together with the interaction via the phonons results in an effective particle-particle interaction

$$V(\mathbf{q},\omega) = \frac{v(q)}{\epsilon_\infty} + V_{ph}(\mathbf{q},\omega) = \frac{v(q)}{\epsilon_\infty} + \frac{g^2(q)2\omega_q}{\hbar(\omega^2 - \omega_q^2 + i\eta)} . \qquad (2.19)$$

In our approximation $\omega_q \equiv \omega_L$ and thus

$$V(\mathbf{q},\omega) = \frac{v(q)}{\epsilon_L(\omega)} , \qquad (2.20)$$

where the lattice dielectric function ϵ_L is given by

$$\epsilon_L(\omega) = \frac{\epsilon_\infty \epsilon_0 (\omega^2 - \omega_L^2)}{\epsilon_0 \omega^2 - \epsilon_\infty \omega_L^2 + i\eta} , \qquad (2.21)$$

and represents the combined screening from the background (from valence and core electrons) and from the optical phonons. For frequencies much higher than the optical phonon frequencies, the phonons give no contributions to the screening and ϵ_L tends to the background screening constant ϵ_∞. In the zero-frequency limit ϵ_L tends to ϵ_0.

For the following equations it is convenient to define the following dielectric function:

$$\epsilon(\mathbf{q},\omega) = 1 - \frac{v(q)}{\epsilon_L(\omega)}\chi^0(\mathbf{q},\omega) = 1 + \alpha(\mathbf{q},\omega)/\epsilon_L(\omega) , \qquad (2.22)$$

where $\chi^0(q,\omega)$ is the sum of the polarizabilities from the particles (donor electrons and few VB holes). This is the same dielectric function as used in our earlier work,[1-5] except that the background dielectric function is now replaced by the frequency-dependent lattice dielectric function. The total dielectric function is the product of the two in Eqs. (2.21) and (2.22). This total dielectric function can be extended to include contributions from other types of coupling like, e.g., the piezoelectric coupling to the acoustical phonons. This is described in detail in Ref. 21 and we refer the reader to this reference for more material on the polaron problem.

We found in Eq. (2.15) that the ion contribution to the energy is the same as in a nonpolar semiconductor[4] except that the background screening-constant is replaced by ϵ_0, i.e.,

$$E_{ion} = \frac{n}{2} \sum_q{}' \frac{v(q)}{\epsilon_0}[\epsilon^{-1}(\mathbf{q},0) - 1] . \qquad (2.23)$$

The exchange and correlation energy in Eq. (2.3) of Ref. 4 is now replaced by a contribution also containing the effects of the interactions with the phonons. We denote it by E_{int}, and it can be expressed as

$$E_{int} = -\int_0^1 \frac{d\lambda}{\lambda} \left[\sum_q{}' \left[\int_0^\infty d\omega \frac{\hbar}{2\pi i}[\epsilon_\lambda^{-1}(\mathbf{q},\omega) - 1] \right. \right.$$
$$\left. \left. + \frac{\lambda(N + N_h)v(q)}{2\epsilon_\infty \Omega} \right] \right] . \qquad (2.24)$$

The function ϵ_λ is the dielectric function of Eq. (2.22) with $v(q)$ replaced by $\lambda v(q)$, where λ is the coupling constant. The only modification caused by the phonons is that the background screening constant is now replaced by the lattice dielectric function in the first term of the integrand in Eq. (2.24).

We now introduce the dielectric function $\epsilon_0(\mathbf{q},\omega)$ and rewrite the last term of Eq. (2.24) the same way as we did in Ref. 4, i.e.,

$$E_{int} = -\int_0^1 \frac{d\lambda}{\lambda} \sum_q{}' \int_0^\infty d\omega \frac{\hbar}{2\pi i}\{[\epsilon_\lambda^{-1}(\mathbf{q},\omega) - 1]$$
$$- [\epsilon_{0,\lambda}^{-1}(\mathbf{q},\omega) - 1]\} . \qquad (2.25)$$

The dielectric function $\epsilon_0(\mathbf{q},\omega)$ is the dielectric function one would have for the system if the particles were occupying the same states as now, but were noninteracting and were not obeying the Pauli principle. With the last statement we mean that they are allowed to scatter into states already occupied. As it stands now it is valid in the absence of the polar coupling. Written in this way the physics becomes more transparent. In general, the energy in a specific approximation can be written as

$$E_{approx} = -\int_0^1 \frac{d\lambda}{\lambda} \sum_q{}' \int_0^\infty d\omega \frac{\hbar}{2\pi i}[\epsilon_{approx,\lambda}(\mathbf{q},\omega) - 1] , \qquad (2.26)$$

where $\epsilon_{approx}(\mathbf{q},\omega)$ is the dielectric function in that particular approximation. If $\epsilon_0(\mathbf{q},\omega)$ is used, one gets the energy of our reference system, i.e., the energy in the absence of both interactions and fulfillment of the Pauli principle. In the next higher order of approximation, the Hartree-Fock approximation, the particles are still noninteracting but obey the Pauli principle. The improvement in energy by using the Hartree-Fock dielectric function is the exchange energy, and is due to the fact that the particles are not allowed to scatter into already occupied states. The further improvement in energy by using, e.g., the RPA dielectric function (as we do here) is the correlation energy, and it arises from the fact that the particles are now treated as interacting, i.e., they contribute screening in the scattering processes. The final improvement in Eq. (2.25) is that the phonons are also contributing to this screening.

The integration of Eq. (2.25) over the coupling constant gives

$$E_{\text{int}} = -\sum_{\mathbf{q}}' \int_0^\infty d\omega \frac{\hbar}{2\pi i} \{ -\ln[\epsilon(\mathbf{q},\omega)] - [\epsilon_0^{-1}(\mathbf{q},\omega) - 1] \} \ . \tag{2.27}$$

The last expression within the curly brackets can be written as

$$\epsilon_0^{-1}(\mathbf{q},\omega) - 1 = -\frac{1}{\epsilon_\infty} \sum_{j,\mathbf{k},\sigma} \alpha_{\mathbf{k},\sigma}^j = -\frac{v(q)}{\epsilon_\infty \hbar \Omega} \sum_{j,\mathbf{k},\sigma} n_{\mathbf{k},\sigma}^j \left[\frac{1}{\omega + \omega^j(\mathbf{k},\mathbf{q}) + i\eta} - \frac{1}{\omega - \omega^j(\mathbf{k},\mathbf{q}) + i\eta} \right] , \tag{2.28}$$

where

$$\omega^j(\mathbf{k},\mathbf{q}) = (e_{\mathbf{k}+\mathbf{q}}^j - e_{\mathbf{k}}^j)/\hbar \ , \tag{2.29}$$

i.e., $\hbar\omega^j(\mathbf{k},\mathbf{q})$ is the change in kinetic energy for a particle of type j in going from state $\mathbf{k}$ to state $\mathbf{k}+\mathbf{q}$.

In our final expression for E_{int} we regroup the terms in the following way:

$$E_{\text{int}} = -\sum_{\mathbf{q}}' \int_0^\infty d\omega \frac{\hbar}{2\pi i} \left[\left[-\ln[\epsilon(\mathbf{q},\omega)] + \frac{1}{\epsilon_L(\omega)} \alpha_0(\mathbf{q},\omega) \right] - \alpha_0(\mathbf{q},\omega) \left[\frac{1}{\epsilon_L(\omega)} - \frac{1}{\epsilon_\infty} \right] \right] , \tag{2.30}$$

i.e., we introduce the phonon screening in the first step of the improvement of the energy over that of our reference system. The last term gives the change in the infinite self-energy of the particles due to the polar coupling.

We now define the quasiparticle energy E_p for a particle in state p according to the following definition:

$$E_p = \frac{\delta E}{\delta n_p} = e_p + \Sigma_p \ , \tag{2.31}$$

where e_p and Σ_p are the kinetic energy and self-energy, respectively, for a particle in state p.

The self-energy contribution from E_{int} is

$$\Sigma_p^{\text{int},j} = -\frac{1}{\Omega} \sum_{\mathbf{q}}' v(q) \int_{-\infty}^\infty d\omega \frac{1}{2\pi i} \left\{ \frac{1}{\epsilon_L(\omega)} \left[G_0^j(\mathbf{p}+\mathbf{q}, e_p^j/\hbar + \omega)/\epsilon(\mathbf{q},\omega) \right. \right.$$

$$\left. + \frac{1}{2} \left[\frac{1}{\omega + \omega^j(\mathbf{p},\mathbf{q}) - i\eta} - \frac{1}{\omega - \omega^j(\mathbf{p},\mathbf{q}) + i\eta} \right] \right]$$

$$\left. - \left[\frac{1}{\epsilon_L(\omega)} - \frac{1}{\epsilon_\infty} \right] \frac{1}{2} \left[\frac{1}{\omega + \omega^j(\mathbf{p},\mathbf{q}) - i\eta} - \frac{1}{\omega - \omega^j(\mathbf{p},\mathbf{q}) + i\eta} \right] \right\} . \tag{2.32}$$

Let us concentrate on the first term. It is exactly the same as that for a nonpolar semiconductor [see Eq. (2.19) of Ref. 4] apart from the fact that the background screening constant is replaced by the lattice dielectric function. We deform the integration path the same way as in the case of a nonpolar semiconductor (see Fig. 2 of Ref. 2). The lattice dielectric function now adds new poles to the problem, but they all fall outside the integration path and cause no problems. We obtain one contribution from the integration along the imaginary ω axes, which we call Σ_{line}, and each of the subterms gives a residue contribution as their poles fall inside the integration path. We denote the contribution from the second subterm by Σ_{res1}, and that from the first subterm by Σ_{res2}.

The contribution from the second term in Eq. (2.32) is the polaron self-energy and it can either be calculated directly or it can be obtained after the deformation of the integration path in the same way as in the treatment of the first term. If the last method is used it gives rise to two contributions, a line part and a residue part. We do not make that separation here, in the presentation of the results. We name the polaron contribution Σ_{pol}.

The remaining self-energy contribution, i.e., the contribution from the interaction with the ions, is the same as

for a nonpolar semiconductor, except that the background screening-constant is here replaced by ϵ_0, i.e.,

$$\Sigma_p^{\text{ion},j} = \frac{n}{\hbar\Omega} \sum_{\mathbf{q}}' \left[\frac{v(q)}{\epsilon_0 \epsilon(\mathbf{q},0)} \right]^2 G_0^j(\mathbf{p}+\mathbf{q}, e_p^j/\hbar) \ . \tag{2.33}$$

We have now reached a point where we can present numerical results for the self-energy shifts in CdS. For ϵ_0, ϵ_∞, $\hbar\omega_L$, m_e, and m_h, we used the values 8.58, 5.26, 36.8 meV, 0.184, and 0.8389, respectively. These data have been obtained from Ref. 22. The masses are in units of the electron mass and are the bare masses, obtained from the experimental polaron masses by eliminating the polaron mass enhancement. The VB mass is unfortunately anisotropic. Here we have used the density-of-states effective mass and neglected the effects from the anisotropy, beyond that.

In Fig. 1 are shown all the different self-energy contributions, introduced above, in the CB of CdS for a donor density of 5×10^{18} cm^{-3}. The curves labeled $a-e$ represent Σ_{res1}, Σ_{res2}, Σ_{ion}, Σ_{pol} and Σ_{line}, respectively. The total self-energy shift, Σ_{tot}, and the quasiparticle energy, E_{tot}, are also shown. All the contributions are given as functions of the particle wave number in units of twice
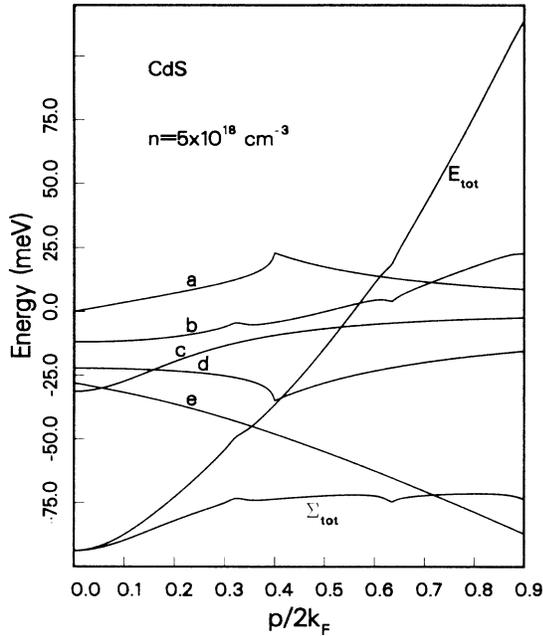
FIG. 1. Different contributions to the self-energy shift in the CB of n-type CdS as functions of wave number. The wave numbers are given in units of twice the Fermi wave number. The curves denoted by a through e represent Σ_{res1}, Σ_{res2}, Σ_{imp}, Σ_{pol}, and Σ_{line}, respectively, all defined in the text. Also shown is the sum of their contributions, Σ_{tot}, and the total particle energy, E_{tot}.

electron, in which the electron can fall downwards in energy through the excitation of a phonon.

In the doped case Σ_{res1} has a structure that completely cancels the structure in Σ_{pol}. Curves a and d are both independent of doping concentration. In the doped case the corresponding structure should appear at an energy of $\hbar\omega_L$ above the Fermi level. This is also what is found. Σ_{res2}, curve b, has a structure at this position. It also has a new structure below the Fermi level. This structure appears at the energy where a new channel opens up for the deexcitation of a hole, in which the hole can fall upwards in energy through the excitation of a phonon. It appears, in other words, at an energy $\hbar\omega_L$ below the Fermi level. For very low doping levels where the Fermi energy is lower than the phonon energy this last structure disappears. There is one more thing to note in Fig. 1. At the right edge of the figure, curve b starts to curve downwards again. This is because there is one further structure just outside the figure, caused by the coupling to the donor-electron plasmons. At an energy equal to the plasmon energy above the Fermi level a new channel is opened up, in which the electron can decay via the emission of a plasmon. Finally, we note that the contribution from the ion interaction is increasing towards the bottom of the band. This causes a stretching of the DOS.

Figure 2 presents the same results as did Fig. 1, but now in the ϵ_0 approximation. In this approximation $\epsilon_L(\omega)$ and ϵ_∞ are both replaced by ϵ_0. The a and d contributions vanish, as do the structures in curve b caused by the phonons. Curve c is completely unchanged. It should be noted that the structure in curve b, which is
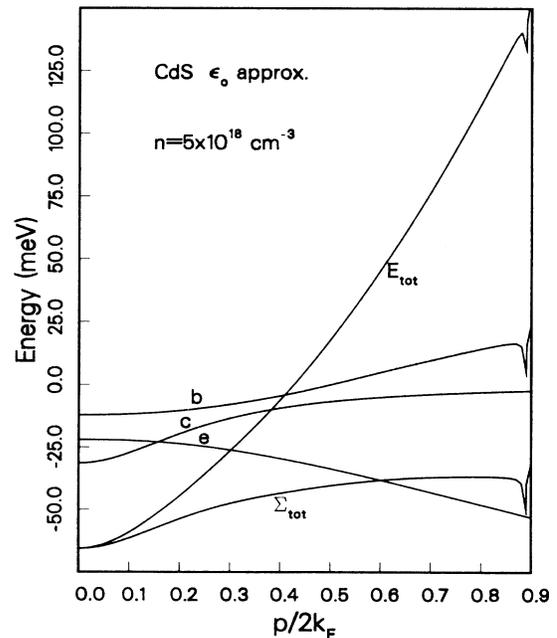
the Fermi wave number. We have limited ourselves to the real part of the contributions. For a single electron in the CB of the undoped crystal, the only self-energy contribution is Σ_{pol}, curve d in Fig. 1. Its value at the band edge is called the polaron ground-state energy and is given by

$$E_{pol} = e^2 \left[\frac{m\omega_L}{2\hbar} \right]^{1/2} \left[\frac{1}{\epsilon_0} - \frac{1}{\epsilon_\infty} \right] = -\alpha\hbar\omega_L , \quad (2.34)$$

where the parameter α is a measure of the strength of the polar coupling. As the value of α depends on the particle mass, it is not the same for the CB and VB. It takes on the values 0.61 and 1.30 for the CB and VB, respectively. Our treatment here corresponds to the Rayleigh-Schrödinger perturbation method for the polaron energy. As is discussed in Sec. 6.1 of Ref. 23, this method gives quite accurate results both in the weak- and intermediate-coupling regimes. The Tamm-Dancoff approximation gives much poorer results. Curve d is clearly curved downwards for small wave numbers, which means that the polaron effective mass is enhanced relative to the bare mass. In our treatment the mass enhancement is equal to $(1-\alpha/6)^{-1}$. Curve d displays a sharp structure for a certain wave number. This structure appears for a wave number where the electron is at an energy of exactly $\hbar\omega_L$ above the bottom of the CB. The structure appears at the point where a new deexcitation channel opens up for the

FIG. 2. The same as in Fig. 1, with unchanged notation, but now in the ϵ_0 approximation. In this approximation the contributions denoted by a and d in Fig. 1 vanish and are left out.

caused by the coupling to the plasmons, has moved down somewhat in energy, and is now inside the figure.

Figure 3 is the same as Fig. 1, but now for the VB. To be noted here is that only one structure, from the coupling to the phonons, appears in the doped case. Furthermore it remains in its unshifted position and is weakened due to the screening from the donor electrons. This is as it should be, as the VB is empty and no particles are blocking the states near the band edge. The structure due to the plasmons is also present here but is outside the figure. Due to the heavier mass in the VB this structure appears at a larger wave number. Of particular interest is the much stronger effect the ion interaction has in the VB compared to that in the CB. This difference is caused by the difference in mass. The larger the mass, the larger the contribution. The same effect was found for n-type doped GaAs in Ref. 3. On the other hand, if the screening particles are the ones with the heavy mass, the ion contribution is drastically reduced. This was found in p-type doped GaAs in Ref. 4.

For completeness we show, in Fig. 4, the results for the VB in the ϵ_0 approximation. Here also the a and d contributions vanish, as do the phonon structures in curve b.

The changes in the energy dispersions cause modifications in the DOS. The results are shown in Figs. 5 and 6 for the CB and VB, respectively.

In Fig. 5 the dashed, dotted, and solid curves represent the DOS in the absence of all interactions (and the absence of the polar coupling), in the ϵ_0 approximation, and in the full calculation, respectively. The DOS is given in units of the noninteracting value at the Fermi level. The energy unit is the Fermi energy E_F, and all curves have
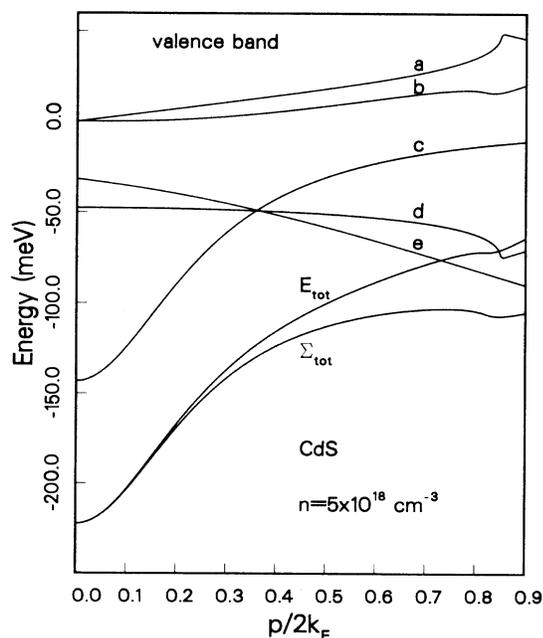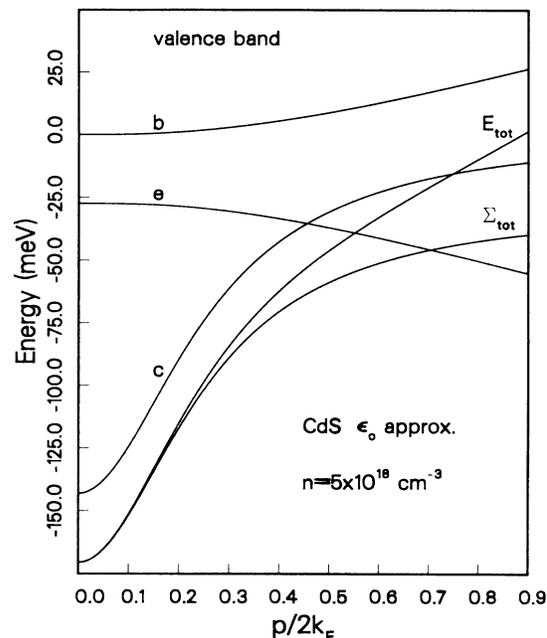

FIG. 4. The same as in Fig. 2 but now for the VB of n-type CdS.

been shifted in energy so that the Fermi levels coincide. The dashed, vertical line separates the occupied states from the unoccupied ones. Apart from the structures caused by the phonons, the dashed and solid curves are rather similar. The ion interaction causes the stretching of the DOS. The structure at the right edge of the figure
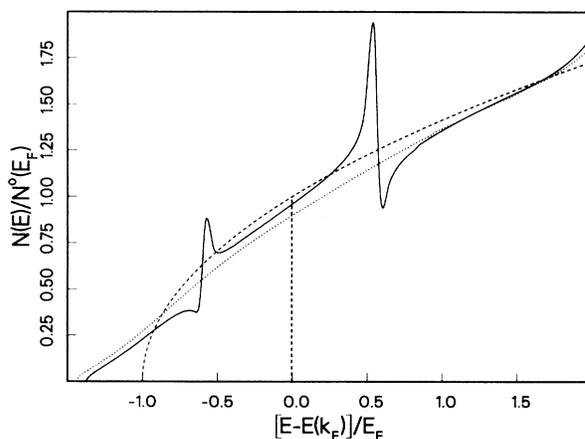

FIG. 3. The same as in Fig. 1 but now for the VB of n-type CdS.

FIG. 5. The CB DOS in n-type doped CdS for a donor density of 5×10^{18} cm^{-3}. The curves are given in units of the value of the DOS at the Fermi level in the unperturbed band, and the energy unit is the Fermi energy. All curves have been shifted in energy so that the Fermi levels coincide. The dashed curve denotes the DOS in the unperturbed band. The solid and dotted curves are the results from the full calculation and those from using the ϵ_0 approximation, respectively. The vertical, dashed line separates the occupied states from the unoccupied ones, in all three cases.

is due to the plasmon interactions and appears in both of the treatments. The results are valid for the same donor density as in Figs. 1–4. We have also obtained results for other densities but have not shown them here. For lower densities the structures move away from the Fermi level (if the energy unit is E_F, as here) and their sizes increase. For low enough density, the structure below the Fermi level moves outside the region of nonvanishing DOS, and disappears. For higher doping levels the structures move closer to the Fermi level and weaken. We believe that at least some traces of these structures should be observable in experiments. Here we have made some idealizations that make the structures particularly sharp. We have neglected the phonon dispersions. Any phonon dispersion would spread out the structures somewhat in energy. Furthermore, our treatment corresponds to having a delta-function-like optical absorption at the transverse optical phonon energy. In the real situation there are always some broadening effects. A more realistic treatment would reduce the structures further, but not completely wipe them out.

Finally, in Fig. 6 is displayed the DOS in the VB. The notation is the same as in Fig. 5, but here the dashed, vertical line indicates the position of the state with CB Fermi wave vector. The structure from the coupling to the phonons is very strong here but outside the figure. We have chosen to show, instead of the structure, the extremely strong stretching of the DOS. The energy spread for the states with wave numbers less than the CB Fermi wave number has increased by a factor of between 9 and 10. Even though our obtained DOS has a very strong tailing in the minority-carrier band the band is still bounded. In this connection it is interesting to mention a
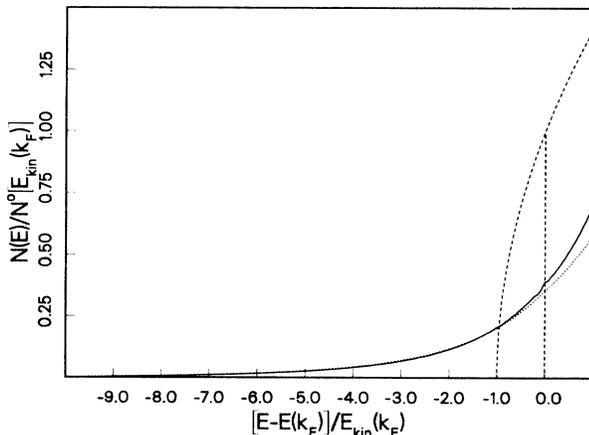

FIG. 6. Same as in Fig. 5, but for the VB. There are, however, some differences in the notation. The curves are normalized to the value of the DOS in the unperturbed VB at the energy for a state with its wave number equal to the Fermi wave number in the CB. The energy unit is chosen as the kinetic energy of such a state. All curves have been shifted so that the energies for states with Fermi wave number coincide. The vertical, dashed line separates, in this case, states with wave numbers smaller than k_F from those with wave numbers larger than k_F.

very recent theoretical work[24] on the minority-carrier states in which one finds that the energy of the states at the band edge is unbounded. The strong tailing obtained here means that one has to be very careful when interpreting experimental results. Another point worth discussing is that the DOS we have presented here is the quasiparticle DOS defined without involving lifetime effects. Depending on the experiments other definitions of the DOS might be more appropriate. Here we have not treated the imaginary parts of the self-energy shifts, causing the lifetime effects. These are important for the interpretation of most experiments. As we do not perform comparisons with experiments in this work we have chosen to leave the imaginary parts out. We just briefly mention that the polar coupling gives rise to imaginary parts in the regions outside the two structures in the DOS, but not in the region in between.

This completes the discussion of the self-energy shifts and the DOS. In the next section we will discuss the BGN, and in particular we will study how good the ϵ_0 approximation is in this context.

III. BAND-GAP NARROWING

In this section we give a brief presentation of our results for the BGN. We present results both from the full calculation and from a calculation in the ϵ_0 approximation. We introduce three different band gaps in the doped semiconductor. The band gap E_g in the undoped CdS is the gap in energy between the bottom of the CB and the top of the uppermost VB at the Γ point. It is the lower threshold both for luminescence and optical absorption. In the doped case it is more complicated and different band gaps enter in different experiments. We name the band gap defined as the distance in energy between the band edges $E_{g,1}$. This is still the lower threshold for luminescence. In an absorption experiment, on the other hand, the lower threshold or optical band gap is different. It is the difference in energy between states in the CB and the VB, both of which have the Fermi momentum. We denote this band gap with $E_{g,3}$. If the bands were shifted rigidly, from the interactions, the difference between these two band gaps would be the sum of the kinetic energies for the states with Fermi momentum in the CB and in the VB. One can define a third band gap, $\Delta E_{g,2}$, as the energy distance between the VB edge and the Fermi level in the CB. This would correspond to the lower threshold for absorption if nonvertical transitions were allowed. It is also the upper threshold for luminescence from thermalized holes if nonvertical processes are allowed. We define the BGN's from the introduced band gaps in the following way:

$$\Delta E_{g,i} = E_{g,i} - E_g, \quad i = 1, 2, 3 . \tag{3.1}$$

The various BGN's are easily obtained from the self-energy shifts derived in the preceding section. However, one must keep in mind that E_g already contains a self-energy shift, viz., the shift caused by the coupling to the phonons. This means that the polaron ground-state energy has to be subtracted from the self-energy shifts.

Thus the BGN's are obtained as

$$\Delta E_{g,1} = \Sigma^e(0) + \Sigma^h(0) - E_{pol}^e - E_{pol}^h \ , \qquad (3.2)$$

$$\Delta E_{g,2} = e^e(k_F) + \Sigma^e(k_F) + \Sigma^h(0) - E_{pol}^e - E_{pol}^h \ , \qquad (3.3)$$

and

$$\Delta E_{g,3} = e^e(k_F) + \Sigma^e(k_F) + e^h(k_F)$$
$$+ \Sigma^h(k_F) - E_{pol}^e - E_{pol}^h \ . \qquad (3.4)$$

Our BGN's obtained for CdS are shown as functions of the donor concentration in Fig. 7. The solid curves are the results from the full calculation and the dashed ones are from the ϵ_0 approximation. As can be seen, the ϵ_0 approximation produces BGN's which are too small, but the discrepancies are very small in the low-density region.

The structure in the energy dispersion for a VB hole due to the coupling to the optical phonons appears at a fixed particle wave number. As the doping level increases, the wave number for the hole involved in the definition of $\Delta E_{g,3}$ increases. One might expect to find some structure in $\Delta E_{g,3}$ when the hole wave number passes through the region with the structure in the energy dispersion. This occurs at the density 2.5×10^{19} cm^{-3}. A weak anomaly is found in this region in Fig. 7, but the effect is much too weak to have an experimental significance.

Some words of caution are in place here. In the derivation we approximated the VB with an isotropic band characterized by the density-of-states effective mass. The purpose was not to produce results to be compared with experiments, but to study the effects from the polar cou-

pling. In CdS the VB is anisotropic, which means that $\Delta E_{g,3}$ is different for particles moving in different directions in the crystal. This complicates the comparisons with experiments. A different complication appears if one wants to determine $\Delta E_{g,1}$ from the lower threshold of the luminescence from thermalized VB holes. The strong tailing of the DOS that was found in the preceding section might make it difficult to separate the effects on the peak shapes of this tailing from the lifetime effects. This might make it difficult to extract an experimental $\Delta E_{g,1}$.

IV. SUMMARY AND CONCLUSION

We have presented a derivation of the self-energy shifts of the electronic states in a heavily doped, polar semiconductor. We applied the results to n-type doped CdS and showed numerical results for the energy shifts in both the conduction and valence bands. We also presented the density of states obtained for both bands and found interesting new structures. In the conduction band a structure appears at an energy above the Fermi level equal to that of a longitudinal optical phonon. This is at the position at which a new channel opens up for the decay of an excited electron, viz., via the emission of a longitudinal, optical phonon. A similar structure appears below the Fermi level. In this case it is at the energy where, in an analogous way, a new decay channel is opened up for an excited hole. Below this energy a hole can fall upwards via the excitation of a phonon. In the valence band a similar structure (and of the same origin) is present at the phonon energy below the band edge. The structures were found to be strong enough to be detected experimentally. We also found strong tailing of the density of states in the valence band.

The purpose of this work was to study the effects from the polar coupling and the modifications, due to this, of the effects from the doping. The deformation of the density of states is one effect from the interactions. A second effect is an overall shift of the bands, leading to band-gap narrowing. We derived the band-gap narrowing as a function of donor density and compared the results to those obtained in the ϵ_0 approximation. In this approximation one can use the much simpler formalism derived for a nonpolar semiconductor. We found important deviations. The ϵ_0 approximation always gave band-gap narrowings which are too small. The discrepancies were small in the low-density limit but increased with density and were severe at the high-density limit. From this we conclude that it is important to take the full effects of the polar coupling into account if one wants to make detailed comparisons with experiments, at least for higher doping levels.

The numerical results for n-type CdS were presented for the purpose of numerical demonstration of the effects from the polar coupling. In an experimental comparison, the anisotropic valence-band dispersion in CdS would have to be taken into account more accurately than we have done here. Furthermore, we have completely neglected the piezoelectric coupling to the acoustical phonons. CdS is known for having strong piezoelectric cou-
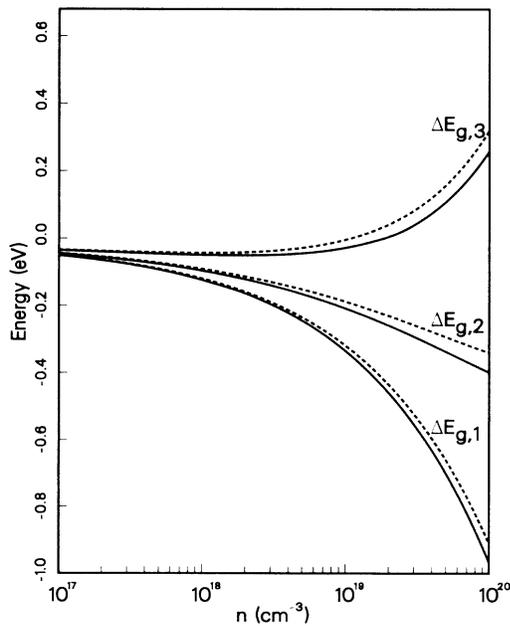

FIG. 7. The three BGN's, defined in the text, as functions of the donor concentration in CdS. The solid curves are the results of the full calculation and the dashed curves are those obtained by using the ϵ_0 approximation.

pling.[21] It is not obvious whether this effect is important for the band-gap narrowing. This should be investigated.

To our knowledge, there is only one published comparison between theory and experiment for the band-gap narrowing in CdS.[25] In this reference, the shift of the upper fluorescence threshold with doping level is studied. A good agreement between the theory and the experiment is found, but the interpretation relies on the dominance of nonvertical transitions. This is in contradiction to our results for GaAs, in Refs. 3–5, where we found that only vertical transitions were of importance.

ACKNOWLEDGMENTS

*Permanent address: Department of Physics and Measurement Technology, University of Linköping, S-581 83 Linköping, Sweden.

[1]B. E. Sernelius and K.-F. Berggren, in *Recent Developments in Condensed Matter Physics*, edited by J. T. Devreese, L. F. Lemmens, V. E. Doren, and J. Van Royen (Plenum, New York, 1981), Vol. 3.

[2]K.-F. Berggren and B. E. Sernelius, Phys. Rev. B **24**, 1971 (1981).

[3]Bo E. Sernelius, Phys. Rev. B **33**, 8582 (1986).

[4]Bo E. Sernelius, Phys. Rev. B **34**, 5610 (1986).

[5]Bo E. Sernelius, Phys. Rev. B **34**, 8696 (1986).

[6]R. A. Abram, A. J. Rees, and B. L. Wilson, Adv. Phys. **27**, 799 (1978), and references therein.

[7]G. D. Mahan, J. Appl. Phys. **15**, 2634 (1980).

[8]P. E. Schmid, Phys. Rev. B **23**, 5531 (1981).

[9]P. E. Schmid, M. L. W. Thewalt, and W. P. Dumke, Solid State Commun. **38**, 1091 (1981).

[10]W. P. Dumke, Appl. Phys. Lett. **42**, 196 (1983).

[11]W. P. Dumke, J. Appl. Phys. **54**, 3200 (1983).

[12]L. Viña and M. Cardona, Phys. Rev. B **29**, 6739 (1984).

[13]J. Wagner, Phys. Rev. B **32**, 1323 (1985).

[14]L. Viña and M. Cardona, Phys. Rev. B **34**, 2586 (1986).

[15]W. Bardyszewski and D. Yevick, Phys. Rev. B **35**, 619 (1987).

[16]I. Hamberg and C. G. Granqvist, Appl. Phys. Lett. **44**, 721 (1984).

[17]I. Hamberg, C. G. Granqvist, K.-F. Berggren, B. E. Sernelius, and L. Engström, Vacuum **35**, 207 (1985).

[18]I. Hamberg, C. G. Granqvist, K.-F. Berggren, B. E. Sernelius, L. Engström, Phys. Rev. B **30**, 3240 (1984).

[19]Bo. E. Sernelius, Phys. Rev. B **36**, 1080 (1987).

[20]Bo E. Sernelius, Ph.D. thesis (Linköping University, Sweden, 1978).

[21]G. D. Mahan, in *Polarons in Ionic Crystals and Polar Semiconductors*, edited by J. T. Devreese (North-Holland, Amsterdam and London, 1972).

[22]*Physics of II-VI and I-VII Compounds, Semimagnetic Semiconductors*, Vol. 17 of *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung (Springer, Heidelberg, 1982), Part b.

[23]G. D. Mahan, *Many-Particle Physics* (Plenum, New York, 1981).

[24]G. D. Mahan (unpublished).

[25]S. M. Girvin, Phys. Rev. B **17**, 1877 (1978).