Ab Initio Calculation of Linear and Nonlinear Optical Properties of Semiconductor Structures

F. Bechstedt, B. Adolph and W. G. Schmidt*

Friedrich-Schiller-Universität, Max-Wien-Platz 1, 07743 Jena, Germany

Received 27 May, 1999

The theoretical and numerical approaches are discussed for ab initio calculations of optical properties. The density functional theory (DFT) combined with the local-density approximation (LDA) allows the calculation of the geometry of crystalline solids and their surfaces with a precision of about one percent. The DFT-LDA band structure and single-electron states therefore provide a reasonable starting point for the calculation of linear and nonlinear susceptibilities within the independent-particle approximation. However, this approach has to be improved by taking into account many-body interactions: self-energy effects, local-field corrections, and electron-hole attraction. Three types of optical spectra are studied: the frequency-dependent dielectric function, the second-harmonic generation, and surface reflectance anisotropy spectra. The systems considered are two-atomic semiconductors, their polytypes and their surfaces.

I Introduction

The recent major advances in the quantitative computation of ground-state properties of solids are essentially related to the development of density-functional theory (DFT) in local-density approximation (LDA) [1]. However, spectroscopic properties, as optical spectra, are not accessible in this theory, since its single-particle eigenvalues or differences of them cannot be directly identified with electronic excitation energies. For instance, energy gaps of semiconductors and insulators are remarkably underestimated within DFT-LDA [2]. Instead, quasiparticle (QP) energies have to be computed. The basic approach to quasiparticles is Hedin’s GW approximation [3], in which the spatially nonlocal and energy-dependent exchange-correlation (XC) self-energy $\Sigma$ is approximated by a convolution of single-particle Green function $G$ and dynamically screened Coulomb interaction $W$. Vertex corrections are neglected. Full numerical calculations of the single-particle excitations can now be performed on real crystals starting from the DFT-LDA results [4,5].

Despite of the complications with the inclusion of the many-body effects, DFT-LDA has been used to determine the atomic structure and to define a reasonable starting point for the description of the electronic structure. Optical spectra have been calculated [6] within the independent-particle approximation [7]. Meanwhile, quasiparticle shifts have been included also in calculations of the complete dielectric function [6,8]. Moreover, the effects of XC beyond the independent-particle approximation have been demonstrated within DFT-LDA for bulk semiconductors, and local-field effects are taken into account [9]. Very recently the compensation effect of the dynamical screening in quasiparticle and vertex corrections has been found [10]. Furthermore, the Coulomb attraction of excited electrons and holes has been indeed calculated in the limit of static screening [11].

In the last years such kind of ab initio calculations in the framework of DFT-LDA and the independent-particle approximation has been also applied to nonlinear optical properties of semiconductors, e.g. the second-harmonic generation (SHG) [12,13], and to semiconductor surfaces [14-16]. The surface activities have been focused on the reflectance anisotropy (RA) spectra [17]. The inclusion of the many-body self-energy and excitonic effects is an unsolved problem. However, at present many activities occur in this field. They mainly concern the inclusion of the QP shifts.

After a brief description of the calculational methods and basic expressions, the status of the ab initio description of the optical properties is demonstrated for three examples, the frequency-dependent dielectric function, the second-harmonic generation and the surface reflectance anisotropy. The semiconducting materials considered are the group- IV semiconductor Si as

*Present address: Department of Physics, North Carolina State University, Raleigh, U.S.A.
well as the compounds SiC and InP. In the SHG case also other III-V compounds are studied. Silicon carbide (SiC) crystallizes in the two-atomic zinc-blende structure (3C) but also in a variety of polytypes (nH), i.e., in natural superlattices with hexagonal (H) symmetry and $n$ differently stacked Si-C bilayers in a unit cell [18,19]. The polar surfaces InP(001) show a variety of reconstructions. The most important ones exhibit a $2 \times 4$ translational symmetry [20], the atomic structure of which varies with the surface stoichiometry.

II Methods

The equilibrium atomic configurations of the SiC polytypes [18,19] and InP(001) surfaces [20] are determined from fully converged self-consistent total-energy calculations [21]. The electron-ion interaction is treated by norm-conserving, $ab$ initio, fully separable pseudopotentials in the Kleinman-Bylander form [22]. The electronic wave functions are expanded in terms of plane waves. The energy cutoffs are 34 Ry (SiC) and 15 Ry (Si, InP, other III-V).

In order to examine optical properties we first calculate the imaginary parts of the elements of the second-rank dielectric tensor $\varepsilon_{\alpha\beta}(\omega)$ or the polarizability $\alpha_{\alpha\beta}(\omega)$. Within the independent-particle approximation (RPA) and the optical limit [6,8] one has,

$$\text{Im}\alpha_{\alpha\beta}(\omega) = \frac{2\pi\varepsilon^2 R^2}{V} \sum_{c,v} \sum_{k} \frac{<ck|v_\alpha|k> <ck|v_\beta|k>}{[\varepsilon_c(k) - \varepsilon_v(k)]^2} \delta[\varepsilon_c(k) - \varepsilon_v(k) - \hbar\omega],$$

where Bloch eigenfunctions $|nk>$ belonging to a band index $n = v,c$ (occupied, empty), the wave vector $k$ in the first Brillouin zone (BZ), and the band energy $\varepsilon_n(k)$ are introduced.

The real parts of the components of the dielectric tensor, $Re\varepsilon_{\alpha\beta}(\omega)$, are calculated from the imaginary parts by means of the Kramers-Kronig transformation. Local-field effects are neglected. To investigate many-body effects, the DFT-LDA eigenvalues $\varepsilon_n(k)$, are shifted by quasiparticle corrections calculated according to a scheme developed by Cappellini et al. [23].

The calculated quasiparticle shifts depend on the band index and the wave vector. The renormalization of the spectral strengths due to the energy dependence of the XC self-energy is treated according to Ref. [6]. The generalization of expression (1) for the case of inclusion of excitons is described elsewhere [11]. Within the independent-particle approximation and using the Coulomb gauge, the imaginary part of the tensor of the SHG susceptibility $\chi_{\alpha\beta\gamma}^{(2)}(\omega) := \chi_{\alpha\beta\gamma}^{(2)}(-\omega,\omega)$ for a crystal of arbitrary symmetry can be described by

$$\text{Im}\chi_{\alpha\beta\gamma}^{(2)}(\omega) = -\frac{\pi}{2V} \left(\frac{e\hbar}{m}\right)^3 \sum_{k} \sum_{c,v} \frac{1}{[\varepsilon_{cv}]^3} \times$$

$$\times \left\{ \sum_{c'} \left( \frac{16G_{c'cv}^{\alpha\beta} \delta(\varepsilon_{cv} - 2\hbar\omega)}{\varepsilon_{cv} + \varepsilon_{c'} + \varepsilon_{c'}} - \frac{G_{c'cv}^{\alpha\beta} \delta(\varepsilon_{cv} - \hbar\omega)}{\varepsilon_{cv} + \varepsilon_{c'}} + \frac{G_{c'cv}^{\alpha\beta} \delta(\varepsilon_{cv} + \hbar\omega)}{\varepsilon_{cv} + \varepsilon_{c'}} \right) \right. \right.$$

$$\left. - \sum_{v'} \left( \frac{16G_{c'cv'}^{\alpha\beta} \delta(\varepsilon_{cv'} - 2\hbar\omega)}{\varepsilon_{cv'} + \varepsilon_{c'v'}} - \frac{G_{c'cv'}^{\alpha\beta} \delta(\varepsilon_{cv'} - \hbar\omega)}{\varepsilon_{cv'} + \varepsilon_{c'v'}} + \frac{G_{c'cv'}^{\alpha\beta} \delta(\varepsilon_{cv'} + \hbar\omega)}{\varepsilon_{cv'} + \varepsilon_{c'v'}} \right) \right\}$$

with interband energies abbreviated by $\varepsilon_{mn'} = \varepsilon_n(k) - \varepsilon_{n'}(k)$ and the combination of the optical matrix elements

$$G_{mm'n'}^{\alpha\beta\gamma} = \text{Im}\{p_{mm'}^{\alpha}(k)p_{n'n'}^{\beta}(k)p_{mn'}^{\gamma}(k) + p_{m'n'}^{\alpha}(k)p_{n'n'}^{\beta}(k)p_{mm'}^{\gamma}(k)\}.$$
The latter ones are given by the matrix elements of the optical transition operator between Bloch states \( \psi_{m,\mathbf{k}}^0 = \langle m | n \rangle < \mathbf{k} | v \rangle \). From the imaginary parts of the SHG susceptibility, the real parts and absolute values can be calculated using a Kramers-Kronig-like transformation. In equation (2) both virtual-electron \( (\nu \sigma') \) and virtual-hole \( (\nu \sigma') \) contributions are taken into account [12,13].

Even for cubic crystals polarization-dependent corrections to the Fresnel reflectance appear. The origin of these reflectance anisotropies is located in the surface region. In the case of normal incidence the polarization-induced relative change is given by [17],

\[
\frac{\delta R}{R} = \frac{R_\alpha - R_\beta}{R} = \frac{4\omega d}{c} \left[ \frac{4\pi}{\varepsilon^b} \left( \alpha_{\alpha\alpha}(\omega) - \alpha_{\beta\beta}(\omega) \right) \right],
\]

where \( \varepsilon^b(\omega) \) is the bulk dielectric function, and \( \alpha_{\alpha\alpha} \) and \( \alpha_{\beta\beta} \) are components of the optical polarizability tensor of the slab with thickness \( d \). In the following we identify the subscripts \( \alpha \) and \( \beta \) with the [110] and [110] directions, respectively.

To model the InP(001) surface we consider a periodic supercell along the surface normal. It contains 12 atomic (001) layers and a vacuum region equivalent in thickness to 8 atomic layers. The dangling bonds at the bottom layer of the slab are saturated with fractionally charged pseudohydrogens [20].

### III Optical properties

#### A. Frequency-dependent dielectric function

As typical results of \textit{ab initio} calculations within the independent-particle or independent-quasiparticle approximation the spectral variation of the dielectric function is plotted in Fig.1 for the semiconductor silicon, crystallizing in diamond structure. In addition to the real and imaginary part, also the reflectivity \( R(\omega) \) is plotted according to the Fresnel formula. The theoretical spectra are compared with experimental ones [24,25]. QP effects are taken into account in the framework of the scissors operator approach. The theoretical spectra have been shifted to higher energies by 0.47 eV, which brings the zeros in the real part together.

The spectra in Fig.1, calculated within the independent-QP approach, reproduce all main structures of the experimental spectra. This holds in particular for the \( E_g \) peak at about 4.3 eV in the absorption spectrum. The \( E_1 \) peak at about 3.4 eV appears only as a shoulder in the spectrum. For many years we know that the underestimation of its oscillator strength is related to the neglect of excitonic effects [26]. The binding of the saddle-point excitons near \( E_1 \) is enforced due to the nearby parallel bands along the \( \Gamma L \) direction in the BZ. The low-photon-energy region is accompanied with an overestimation for higher energies. A treatment beyond the independent-QP approach [11] strongly improves the agreement between theory and experiment in the case of the dielectric function. On the other hand, the Fresnel reflectivity is already reasonably described within the independent-QP approach. This holds for its absolute values but also its spectral variation.

![Figure 1. Real and imaginary part of the dielectric function and the Fresnel reflectivity versus photon energy for silicon. Solid line: calculation within independent-QP approach, dotted line [24] and dashed line [25]: experimental data.](image)

The atomic geometry strongly influences the electronic structure and thereby the resulting optical properties of crystals. A prototypical material to study the influence of the crystal structure on optical spectra is SiC. Silicon carbide occurs in more than 200 polytypes [18]. The two most extreme polytypes are zinc blende (3C) with pure cubic stacking of the Si-C bilayers in the [111] direction and wurtzite (2H) with pure hexagonal stacking in the [001] direction. Other hexagonal (H) polytypes, nH, represent combinations of these stacking sequences with a periodicity of \( n \) double layers in the stacking direction [19,27]. It is well known that the polytypism strongly influences the physical and chemical properties. For example, the energy gaps and the location of the conduction-band minima in \( k \) space vary with the crystal structure (cf. Refs. [19,23] and references therein). With a change of the indirect energy gap of about 1 eV between the 3C and 2H polytypes
SiC represents an extraordinary example for the polytype influence on the electronic structure. Optical spectroscopy should make visible the drastic changes in the electronic properties with the polytype. A systematic investigation of the vacuum-ultraviolet reflectivity has been presented for 3C, 4H, and 6H crystals [30,31]. Recently complete dielectric functions have been also measured by means of the spectroscopic ellipsometry [32,33].

Figure 2. Real and imaginary part of the dielectric function for light polarization perpendicular to the c-axis of the SiC polytypes. Results of the spectroscopic ellipsometry are indicated by dotted lines [33].

Figure 3. Reflectivity for normal incidence and light polarization perpendicular to the c-axis of the SiC polytypes. Measured spectra are given as dotted lines [30] or dashed lines [31].

Results calculated according to expression (1) are presented in Fig. 2 for the dielectric function and in Fig. 3 for the reflectivity for normal incidence and light polarization perpendicular to the c-axis. To compare with experimental spectra [30,31,32], the theoretical spectra are shifted to higher photon energies by a constant quasiparticle shift of about 0.75 eV. The spectra are strongly influenced by the polytypism. For the hexagonal polytypes 4H and 6H, the pronounced double peak structure in the imaginary part of 3C shifts to lower energies and becomes broadened. A similar tendency is observed for the real part and the reflectivity. We conclude that the BZ folding effect accompanying the enlargement of the unit cell induces an effective smearing out of the structures. It is therefore related to the period of the translational symmetry in c-axis direction, but not to the relative number of the twisted bonds, i.e., the hexagonality of the polytype. The hexagonality much stronger influences the polarization dependence of the spectra.

B. Second-harmonic generation

The calculation of nonlinear optical properties is much more complicated than the same procedure in the linear case. The difficulties concern both the numerics and the physics. The k-space integration in expression (2) has to be performed more carefully using a generalization of the tetrahedron method [12,13]. More conduction bands have to be taken into account to reach the same accuracy. In principle, there is still no rigorous ansatz to describe local-field effects in contrast to the linear case [34]. Making use of the Coulomb gauge of the electromagnetic field and the quadratic-response theory [13,35] the second-order optical properties can be related to three-particle Green functions. A certain
decoupling may indicate the way to introduce correctly excitonic effects. On the other hand, there are also complications to determine the absolute value and the frequency dependence of nonlinear optical coefficients from the spectroscopic data.

Figure 4. Imaginary part of the SHG coefficient for four III-V compounds.

In order to test the description of nonlinear coefficients within the independent-particle approximation we compare the SHG coefficient (cf. expression (2)) with results of measurements in Fig.5 for several III-V compounds crystallizing in zinc-blende structure. The calculations are performed using 505 \( k \) points and 24 bands. The fact that the SHG coefficient (expression (2)) is related to three optical transitions has remarkable consequences. The simplified picture of superposing two combined densities of states at \( 2\omega \) and \( \omega \) according to the contributions to Eq. (2) is not valid. First, the \( 2\omega \)-term dominates the SHG coefficient. Second, the spectral behaviour of \( \text{Im} \chi^{(2)}_{xyz}(\omega) \) (Fig.4) can be explained by the two prominent categories of optical transitions, \( E_1 \) and \( E_2 \), in the linear absorption of, at least, the III-V compounds. However, matrix element effects are much stronger in the nonlinear case. In comparison to the linear absorption the peaks are sharpened and their sign could be negative in accordance to the prefactor. The sign of the imaginary part of the product of three momentum matrix elements is not fixed. The spectra for the four III-V compounds are rather similar. They look like the superposition of the spectra of more or less four pronounced oscillators with resonance frequencies close to the \( E_1 \)- and \( E_2 \)-like structures appearing in the \( 2\omega \) and \( \omega \) - terms of the imaginary parts.

Figure 5. Module of the SHG susceptibility for four III-V compounds versus photon energy. Calculation: solid line, measurements: GaP: dots [36], GaAs and InAs: open circles [37], dotted line [38], dashed line [39], squares [40].

The comparison of the module of the total SHG coefficients with experimental data [36-40] for GaP, GaAs, and InAs (Fig.5) indicates a reasonable description of the SHG spectra by the presented theory. This holds especially for the order of magnitude and the principal spectral behaviour. Consequently, we conclude that the presented theoretical description of the SHG coefficient should allow the prediction of the second-harmonic generation for new materials.

As an example of such a prediction the SHG coefficients of the most important SiC polytypes are given in Table I for incident light with a frequency small compared to the energy gap. The two independent tensor components zzz and zxx are listed for \( \omega = 0 \). The three sets of theoretical values agree with respect to the absolute value, the sign and the trend with the polytype. This holds in particular for the comparison of our data and those of Rashkeev et al. [41], which also nearly agree with respect to the absolute values, though completely different basis sets have been used for the single-particle wave functions. The small deviations to the results of Chen et al. [42] may be traced back to their trial to consider local-field effects. The inclusion of the QP shifts of the bands gives rise to a reduction of the coefficients independent of the quality of the description of the QP effects. The larger reduction in our full calculation of the XC self-energy in GW approximation is mainly due to the larger gap openings (average value:
1.68 eV) compared to the scissors operators used in Ref. 41 (Δ = 1 eV) and Ref. 42 (Δ = 1.04...1.27 eV). The reduction of the static SHG coefficients itself is mainly related to the energy denominators in expression (2).

Table I. Static SHG coefficient (in pm/V) of independent tensor components for several polytypes. Values are calculated within DFT-LDA as well as including QP corrections within a GW treatment or a scissors operators approach. The results are compared with other theoretical (Rashkeev et al. [41], Chen et al. [42]) and experimental (Lundquist et al. [43], Niedermeyer et al. [44]) data.

<table>
<thead>
<tr>
<th>Polytype</th>
<th>2H</th>
<th>4H</th>
<th>6H</th>
<th>3C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hexagonality [%]</td>
<td>100</td>
<td>50</td>
<td>33</td>
<td>0</td>
</tr>
<tr>
<td>Component</td>
<td>χ(zzz)</td>
<td>χ(zzx)</td>
<td>χ(zzz)</td>
<td>χ(zzx)</td>
</tr>
<tr>
<td>this work</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFT-LDA</td>
<td>4.5</td>
<td>-1.7</td>
<td>15.5</td>
<td>-9.7</td>
</tr>
<tr>
<td>QP (GW)</td>
<td>2.3</td>
<td>-3.7</td>
<td>9.4</td>
<td>-5.9</td>
</tr>
<tr>
<td>Rashkeev et al.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFT-LDA</td>
<td>3.6</td>
<td>-6.1</td>
<td>14.5</td>
<td>-8.9</td>
</tr>
<tr>
<td>QP (scissors)</td>
<td>2.5</td>
<td>-3.9</td>
<td>9.2</td>
<td>-5.7</td>
</tr>
<tr>
<td>Chen et al.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFT-LDA</td>
<td>8.6</td>
<td>-13.2</td>
<td>23.3</td>
<td>-14.8</td>
</tr>
<tr>
<td>QP (scissors)</td>
<td>5.8</td>
<td>-8.8</td>
<td>15.6</td>
<td>-10.0</td>
</tr>
<tr>
<td>Lundquist et al.</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Niedermeyer et al.</td>
<td>-</td>
<td>-</td>
<td>18±8</td>
<td>-4±2</td>
</tr>
</tbody>
</table>

The comparison with recent experimental values is also rather successful for the static SHG coefficients of the three polytypes 3C, 6H, and 4H [43,44]. The agreement of the absolute values and the signs is reasonable considering the experimental and theoretical uncertainties. This holds in particular for the zzz tensor components. The zzz tensor components seem to be overestimated by the procedure of data extraction. This is already obvious for the cubic SiC polytype, where the ratio χ(zzz)/(zzx) = -2 should be exactly fulfilled. Theoretically one can show that the absolute value of this ratio should decrease with increasing hexagonality of the polytype.

C. Reflectance anisotropy spectra

Since in the case of cubic crystals the only origin of an reflectance anisotropy is the surface region, RA spectra can be used to clarify the atomic structure and the stoichiometry of a surface. We probed a variety of structural models proposed for the InP(001)2×4 surface [16,20]. It turns out that dimer reconstruction models are energetically most favoured. For In-rich surfaces we predict the formation of mixed In-P dimers on top of an In-terminated surface (Fig.6). Under less In-rich conditions the structure gives the lowest total energy. In the case of intermediate preparation conditions the top-P dimer and the α structure (Fig.6) cannot be excluded. Meanwhile, it has been found that a modified β2 structure with only one top -P dimer, the so-called δ structure [45], is more favourable from the energetical point of view [46].
In Fig 7 we show the RA versus photon energy calculated according to expression (3) for the four favourable surface geometries. The explicit calculations of the slab polarizabilities are performed using a recently developed real-space multigrid method [47]. This approach provides for effective convergence acceleration and preconditioning on all length scales and, hence, reduces the numerical effort. Furthermore, it allows for an efficient parallelization and is thus particularly suitable for large surface reconstructions as studied here.

In Fig 7 a negative anisotropy appears around 1.7 eV for the mixed-dimer model, which contains 6 In-In surface bonds along [110]. This anisotropy is reduced in strength for the top-P-dimer model with 4 In-In surface bonds and is even weaker in case of the α structure with 2 In-In bonds. This feature is correlated to the existence of cation-cation bonds at the surface and caused by transitions between σ-like In-In bonding states and empty dangling bonds localized at the surface cations. The RA spectra show also a dependence on structural details for higher energies: A relatively broad positive anisotropy between about 2 and 4 eV appears and the “three-buckle” shape in the high-energy region flattens for the β2 (and also δ, not shown) geometry with 3 P-P dimers parallel to [110]. This is due to transitions between P-P dimer states and surface resonances. The structure, with 2P-P dimers, gives rise to a very similar line shape. Its spectrum, however, is somewhat downshifted. All structures exhibit a derivative-like feature around 2.5 eV, caused by transitions between surface-modified bulk states. Its energetical position coincides with the calculated position of the $E_1$ peak. The described evolution of the spectra parallels closely the experiments [48,49], which observe similar trends in the RA of InP(001)2×4 surfaces exposed to varying P partial pressures or different temperatures. This supports the view that the corresponding “fingerprints” in the measured spectra are caused by the dimerized In and P atoms.

More in detail, the evolution of the RA spectra with annealing temperature between 470 and 590 °C [48] clearly exhibits a similar variation as the spectra calculated for the mixed-dimer model 2 and the structure. Therefore, we identify the mixed-dimer model (β2 structure) with the stable 2×4 reconstruction at In-rich (less In-rich) preparation conditions of the InP(001) surface.

Figure 7. Calculated RA spectra of favourable InP surface structures.

IV Summary

In conclusion, we have shown that in principle it is possible to calculate optical spectra of semiconductor structures without taking into account any experimentally determined parameter. The resulting agreement between ab initio results and experimental data is reasonable considering a wide spectral region. For structures like two-atomic crystals, natural superlattices with layer thicknesses up to 12 atomic layers, and surfaces of polar semiconductors we found that a reliable description of their linear optical spectra over
a wide frequency range is already possible within the independent-particle approximation. In certain cases the optical transition energies have to be enlarged by quasiparticle effects. However, for the correct description of the fine structure of the spectra, excitonic and local-field effects have to be considered as well. The independent-quasiparticle approach also seems to be a reasonable starting point for \textit{ab initio} calculations of nonlinear optical coefficients. The spectral behaviour and the absolute values of SHG coefficients can be predicted with a reliable accuracy. Only the numerical efforts concerning number of \textit{k} points and bands have to be enforced. Similar experiences have been made with a parameter-free description of the spectral dependence of the surface anisotropy. However, the accuracy of the calculations is already high enough to predict spectra for certain surface reconstructions and, hence, identify the real surface geometries and stoichiometries.

\section*{Acknowledgement}

This work was financially supported by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 196, Project No. A8).

\section*{References}


