PHYSICAL REVIEW B VOLUME 60, NUMBER 11 15 SEPTEMBER 1999-I

Electronic structure of $ZnGeP_2$: A detailed study of the band structure near the fundamental gap and its associated parameters

Sukit Limpijumnong, Walter R. L. Lambrecht, and Benjamin Segall
Department of Physics, Case Western Reserve University, Cleveland, Ohio 44106-7079

(Received 8 April 1999; revised manuscript received 19 May 1999)

Full-potential linear muffin-tin orbital band-structure calculations of $ZnGeP_2$ in the local-density approximation (LDA) indicate several close-lying conduction-band minima at different points in the Brillouin zone. Quasiparticle results available for the zinc-blende ''parent'' compound GaP are used to estimate corrections beyond the LDA. Even including these corrections, $ZnGeP_2$ appears to be truly indirect rather than pseudodirect, as has been suggested in the literature. The experimental evidence is reviewed. The standard assignment of features in optical data pertaining to the absorption edge is questioned and a tentative new interpretation is presented in light of our multiple conduction-band minima model. Related band-structure parameters, such as band-gap deformation potentials, effective masses at each of the minima, and the valence-band effective Hamiltonian are determined. [S0163-1829(99)01835-4]

I. INTRODUCTION

$ZnGeP_2$ is a ternary semiconductor compound with chalcopyrite (CKP) structure which has so far been studied primarily for its potential as nonlinear optical material for frequency conversion applications.[1,2] Recently, significant progress has been made in optimizing its crystal growth[3] and in identifying and reducing its defects,[4] drawing renewed attention to it and its related materials. In spite of many years of study,[5–18] there is still not a consensus on the nature of its band structure. Early work was based on comparison to the zinc-blende (ZB) parent III-V compound GaP and suggested either a so-called pseudodirect band gap or an indirect band gap. The conduction-band minimum of GaP at one of the X points of the ZB Brillouin zone is folded onto the Γ point of the chalcopyrite superstructure, while the other ones fall at T. All of the X_1 derived states [that is, the two T_1+T_2 (time-reversal degenerate) states and the Γ_3 state using the symmetry notation of Ref. 19] are thus candidates for the conduction-band minimum in $ZnGeP_2$. Because the c/a ratio is slightly lower than 2, implying that there is a compressive uniaxial strain effect on the X_{ZB} state along the z axis, it was argued that the state folded onto Γ should fall below the ones that fold onto T.[12] On the other hand, interactions with other folded bands suggest that the T_1+T_2 state would be lower.[7] Experimentally, one observes a weak and slowly increasing absorption edge, which is consistent with either an indirect transition or a very weak direct transition.[10] Based on the resulting splittings of the low-lying conduction-band states and the valence-band maximum at Γ, a model has also been proposed to explain the photoluminescence data.[20–22]

The first local-density-functional calculations were performed by Jaffe and Zunger using a mixed basis-set method and obtained a pseudodirect Γ_3 gap, followed by minima at N and T. The empirical pseudopotential calculations of Varea de Alvarez and Cohen[15] found a direct gap Γ_1 closely followed by an indirect gap at N. More recently, more or less simultaneously with our present work, of which some preliminary results were first reported in Ref. 23, Zapol[18] performed all-electron calculations using a Gaussian orbital basis set and found the minimum location to depend sensitively on the unit-cell volume. At the experimental equilibrium volume he obtains the order N_1, Γ_1, T_1+T_2.

In the present paper, we use all-electron full-potential linear muffin-tin orbital calculations in the local-density approximation. We find several close-lying conduction-band minima of which the lowest one is at a low symmetry point of the Brillouin zone. We note that the differences between the various minima are of order 0.1 eV, and may thus be modified by differences in quasiparticle self-energy corrections to the LDA eigenvalues. By examining the relations between the band structure of $ZnGeP_2$ and that of GaP, we estimate the corrections beyond the LDA from the differences between our LDA results and the quasiparticle calculations available for GaP.[24] We then show that our results may account for the available experimental data and point out some problems with the traditional pseudodirect gap interpretation. Our analysis includes a discussion of the crystal-field and spin-orbit splittings of the valence-band maximum.

Unfortunately, we must admit that the present computational uncertainty does not permit us to draw a firm conclusion on the pseudodirect or indirect nature of the gap. Rather, we think that a further experimental study is required. To facilitate further analysis and also because future modeling of the transport in this material may involve several of the minima, we present additional results for the deformation potentials for all the relevant gaps and for the effective masses of the various conduction-band minima and the Kohn-Luttinger type effective-mass Hamiltonian parameters[26] for the valence-band maximum manifold of states.

II. COMPUTATIONAL METHOD

The details of the computational method are as follows. We use the density-functional theory in the local-density approximation[27] as parametrized by Hedin and Lundqvist.[28] Since we are primarily interested in the band structure, we

0163-1829/99/60(11)/8087(10)/$15.00

choose to perform all calculations at the experimental lattice parameters. Typically, the LDA provides good values for c/a and internal structure parameters such as u but underestimates the overall absolute length scale of the system or the volume per unit cell by up to a few percent. We have minimized the energy versus volume, c/a, and u to determine how these parameters depend on each other. This then allowed us to also determine some of the band-gap hydrostatic and uniaxial deformation potentials at the experimental lattice constant.

The full-potential (FP) linear muffin-tin orbital (LMTO) method was used in the implementation of Methfessel and van Schilfgaarde.[29] It uses nearly touching muffin-tin radii with empty spheres inserted in the usual interstitial sites for a tetrahedrally coordinated material. The Zn 3d orbitals are treated as valence bands and an additonal so-called ‘‘local orbital’’[30] Zn 4d is included in the basis set. The deeper Ge 3d orbitals are treated as core states. Similarly for the auxiliary calculations of GaP, we treat the Ga 3d as valence orbitals and the Ga 4d as local orbitals. The local orbital treatment means that the $(\phi,\dot{\phi})$ linear combination of the linear method (ϕ and $\dot{\phi}$ being the radial wave-function solutions of the spherical average of the potential inside the sphere and its energy derivative at a chosen energy) is chosen such that the wave function vanishes outside the sphere rather than properly matching to the outside wave function at the boundary of the sphere. This is a sufficient approximation because the Zn 4d lies far from the bands of interest near the gap and therefore the precise scattering boundary conditions on that wave function are not important. Its inclusion nevertheless adds additional degrees of freedom to the wave function inside the sphere. This insures a better treatment of the conduction-band and valence-band maximum.

The angular momentum cutoff used for the interstitial region Hankel function basis set is $l_{\max}=6$. Brillouin-zone integration was carried out with a regularly spaced mesh of $8\times8\times8$ points in the reciprocal unit cell shifted from the origin as in the Monkhorst-Pack method[31] and reduced by symmetry to a set of irreducible $\mathbf{k}$ points. For the calculations of the spin-orbit splitting, we used the atomic-sphere approximation (ASA) to the LMTO method[32] after checking that the ASA results without spin-orbit coupling were in good agreement with the full-potential results for the bands of interest. Since spin-orbit coupling arises primarily from the inner part of the atomic spheres, where the potential is very close to being spherically symmetric, this should be an adequate approximation.

III. RESULTS

A. Band structures

The band structure of ZnGeP$_2$ is shown in Fig. 1. The corresponding Brillouin zone is shown in Fig. 2 along with that of the ZB. The CKP crystal structure is shown in Fig. 3. It is a superlattice of the zinc-blende structure with a specific ordered arrangement of the Zn and Ge cations accompanied by small structural distortions. It can be described in a body-centered-tetragonal primitive unit cell. The corresponding lattice vectors, $\mathbf{a}_1=(-a/2,a/2,c/2)$, $\mathbf{a}_2=(a/2,-a/2,c/2)$, and $\mathbf{a}_3=(a/2,a/2,c/2)$, are indicated in Fig. 3. The structural parameters are the a lattice constant, the c/a ratio, and the internal structural parameter u, which determines the position of the anion in its nearest neighbor tetrahedron. For example, the atom in the lower left corner has coordinates $(a/4,ua,c/8)$. In the ideal structure, $c/a=2$ and $u=1/4$.

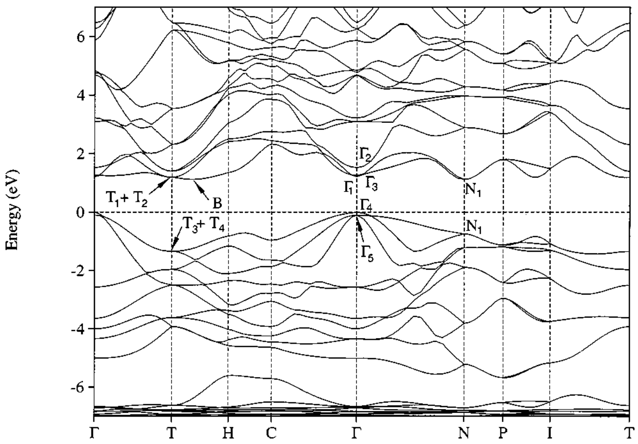

FIG. 1. Band structure of ZnGeP$_2$ in the local-density approximation at experimental lattice parameters. Some states of interest near the gap are labeled for further reference in the text.

We note that the lowest conduction-band state occurs at a point B, a point slightly away from T, closely followed by minima at N and at Γ. To better understand the nature of the bands, it is of interest to compare the band structure to that of the parent compound GaP. In fact, ZnGeP$_2$ can be thought of as being derived from the III-V compound GaP by replacing pairs of neighboring group-III elements (Ga) by a group II (Zn) or group IV (Ge) atom. This substitution can be considered as a relatively small perturbation. Thus, it is of interest to display the band structure of GaP in the CKP Brillouin zone. This can simply be done by performing the appropriate foldings of the ZB Brillouin zone. Specifically, we note that the X_{ZB} in the z direction (the direction in which the cubic unit cell is doubled in CKP) folds onto Γ. The X_{ZB} of the x and y directions folds onto T and the L_{ZB} point folds onto N.

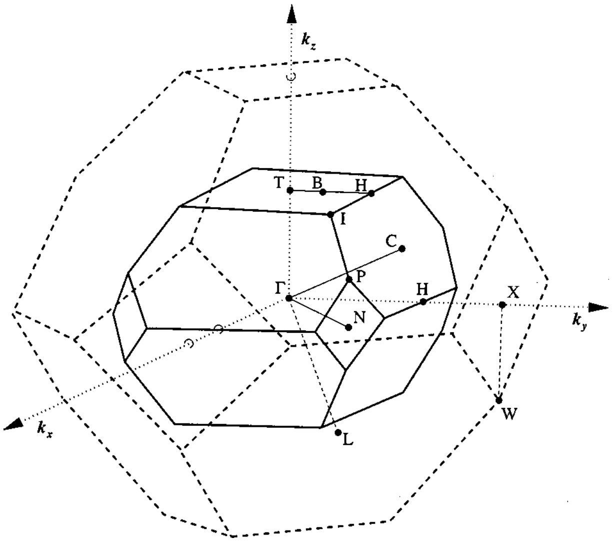

FIG. 2. First Brillouin zone (BZ) of chalcopyrite structure (solid lines) and its relation to that of the zinc-blende (ZB) structure (dashed lines). The points X, L, and W lie on the ZB BZ.

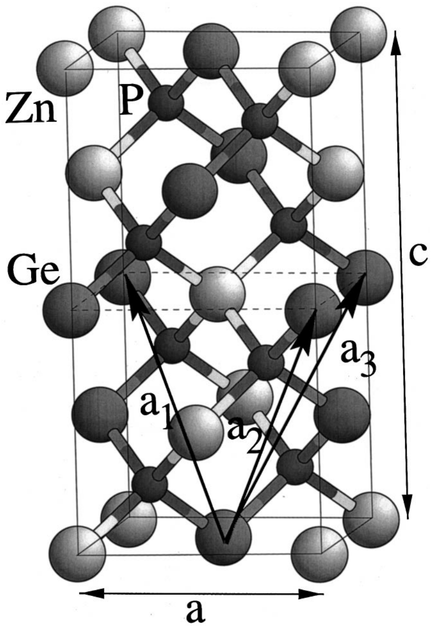

FIG. 3. The chalcopyrite crystal structure.

TABLE I. Selected conduction- and valence-band eigenvalues measured from the valence-band maximum in $ZnGeP_2$ and GaP.

		GaP			$ZnGeP_2$	
ZB	CKP	LDA	GW[a]	Expt.	LDA	corrected
Γ_1	Γ_1	1.60	2.85	2.89	1.24	2.49
$X_1(z)$	Γ_3	1.51	2.55	2.39	1.27	2.31
$X_3(z)$	Γ_2	1.68	2.81	2.75	1.54	2.67
$X_1(xy)$	T_1+T_2	1.51	2.55	2.39	1.20	2.24
L_1	N_1	1.50	2.67	2.64	1.14	2.31
$\Delta_{1\ \mathrm{min}}$		1.45				
	B_{min}				1.13	2.17
Γ_{15}	Γ_4	0.00	0.00	0.00	0.00	0.00
	Γ_5				−0.10	

[a] Reference 24.

[b] Because of the close proximity of B_{min} to the T point, we use the same correction for it as for the T point.

In Fig. 4, we show the band structures of GaP and $ZnGeP_2$ along a smaller portion of the CKP Brillouin zone. The states are labeled according to their ZB ''folding'' origin. Table I provides the correspondence between the CKP space group (D_{2d}^{12}) notation for the states to the ZB notation along with the numerical values of the states in question. We now see that in GaP, several locations compete for the minimum of the conduction band. At Γ, we note that the folded state X_1 is lower than the Γ_1 state. But also note that the L gap is very low and that the minimum actually occurs slightly away from X along the Δ_{ZB} axis. This location of the minimum, along the Δ axis, is in agreement with experiment.[33] This is related to the small splitting of the X_1 and X_3, which in turn is related to the low ionicity. This makes the band structure similar to that of Si near the X point. In $ZnGeP_2$, on the other hand, we find the Γ_1 state to be below the Γ_3 (folded X_1) state. However, both the minima at T and N are actually lower and the lowest state of all, as mentioned above, is at B.

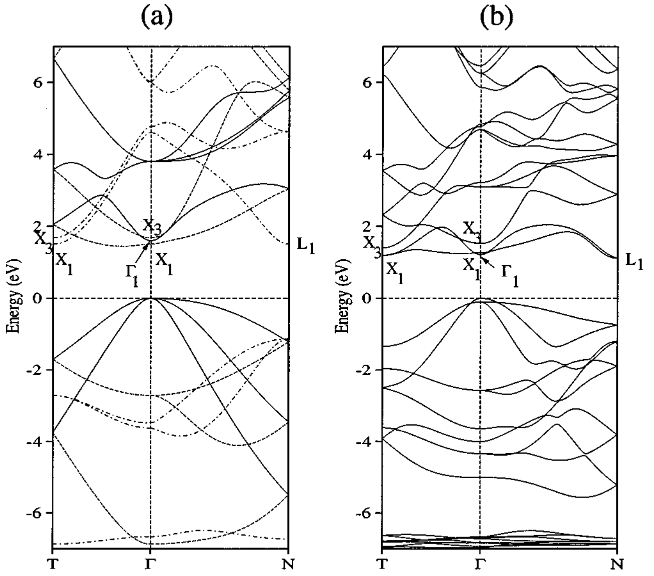

FIG. 4. Band structure of GaP (a) compared to that of $ZnGeP_2$ (b), both displayed in the chalcopyrite BZ but using ZB notation for labeling of the states of interest. In (a) the dashed and dash-dotted lines are obtained by folding.

Closer inspection reveals that at T the next higher band is of symmetry T_5. This corresponds to the folded $X_3(xy)$ of the parent compound (GaP). Along $T-H$ or B, the symmetry group permits only two representations and the order of the bands is, e.g., B_1, B_2, B_1, B_2. The closeness of the two B_1 bands is what leads to their repulsive interaction and thus to the minimum being displaced from T. Also, at N, the minimum is not strictly at N but slightly displaced. We will discuss this further in Sec. III D.

It is well known that the band gaps are not given accurately by the Kohn-Sham eigenvalues in the LDA because these do not truly represent quasiparticle excitation energies. The quasiparticle equation contains a nonlocal and energy-dependent self-energy operator $\Sigma_{\mathrm{xc}}(\mathbf{r},\mathbf{r}',\omega)$ instead of the exchange-correlation potential $v_{\mathrm{xc}}(\mathbf{r})$. For GaP, calculations of the quasiparticle excitations have been performed within the GW (Ref. 34) approximation.[24] We can use this information to estimate the band-gap corrections in $ZnGeP_2$. In fact, from the general aspects of GW theory[35] it is clear that the difference between the self-energy operator and the exchange-correlation potential depends primarily on the dielectric constant and the overall electron density, both of which are expected to be similar in these two compounds. Furthermore, the first-order correction $\langle \psi_i | \Sigma_{\mathrm{xc}} - v_{\mathrm{xc}} | \psi_i \rangle$ (using LDA eigenfunctions) should be similar for similar states. In terms of the folding we can establish a one-to-one correspondence between states in GaP and $ZnGeP_2$, which should differ only slightly. The GW calculations for GaP (Ref. 24) and the experimentally known optical transitions in GaP indicate that the self-energy corrections beyond the LDA are slightly larger at Γ than at X, with those of L falling in between. In fact, this ordering is quite general in the III-V semiconductors. Thus, we assume the same LDA→GW eigenvalue corrections in $ZnGeP_2$ as found in GaP for the corresponding states. This should work well since we treated GaP and $ZnGeP_2$ in exactly the same manner within the LDA.

Applying these corrections, we obtain the results indicated in Table I. We see that this interchanges the ordering of the levels. At Γ, we restore the previously accepted picture of the folded Γ_3 state being lower than Γ_1. Nevertheless, we find the N-point minimum to be below Γ_3 and the lowest state is still at B.

In the next section, we will consider the implications of these results for optical transitions and compare to experimental values. Before we do so, however, we need to discuss the valence-band splittings. Without spin-orbit interaction, the threefold-degenerate Γ_{15} state of ZB the splits into a singlet Γ_4 and doublet Γ_5 in CKP. Our calculated crystal-field splitting is -0.1 eV, where the minus sign means that the singlet lies above the doublet. When spin-orbit coupling is included, we switch to double group notation:[5,19] $\Gamma_4 \rightarrow \Gamma_6$ and Γ_5 splits into a Γ_7 and Γ_6 state, $\Gamma_1 \rightarrow \Gamma_6$, $\Gamma_2 \rightarrow \Gamma_6$, and $\Gamma_3 \rightarrow \Gamma_7$. We can adequately describe the splitting of the zinc-blende Γ_{15} with a quasicubic model, which includes a tetragonal crystal-field splitting Δ_c and a single spin-orbit parameter Δ_s. This model only neglects the possible anisotropy in the spin-orbit coupling, which should be small. One obtains the eigenvalue splittings:

$$E(\Gamma_7)-E(\Gamma_6)=\frac{\Delta_c+\Delta_s}{2}\pm\frac{1}{2}\sqrt{(\Delta_c+\Delta_s)^2-\frac{8}{3}\Delta_c\Delta_s}. \tag{1}$$

Using this model, we extract a spin-orbit and crystal-field splitting in $ZnGeP_2$ of 87 meV and -97 meV, respectively, from our first-principles calculation. This value of the crystal-field splitting is in good agreement with the one obtained directly in the FP-LMTO calculation without spin-orbit coupling, which is -100 meV. The value of the spin-orbit splitting is close to that of GaP (80 meV).[25]

B. Optical transitions

Optical absorption,[20] thermoreflectance,[11] electroreflectance,[13] wavelength modulated absorption,[14] cathodoluminescence,[20] and photoluminescence spectra[21,22] of the near-band-gap optical transitions have in the past been interpreted in terms of three series of (more or less) identically spaced transitions labeled $\{A',B',C'\}$, $\{A'',B'',C''\}$, and $\{A,B,C\}$ (in increasing order of energy). The most widely used model is that each of these series corresponds to transitions from the valence bands to a different conduction-band state at Γ, namely $\Gamma_6(3)$, $\Gamma_7(2)$, and $\Gamma_7(1)$, respectively. Here the subscript indicates the double group notation and the index in parentheses indicates the single group state from which it is derived. Sometimes, the lowest of these was ascribed as possibly being a combination of indirect transitions to the T_1+T_2 states and the pseudodirect transitions to the Γ_3 states. Within this model, the small splittings between A, B, and C should be the same as those in the corresponding primed and double-primed series, because they correspond to the splitting of the valence-band maximum. The transitions labeled A (with or without primes) originate from the valence-band maximum $\Gamma_6(4)$, the transitions labeled B originate from the $\Gamma_7(5)$, and the transitions labeled C originate from $\Gamma_6(5)$.

We note that not all lines are clearly resolved in all spectra. The unprimed $\{A,B,C\}$ series clearly corresponds to transitions to the Γ_1 state because they correspond to the strongest peaks in reflectance and therefore must be the dipole-allowed direct transitions. They are clearly visible in electroreflectance[13,12] and provide a good basis for assessing the valence-band splitting. Also, the polarization dependence of these states was clearly determined and confirms this interpretation, the A transitions appearing predominantly in parallel to **c**-axis polarization and the B exclusively in perpendicular.

Our present calculations give valence-band splittings A-B and B-C of 80 and 70 meV, respectively. The corresponding values obtained by Shay *et al.*[13] are 60 meV and 80 meV as derived from the strongest set of lines (the unprimed series). Shileika,[12] who summarized results of these splittings for a whole series of chalcopyrites, gives 70 and 60 meV for these splittings in $ZnGeP_2$. Gorban[20] gives slightly smaller splittings 50 and 40 meV as deduced from the $\{A',B',C'\}$ series and 40 and 40 meV as deduced from the $\{A'',B'',C''\}$ series. The discrepancy of this set of data obtained from cathodoluminescence and absorption spectroscopy already suggests that the interpretation of the primed series of peaks is not quite consistent with the model.

Our calculations would predict five distinct series of $\{A,B,C\}$ lines: one strong one corresponding to Γ_1, two weak ones very closely spaced at 172 and 176 meV below it corresponding to transitions to Γ_3 and N_1, another series at 239 meV below the top one corresponding to T_1+T_2, and finally one at 313 meV corresponding to transitions to B. Including the weak (pseudodirect) transitions to the Γ_2 would predict another weak series 182 meV above the main Γ_1 series. However, because this series would lie in the region above the strong direct absorptions, it would presumably be very difficult to observe. Figure 5 displays how one might roughly expect these series to add up. In this figure, we have simply used vertical lines of arbitrary but coded height for each series of lines. The series corresponding to the transitions to Γ_1 is indicated by a high intensity because these transitions are allowed. The other ones have been given heights of 0.2, 0.3, 0.4, 0.5, 0.6 for, respectively, transitions to Γ_2, B, T_1+T_2, N_1, Γ_3. Full lines correspond to transitions from $\Gamma_6(4)$, dotted lines to transitions from $\Gamma_7(5)$, and dashed lines to transitions from $\Gamma_6(5)$. One can notice that several of these transitions are predicted to fall very close to each other and to form a more or less evenly spaced set of features, suggestive of $\{A,B',C',A'',B'',A,B,C\}$ and thus not inconsistent with the data. We note in particular that except for the cathodoluminescence data above the edge as measured by Gorban *et al.*,[20] nobody seems to resolve the C'' line.

In the photoluminescence data of McRae *et al.*[21] and Petcu *et al.*,[22] one observes a fairly broadband emission with peaks at 1.4 eV and 1.6 eV on which different fine structure emerges depending on the excitation source. In Petcu *et al.*,[22] the near-gap luminescence is observed to peak at 2.35 eV with minor features interpreted as A',B',C', and B''. This

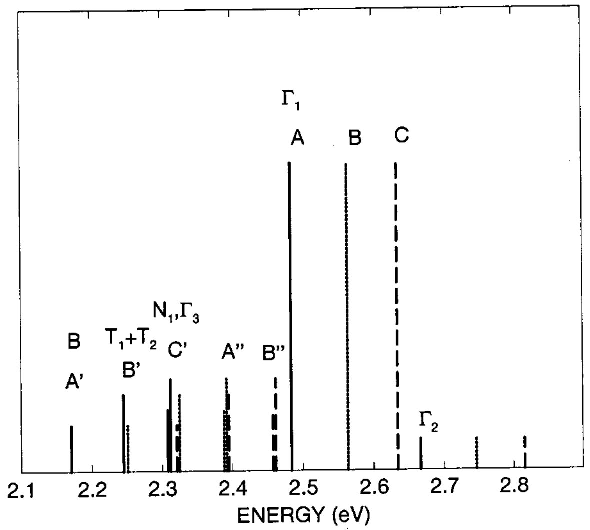

FIG. 5. Predicted optical transitions and their tentative interpretation of the observed experimental features. The full, dotted, and dashed lines correspond, respectively, to transitions from the $\Gamma_6(4)$, $\Gamma_7(5)$, and $\Gamma_6(5)$ valence bands, usually labeled as A,B,C. The heights of the bars are coded to represent transitions to Γ_2, B, T_1+T_2, N_1, Γ_3, Γ_1, as indicated for the full line series by the labels on top. They do not represent oscillator strengths, except that the transitions to Γ_1 were chosen to be drawn significantly stronger than all the other ones because it is the only truly direct dipole-allowed transition. The second label provides a connection to the peak labels used by experimental studies in the past.

corresponds to the room-temperature luminescence which is dominant at room temperature over the defect-related transitions. We note that their *A* feature, which appears at 2.35 eV, is consistent with the reflectivity data of Shay *et al.*[13] (at 2.34 eV) but differs from our position of the *A* feature at 2.49 eV because our calculation corresponds to absolute zero temperature and does not include the excitonic binding energy. The temperature-dependent cathodoluminescence data of Gorban *et al.*[20] indicate a shift of about 70 meV between 300 K and 4 K. By including an exciton binding energy of a few 10 meV, one can explain the 0.1 eV difference between theory and experiment of the direct Γ_1 gap. Moreover, we note that even the *GW* corrected band gaps have uncertainties of at least 0.1 eV.

From Fig. 5 one might expect the C' peak to be the next strongest feature in absorption or reflectivity after the *A* peak because four different indirect and pseudodirect transitions contribute to it. The C' peak in Fig. 5 occurs at about 0.2 eV below *A*, consistent with the data of Shay *et al.*,[13] who place *A* at 2.34 eV and C' at 2.11 eV.

In the present model the transitions are mostly mixtures of different indirect and pseudodirect transitions and very little if any polarization dependence of these peaks is expected. Because C' among others contains the $\Gamma_6(5)\rightarrow\Gamma_6(3)$ transition, which is purely perpendicularly polarized, one may expect a slightly stronger line in perpendicular polarization. This is consistent with the data of McRae *et al.*[21] for the peak centered around 1.4 eV. In their spectra, the more the excitation shifts towards shorter wavelengths or higher energy (including the cathodoluminescence), the more the transitions originating in the higher conduction-band levels show up. This is consistent with the expected thermalization of the carriers to the bottom of the band.

To gain further insight into the nature of the pseudodirect transitions, we have calculated the dipole transition matrix elements from the valence-band maximum to the conduction bands at Γ. This reveals that the transitions directly to Γ_3 should be a factor 10^4 weaker than those to Γ_1. Such a large difference is not observed. The indirect transitions generally involve a phonon to satisfy momentum conservation. It is safe to assume that the largest contribution comes from transitions involving an allowed transition to an intermediate state such as Γ_1 followed by a phonon emission or absorption. It should be recalled that energy conservation is not required for the transitions to and from the intermediate (or virtual) state, but only with the overall process involving the final indirect excitonic state including the emitted (or absorbed) phonon. Similarly, the fact that the pseudodirect transitions to Γ_3 are not as weak as calculated directly without phonons suggests that this must also be a phonon-assisted process. Further experimental investigations, such as piezomodulated transmission, may enable one to identify the phonons involved in these transitions and thereby may assist in determining which transitions are indirect or pseudodirect by providing signatures of specific phonons. See, for example, Glembocki and Pollak[36] and Alawadhi *et al.*[37] for studies of this type in GaP.

C. Deformation potentials

Because the various conduction-band minima under consideration are very close in energy to each other, it is important to consider their deformation potentials. In fact, their ordering could be changed by strain. As already mentioned, the values for the gaps given in Sec. III A used the experimental lattice constants. In order to determine the hydrostatic band-gap deformation potentials, we first need to investigate how the internal structural parameters *u* and the c/a ratio vary with unit-cell volume.

Our total-energy results are summarized in Table II. The *u*

TABLE II. Lattice parameters of $ZnGeP_2$.

	Present calc.	Expt.[a]	Relative error	CRYSTAL95	FLAPW
u	0.250	0.2582	-3%	0.251 (0.251)	0.250
c/a	1.9765	1.965	$+0.5\%$	1.986 (1.97)	
a(Å)	5.396	5.463	-1%	5.432 (5.62)	

[a] Reference 39.

[b] Zapol (Ref. 18), first number LDA, second (in parentheses) generalized gradient approximation (GGA).

[c] Continenza *et al.* (Ref. 38).

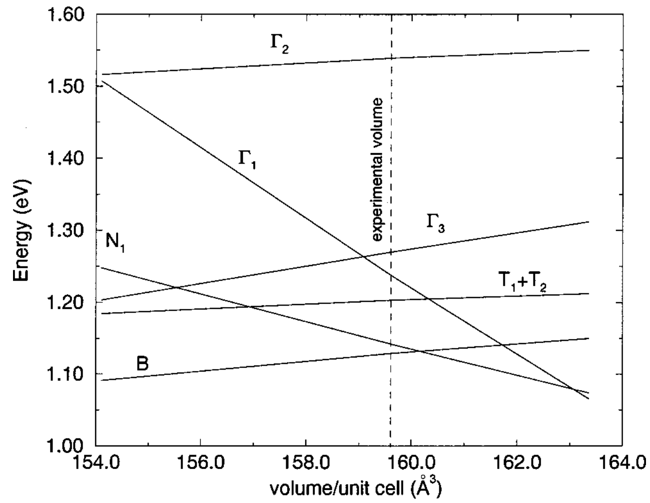

FIG. 6. LDA band gaps as functions of volume at the experimental c/a and u.

value is found to be close to ideal and the c/a value is found to be slightly too large. The overall length scale set by a is underestimated by about 1% as is usual for the LDA. These results are in very good agreement with other LDA calculations including the tendencies of the errors indicating that these are indeed due to the LDA and not to FP-LMTO convergence parameters. Our calculations further indicate that under compression the c/a decreases by only 1% per percent change in unit-cell volume: $d(c/a)/d\ln\Omega=0.01$ and that the equilibrium u remains unchanged as a function of lattice constant and c/a. Since these deviations are smaller than the errors in the calculated c/a and u compared to experiment, we decided to calculate the changes in the band gaps as a function of unit-cell volume with c/a and u kept fixed at their experimental values. The behaviors of the various gaps as a function of unit-cell volume are shown in Fig. 6. Note that these are the LDA gaps rather than the corrected gaps. One may note that several crossings occur. This means that the location of the minimum gap may change with strain. The corresponding deformation potentials $a_i=dE_i/d\ln\Omega$ are given in Table III. The values are accurate here to the extent that the corrections beyond the LDA are much less volume dependent than the LDA eigenvalues. The crystal-field splitting Δ_c is found to decrease with unit-cell volume by $d\Delta_c/d\ln\Omega=30$ meV.

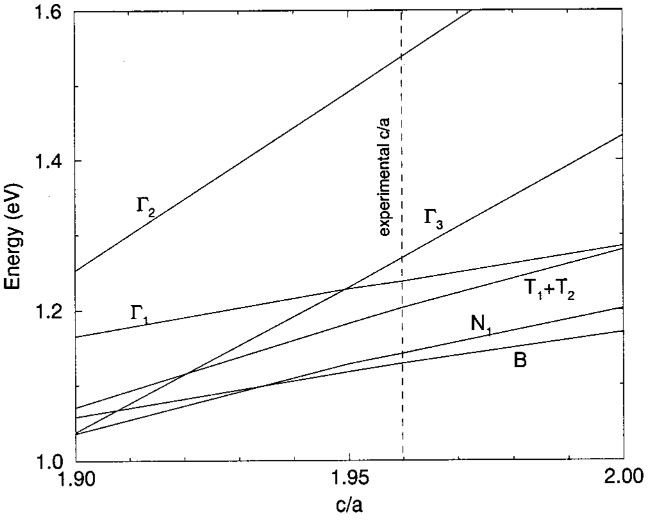

FIG. 7. LDA band gaps as functions of c/a at fixed experimental volume and u.

TABLE III. Band-gap deformation potentials in ZnGeP$_2$.

	$dE/d\ln\Omega$(eV)	$dE/d(c/a)$(eV)
Γ_1	-7.7	1.2
Γ_3	1.9	4.0
Γ_2	0.6	4.8
T_1+T_2	0.5	1.9
N_1	-3.0	1.5
B_{min}	1.0	1.0

Because in thin-film growth the strain in a film is often uniaxial, we have also calculated the deformation potentials $b_i=dE_i/d(c/a)$ at fixed experimental volume. Again, we have kept u fixed at the experimental value. These values are also given in Table III and the eigenvalues as a function of c/a are shown in Fig. 7.

D. Conduction-band effective masses

The conduction-band masses are summarized in Table IV. The situation is rather complex because of the numerous local minima. At Γ, the fourfold symmetry dictates that to second order in k, the bands are described by an ellipsoid of revolution, characterized by two masses, one parallel to $\mathbf{c}(m_\parallel)$ and one perpendicular to $\mathbf{c}(m_\perp)$. At T_1+T_2, the two bands become degenerate. As already mentioned, this zone-boundary degeneracy is related to time-reversal symmetry and the nonsymmorphic nature of the group, or the presence of a fourfold screw axis. The point is not a minimum but a saddle point for one of the two bands. The lower band becomes extremely flat in the direction parallel to $\Gamma-N$, making it nearly impossible to extract a significant mass, and

TABLE IV. Conduction-band effective masses.

state	m_1	m_2	m_3
Γ_1[a]	0.097	0.109	
Γ_3	-6.7	0.26	
Γ_2	0.40	0.24	
T_1+T_2[b]	0.46	-0.34	
	0.35	0.27	
N[c]	0.50	0.21	0.37
B_{min}[d]	1.03	0.27	0.28

[a] Near Γ each of the bands to order k^2 is described by an ellipsoid of revolution with $m_\parallel=m_1$ and $m_\perp=m_2=m_3$.

[b] At T there is a degeneracy between a band with a minimum and another one with a saddle point. m_1 corresponds to $T-\Gamma$, m_2 to $T-I$.

[c] The minimum actually occurs slightly (within $0.008\times2\pi/a$) away from N along $N-\Gamma$; m_1 is the longitudinal mass along $N-\Gamma$, m_2 is along $N-P$, and m_3 in the direction orthogonal to $N-\Gamma$ and $N-P$.

[d] m_1, m_2, m_3 along k_x, k_y, and k_z, respectively.

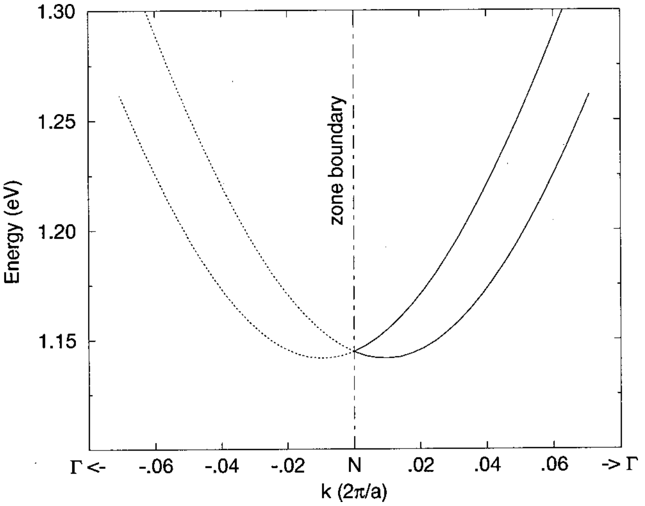

FIG. 8. Lowest conduction bands near the *N* point.

obtains a negative mass in the direction $T-H$. This leads to the actual minimum at *B* along $T-H$. At *B*, three different masses are required to describe the bands to second order. Near *N*, the minimum is actually very slightly removed from the symmetry point itself because of the symmetry-imposed degeneracy at the zone boundary. See Fig. 8. At the minimum, the lower band is characterized by three effective masses, one in the longitudinal direction $N-\Gamma$ and two in the two directions perpendicular to that, of which one is parallel to the **c** axis.

E. Valence-band effective Hamiltonian

Because of the close relationship of the present band structure to that of its cubic parent compound, we can attempt to describe the valence-band manifold by an effective quasicubic Hamiltonian,

$$\mathcal{H}=\Delta_c L_z^2+\frac{1}{3}\Delta_s \mathbf{L}\cdot\boldsymbol{\sigma}+Ak^2-(A-B)(\mathbf{L}\cdot\mathbf{k})^2$$

$$-2C\sum_{ij}[L_i,L_j]_+ k_i k_j, \tag{2}$$

in which Δ_c is the crystal-field splitting which splits (x,y)-like states from z-like states, Δ_s is the cubic spin-orbit splitting, $\boldsymbol{\sigma}$ is the Pauli spin matrix vector, $\mathbf{L}$ the angular momentum operator with subindices $i=x,y,z$, and the last sum is over the three cyclic permutations and $[L_i,L_j]_+=\frac{1}{2}(L_iL_j+L_jL_i)$. Aside from the crystal-field term, this Hamiltonian is the same as that for a cubic crystal with the A,B,C parameters being the usual cubic Kohn-Luttinger parameters.[26] The effect of the Δ_c terms is the same as that of a tetragonal strain in a cubic crystal.

We first study this Hamiltonian without the spin-orbit coupling term. In terms of the *A*, *B*, *C* parameters, the directional-dependent hole effective masses can be summarized by

$$m(\Gamma_4)=-1/A, \quad m(\Gamma_5)=-1/B \tag{3}$$

for the $\Gamma-T$ direction $(k_x=k_y=0)$, and where $m(\Gamma_4)$ means the mass of the band originating in the Γ_4 state;

TABLE V. Valence-band effective Hamiltonian parameters.

Δ_c (meV)	-97
Δ_s (meV)	87
$A(\hbar^2/2m_e)$	-9.90
$B(\hbar^2/2m_e)$	-2.75
$C(\hbar^2/2m_e)$	-11.65

$$m(\Gamma_4)=-1/B, \quad m_h(\Gamma_5)=-1/B, \quad m_l(\Gamma_4)=-1/A \tag{4}$$

for the direction $\Gamma-H$, $k_y=k_z=0$, and where *h* and *l* subindices stand for heavy and light, respectively, and

$$m(\Gamma_4)=-1/(2B), \quad m_h(\Gamma_5)=-1/(A+B-C),$$

$$m_l(\Gamma_4)=-1/(A+B+C) \tag{5}$$

for the $\Gamma-N$ direction $(k_x=k_y,k_z=0)$. Within this model, certain relations are predicted between the masses along different directions. For example, the mass of the doubly degenerate band originating in Γ_5 along $\Gamma-T$, which we find to be 0.358, should equal the mass of the heavy-hole band originating in Γ_5 along k_x, for which we find 0.383, and of the band originating in Γ_4 along k_x, for which we find 0.364, and should be twice the mass of the Γ_4 derived band along $\Gamma-N$, for which we find 0.171. Similarly, *A* determines the light mass of the Γ_5 derived band along k_x, for which we find 0.1098 and the Γ_4 derived band along k_z, for which we find 0.1060. We see that this model with only three inverse mass parameters is satisfied rather well. The values for *A*, *B*, and *C*, given in Table V, are determined by a least-square fit to a regularly spaced grid of points within a spherical region of **k** space extending to $0.06\times 2\pi/a$, or roughly to energies down to 300 meV below the valence-band maximum. The residual root-mean-square error is only 2 meV and the maximum error is 9 meV.

This model also satisfactorily reproduces the fourfold symmetry of the energy contours of the Γ_5 derived bands in the *xy* plane. Following Lietz and Rössler (LR),[6] the latter can be described by

$$E(\Gamma_5)_{1,2}=\Delta_c+C_1k_z^2+C_2k_\perp^2$$

$$\pm k_\perp^2\sqrt{C3^2+C_4^2+2C_3C_4\cos 4\varphi}, \tag{6}$$

where φ measures the angle from the k_x direction. The parameters in this model are related to our present model by

$$C_1=B, \quad C_2=A+B, \quad C_3=(A-B+C)/4,$$

$$C_4=(A-B-C)/4, \tag{7}$$

as can easily be checked by comparing the expressions for the $\varphi=0$ and $\varphi=\pi/4$ directions.

In LR's model, the bands derived from Γ_4 are given by

$$E=C_1'k_z^2+C_2'k_\perp^2. \tag{8}$$

In our model, $C_1'=A$, $C_2'=B$.

Following Bir and Pikus's theory of invariants,[40] our Hamiltonian matrix could be generalized as follows:

$$\mathcal{H}=\begin{pmatrix} \Delta_c+Ak_x^2+B(k_y^2+k_z^2) & Ck_xk_y & C'k_yk_z \\ Ck_xk_y & \Delta_c+Ak_y^2+B(k_x^2+k_z^2) & C'k_xk_z \\ C'k_yk_z & C'k_xk_z & A'k_z^2+B'(k_z^2+k_y^2) \end{pmatrix}. \tag{9}$$

Within this generalized model, the relation to LR's parameters is obviously that $C_1'=A'$, $C_2'=B'$. The possibility of a C' parameter different from C would allow for the C_3 and C_4 parameters to depend on k_z and goes beyond LR's approach of separating the Γ_4 and Γ_5 derived bands from the start as separate vector spaces. From our fittings of the first-principles bands to the simple quasicubic Kohn-Luttinger model, which does not distinguish between primed and unprimed A,B,C parameters, it appears that the errors due to the non-parabolicity (as apparent from slightly different values obtained depending on the range of **k** space over which we fit) are larger than the differences between the primed and unprimed quantities.

When the spin-orbit coupling terms are added, an additional splitting occurs at Γ, as already discussed in Sec. III A. In addition, the masses are affected because new couplings become allowed by symmetry and the bands become nonparabolic. The bands including spin-orbit coupling deduced from our effective Hamiltonian are shown in Fig. 9 along with the first-principles results. The agreement is excellent, indicating that there is no need to complicate the model with an anisotropic spin-orbit splitting. The effective masses at Γ in various directions are summarized in Table VI. Note that in the $\Gamma-T$ direction, these are close to the ones obtained without spin-orbit coupling (0.11 for the Γ_4 derived band and 0.36 for the Γ_5 derived bands). For the $\Gamma-N$ direction, however, the bands derived from Γ_5 have either a very high (1.34) or very low mass (0.04) without spin-orbit coupling. With spin-orbit coupling these masses become almost equal but the bands are strongly nonparabolic. The Γ_4 derived band would have a mass of 0.17 without spin-orbit coupling.

IV. CONCLUSION

In this paper, we have studied the lower conduction bands and upper valence bands in ZnGeP$_2$ in detail by first-principles FP-LMTO calculations. In particular, we have reinvestigated the question of the pseudodirect or indirect nature of the minimum gap in this semiconductor. While there still remains some uncertainty regarding this issue, due to the presently attainable accuracy in quasiparticle calculations from first principles, our calculations lead to several new findings.

TABLE VI. Hole effective masses including spin-orbit coupling (in units of m_e).

k direction	m_A	m_B	m_C[a]
k_z or $\Gamma-T$	-0.107	-0.364	-0.300
$k_x=k_y$, $k_z=0$, or $\Gamma-N$	-0.329	-0.158	-0.166

[a] A, B, C correspond, respectively, to $\Gamma_6(4)$, $\Gamma_7(5)$, $\Gamma_6(5)$.

First, it is clear that several energetically very closely spaced conduction-band minima occur at different points in the Brillouin zone, all of which should be taken into account on an equal footing in future studies of transport and absorption edge optical properties. The point at which we find the actual minimum has to the best of our knowledge not previously been identified. It appears to be a point slightly away from the T_1+T_2 state on the surface of the Brillouin zone in the k_x direction. The reason for the displacement of the minimum from the high-symmetry point was identified in terms of interactions between close-lying bands of the same symmetry. While we did not explicitly perform GW calculations, we added GW corrections to our LDA results by making the well-justified assumption that these should be very close to those in GaP at the corresponding **k** points. The systematic differences in GW corrections between different **k** points play a significant role in establishing the energetic order of the various conduction-band states under consideration. In other words, a simple constant energy gap shift over the Brillouin zone is not sufficient.

We reexamined the optical-absorption edge data within the light of our model and came to the conclusion that the previously widely accepted assignment is not valid. The traditional assignment is based on three well-separated sets of three closely spaced lines, the separation of the sets being determined by the conduction-band splittings and the splittings within each set being determined by the valence-band splittings. In our model, on the other hand, more conduction-band minima are considered and their spacings are comparable to the valence band splittings, leading to a more complex interweaving of these sets of lines. We show that our model may explain why certain lines in the traditional model

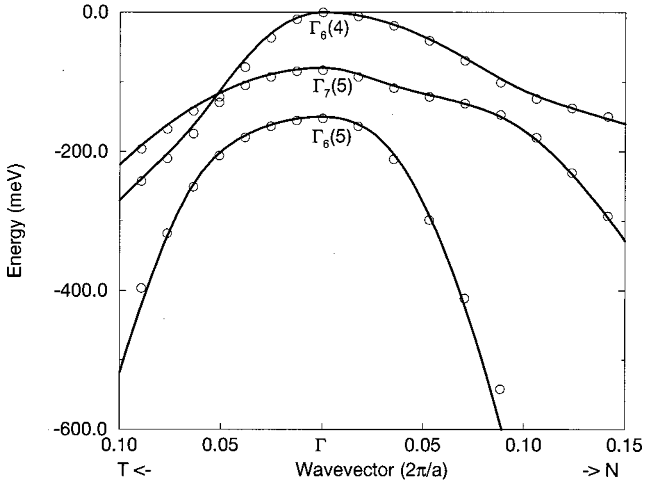

FIG. 9. Valence bands near maximum including spin-orbit coupling: circles, first-principles (ASA-LMTO) results; solid lines, as derived from effective Hamiltonian given in Eq. (2) and using parameters given in Table V.

appear to be missing from the data and why the polarization dependence is not as clearly established as would be expected on the basis of the previous model. We think that this results basically from the fact that some of the observed features are actually superpositions of several transitions with different selection rules. While there still remains some computational uncertainty in our present results as to the precise order of states and their splittings, it is clearly unjustified to simply neglect the indirect minima in favor of the direct (or pseudodirect) transitions at Γ. Thus further experimental work with higher resolution seems to be required to understand the optical data. This would hopefully also lead to the identification of the phonons involved in these phonon-assisted transitions and may thus finally resolve the main question posed in this paper. We also noted that even the pseudodirect transitions to Γ_3 probably are phonon-assisted. In other words, the phonon-assisted processes appear to have higher oscillator strength than the direct process, which we found to be much weaker compared to the truly direct transitions to Γ_1 than indicated by experiment.

Finally, because the various conduction-band minima may play a role in transport and in future interpretations of the optical data, we extracted further details about these minima from our calculations. In particular, we presented effective masses and their anisotropies for the various minima investigated. We also present a relatively simple quasicubic effective Hamiltonian for the upper valence bands around $\mathbf{k}=\mathbf{0}$. Its parameters were determined and it was shown to accurately describe the valence bands in the relevant region. Because the conduction-band minima are so closely spaced, it is also important to study their shifts in energy and possible changes in ordering due to strain. For that purpose, we have calculated the band-gap deformation potentials for hydrostatic and tetragonal uniaxial strains for the various minima.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under Grant No. DMR 95-29376.

---

[1] See, for example, M.C. Ohmer and R. Pandey, MRS Bull. **23**, 16 (1998), for an overview of the literature and history of this class of materials.

[2] J. L. Shay and J. H. Wernick, *Ternary Chalcopyrite Semiconductors: Growth, Electronic Properties, and Applications* (Pergamon Press, New York, 1975).

[3] P.G. Schunemann and T.M. Pollak, MRS Bull. **23**, 23 (1998), and references therein.

[4] N.C. Giles and L.E. Halliburton, MRS Bull. **23**, 37 (1998), and references therein.

[5] R. Sandrock and J. Treusch, Z. Naturforsch. A **19a**, 844 (1964).

[6] M. Lietz and U. Rössler, Z. Naturforsch. A **19a**, 850 (1964).

[7] G.F. Karavaev and A.S. Poplavnoi, Fiz. Tverd. Tela (Leningrad) **8**, 2143 (1966) [Sov. Phys. Solid State **8**, 1704 (1967)].

[8] B. Ray, A.J. Payne, and G.J. Burrell, Phys. Status Solidi B **35**, 197 (1969).

[9] S. Isomura and K. Masumoto, Phys. Status Solidi A **13**, 223 (1972).

[10] R. Bendorius, V.D. Prochukhan, and A. Shileika, Phys. Status Solidi B **53**, 745 (1972).

[11] A. Raudonis, V.S. Grigoreva, V.D. Prochkhan, and A. Šileika, Phys. Status Solidi B **57**, 415 (1973).

[12] A. Šileika, Surf. Sci. **37**, 730 (1973).

[13] J.L. Shay, B. Tell, E. Buehler, and J.H. Wernick, Phys. Rev. Lett. **30**, 983 (1973).

[14] G. Babonas, G. Ambrazevičius, V.S. Grigoreva, V. Neviera, and A. Šileika, Phys. Status Solidi B **62**, 327 (1974).

[15] C. Varea de Alvarez and M.L. Cohen, Phys. Rev. Lett. **30**, 979 (1973); C. Varea de Alvarez, M.L. Cohen, S.E. Kohn, Y. Petroff, and Y.R. Shen, Phys. Rev. B **10**, 5175 (1974).

[16] A. Heinrich, W. Cordts, and J. Monecke, Phys. Status Solidi B **107**, 319 (1981).

[17] J.E. Jaffe and A. Zunger, Phys. Rev. B **30**, 741 (1984).

[18] P. Zapol, Ph. D. thesis, Michigan Technological University, 1998; Avail. UMI Order No. DA9823065.

[19] V.A. Chaldyshev and V.N. Pokrovskii, Izv. Vyssh. Uchebn. Zaved. Fiz. **2**, 173 (1960).

[20] I.S. Gorban, V.B. Grishchuk, I.G. Tregub, and M.V. Chukichev, Fiz. Tekh. Poluprovodn. **18**, 1426 (1984) [Sov. Phys. Semicond. **18**, 892 (1984)].

[21] J.E. McRae, Jr., M.R. Gregg, R.L. Hengehole, Y.K. Yeo, P.H. Ostdiek, M.C. Ohmer, P.G. Schunemann, and T.M. Pollack, Appl. Phys. Lett. **64**, 3142 (1994).

[22] M.C. Petcu, N.C. Giles, P.G. Schunemann, and T.M. Pollack, Phys. Status Solidi B **198**, 881 (1996).

[23] S. Limpijumnong, S.N. Rashkeev, and W.R.L. Lambrecht, Bull. Am. Phys. Soc. **43**, 905 (1998).

[24] X. Zhu and S.G. Louie, Phys. Rev. B **43**, 14 142 (1991).

[25] *Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung, Landolt-Börnstein New Series, Group III, Vol. 22, Part a (Springer, Berlin, 1991).

[26] J.M. Luttinger, Phys. Rev. **102**, 1030 (1956); J.M. Luttinger and W. Kohn, *ibid.* **97**, 869 (1955).

[27] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L.J. Sham, *ibid.* **140**, A1133 (1965).

[28] L. Hedin and B.I. Lundqvist, J. Phys. C **4**, 2064 (1971).

[29] M. Methfessel, Phys. Rev. B **38**, 1537 (1988); M. van Schilfgaarde (unpublished).

[30] D. Singh, Phys. Rev. B **43**, 6388 (1991).

[31] H.J. Monkhorst and J.D. Pack, Phys. Rev. B **13**, 5188 (1976).

[32] O.K. Andersen, Phys. Rev. B **12**, 3060 (1975); O. K. Andersen, O. Jepsen, and M Šob, in *Electronic Band Structure and Its Applications*, edited by M. Yussouf (Springer, Heidelberg, 1987), p. 1.

[33] P.J. Dean and D.C. Herbert, J. Lumin. **14**, 55 (1976).

[34] G stands for the one-electron Green's function and W for the screened Coulomb interaction used in Hedin and Lundqvist's (Ref. 35) first-order approximation to the self-energy operator.

[35] L. Hedin and S. Lundqvist, Solid State Phys. **23**, 1 (1969)

[36] O.J. Glembocki and F.H. Pollak, Phys. Rev. B **25**, 1179 (1982); **25**, 1193 (1982).

[37] H. Alawadhi, R. Vogelgesang, A.K. Ramdas, T.P. Chin, and J.M. Woodall, J. Appl. Phys. **82**, 4331 (1997).

[38] A. Continenza, S. Massidda, A.J. Freeman, T.M. de Pascale, F. Meloni, and M. Serra, Phys. Rev. B **46**, 10 070 (1992).

[39] A. MacKinnon, in *Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung, Landolt-Börnstein New Series, Group III, Vol. 17, Pt. h (Springer, Berlin, 1985).

[40] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (John-Wiley & Sons, New York, 1974).