PROC. 15TH INT. CONF. PHYSICS OF SEMICONDUCTORS, KYOTO, 1980
J. PHYS. SOC. JAPAN **49** (1980) SUPPL. A p. 69–76

MANY-BODY EFFECTS IN THE OPTICAL SPECTRA OF SEMICONDUCTORS

L. J. Sham

Department of Physics
University of California San Diego
La Jolla, California 92093
U.S.A.

The optical spectrum of many a semiconductor is affected by the electron-hole interaction not only in the exciton region near the energy gap but also throughout the continuum, particularly the strength of the E_1 and E_2 peaks. The local-field effect is shown to worsen the discrepancy between the non-interacting approximation and experiment. The electron-hole attraction is necessary to account for the experiment. Both first principle calculations using plane wave and LCAO methods and simple models are discussed.

I. Introduction

The primary process in the optical absorption in a semiconductor is the excitation of an electron into a conduction band, leaving behind a hole in a valence band. The analysis of the absorption spectrum in terms of the electron and hole energies and their wave functions [1,2], we shall call the non-interacting approximation. The Coulomb interaction between electrons modifies the simple absorption process by a non-interacting electron-hole pair. The modification of the one-electron band structure by the many-body effects shall not concern us here. Our main interest is in the additional processes in the optical absorption caused by the electron interaction.

Phillips [1] has given a very prescient description of the physics of the many-body effects on the absorption spectrum. The attraction of the excited electron-hole pair to form excitons in the gap is well known. More important to us is the excitonic effect in the main body of the absorption spectrum, giving rise to what Phillips called the saddle-point exciton. This continuum exciton has been studied in the effective mass approximation by Velicky and Sak [3] and by Kane [4] and in the contact interaction approximation by Velicky and Sak [3], by Hermanson [5] and by Toyozawa et al. [6]. These works have been critically reviewed by Hanke [7]. Only more recent work will be discussed here.

Another many-body effect on the optical spectrum is the local-field effect. If the Lorentz-Lorenz relation is applicable, it certainly affects the spectrum greatly [8]. In a semiconductor, while the electron-hole pair does not act as such a localized dipole as assumed in the Lorentz-Lorenz relation, the pair can, nonetheless, create secondary electric field to further excite other electron-hole pairs. Recent calculations of this effect will also be discussed.

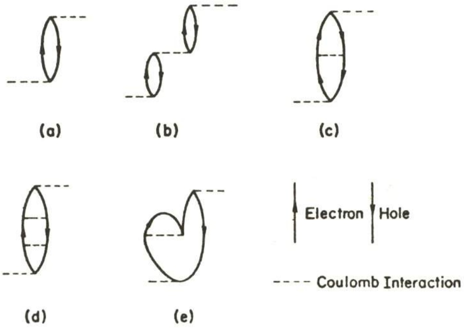

Fig. 1. Polarization diagrams

Fig. (1) defines more precisely the effects discussed above. Fig. (1a) is the polarization by the external field of a non-interacting electron-hole pair. Fig. (1b) is an example of the local-field term. The sum of all such bubbles constitutes the random phase approximation. Fig. (1c) and (1d) depict the electron-hole attraction or the exciton effect. Fig. (1e) illustrates the one-electron energy correction, of no concern here.

II. Optical Spectrum of Non-interacting Electron-hole Pairs

The solid line in Fig. (2) represents the experimental optical

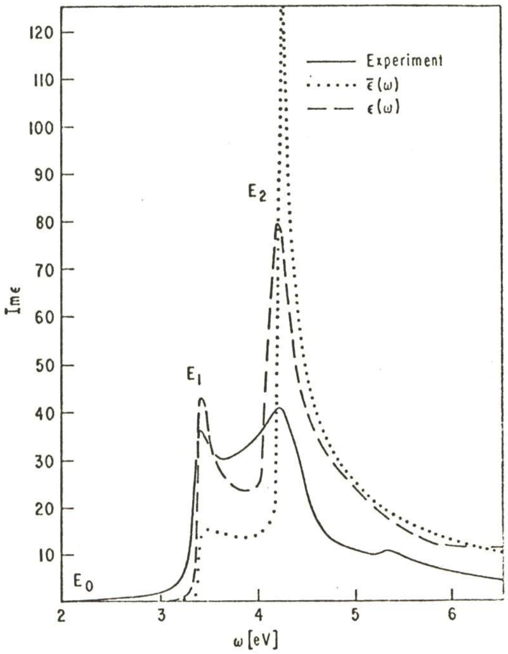

Fig. 2 Absorption spectrum Im ε of silicon deduced from reflectivity and calculated in a simple model without interaction, $\bar{\varepsilon}$, and with interaction ε

absorption spectrum of silicon [9]. The three prominent features are labeled E_0, E_1 and E_2. These features are typical of a wide class of semiconductors of group IV elements and of III-V and II-VI compounds. According to the analysis in the non-interacting approximation [10], the E_0 peak is due to the electronic transition across the direct band gap. (See Fig. (3)). The E_1 peak is due to transitions from the neighborhood of the top valence band to the nearly parallel lowest conduction band in the (111) direction, i.e., between the nearby parallel Λ_{3V} and Λ_{1C} bands. The E_2 peak is due to transitions in a rather large neighborhood of the X point in the Brillouin zone, (100) $2\pi/a$, between X_{4V} and X_{1C} with bands nearly parallel in two directions.

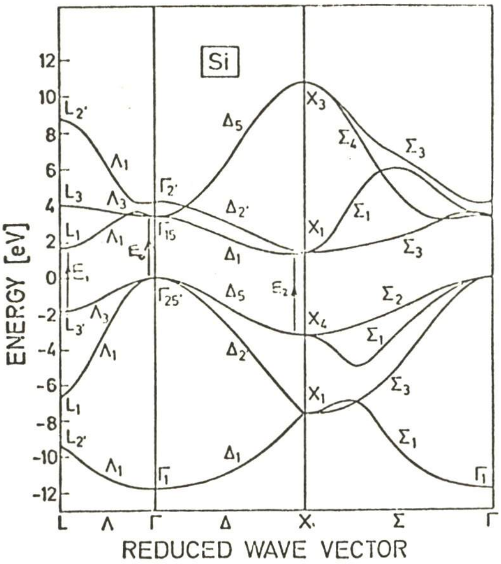

Fig. 3 Silicon band structure, taken from [23]

The dotted curve in Fig. (2) represents a calculation, with details given in Sec. IV, which takes the above description of the E_1 and E_2 peak literally in a simple model. Compared to the measured spectrum, the strength of the E_1 peak is underestimated and that of the E_2 peak is overestimated. A proper calculation of the absorption by one-electron transitions in Si [11] certainly improves the agreement but the qualitative features of a lower E_1 peak and a higher E_2 peak remain in the accurate calculation. This is also true, though with varying extent, for other semiconductors. Some examples are Ge [12,13], GaAs [14,15], InAs, InSb [15], ZnS, ZnSc [16,14]. In some of these calculations the electron energies have been fit to the peak positions, in the pseudopotential frame [14] or in the k·p framework [15], and yet the strength of the peaks has the same feature. One exception is a Fourier sum interpolation of the band structure [17], which reproduces the optical spectrum of Si and Ge exceptionally well. It appears, however, that the over-all band structure, and, hence, the momentum matrix elements, are very different from the other works [18].

The absorption spectrum of diamond [19] is rather different from those of the semiconductors described above. It has a main peak with some weaker structures on the low energy side of the peak. Non-interacting approximation with some band structures [20] yields rather good agreement with experiment [21,22].

III. The Local-field and Excitonic Effects

The sum of all interaction processes which correlates the electron-hole pair illustrated in Fig. (1) will be denoted by the two-particle Green's function S which satisfies an integral equation [23]

$$S = S_o + S_o I S \qquad (1)$$

with S_o being the uncorrelated electron-hole pair and I consisting of two parts: I_a which is the electron-hole attraction, suitably screened and I_b which is the bare Coulomb interaction between one electron-hole pair and another. (See Fig. 4).

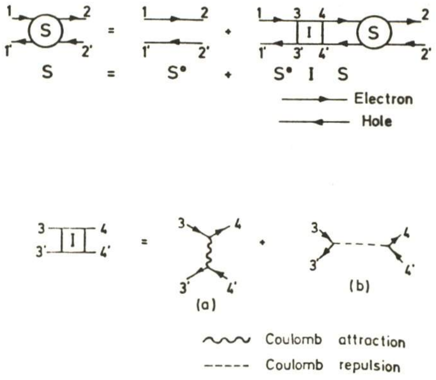

Fig. 4 Integral equation for the two-particle Green's function, eq. (1)

I_b alone in eq. (1) will reproduce the random phase approximation. I_a is the exchange counterpart of I_b, except that the Coulomb interaction is screened in the former. I_a produces the excitonic effect and I_b may be regarded as the exchange term to the electron-hole attraction, which should not be screened [24]. Both points of view have been adopted, the former being prevalent in many-body papers and the latter in exciton papers [25].

Close to the E_o peak, i.e., for the electron transition close to the direct fundamental gap, eq. (1) can be shown to reduce to the effective mass approximation [24], which adequately accounts for the bound exciton lines and continuum exciton effect near E_o [26] with a small exchange correction [25]. The effective mass approximation is, however, invalid around the E_1 or E_2 peak which involves a much larger phase space than the electron-hole pair around E_o which binds into a Wannier exciton. Eq. (1) has to be solved directly.

In the pseudopotential method, the electron wave function is expanded in a series of plane waves and eq. (1) in the RPA, i.e., with only the I_b term in the interaction, is reduced to a matrix equation with the elements ranked by the reciprocal lattice vector. The equation is thus solved for diamond [22] and silicon [11]. With the electron-hole attraction, eq. (1) remains an integral equation, which is yet to be solved for a semiconductor.

In the method of linear combination of local orbitals, the electron wave function is expressed in terms of localized orbitals. Eq. (1), with the interaction including both I_a and I_b, i.e., both the electron-hole attraction and its exchange, is reduced to a matrix equation. Its inversion for diamond [27] and silicon [23] leads to optical spectra including the excitonic and local-field effects.

The local-field effect, as produced by the RPA, i.e., taking only I_b in eq. (1), reduces the intensity of a peak of the optical spectrum in the non-interacting approximation, more severely on the low energy side, resulting in the appearance of a smaller peak with its position shifted to a higher energy. For Si, since the E_1 peak is already too low in the non-interacting approximation, the inclusion of the local-field effect pushes it further away from the measured curve, as shown in Fig. (5), calculated in the LCAO representation [23]. The RPA in the plane wave representation [11] gives qualitatively the same modification, although the depression of the E_1 peak is smaller. The E_2 peak which is too high in the non-interacting approximation is lowered too much by the local-field effect in com-

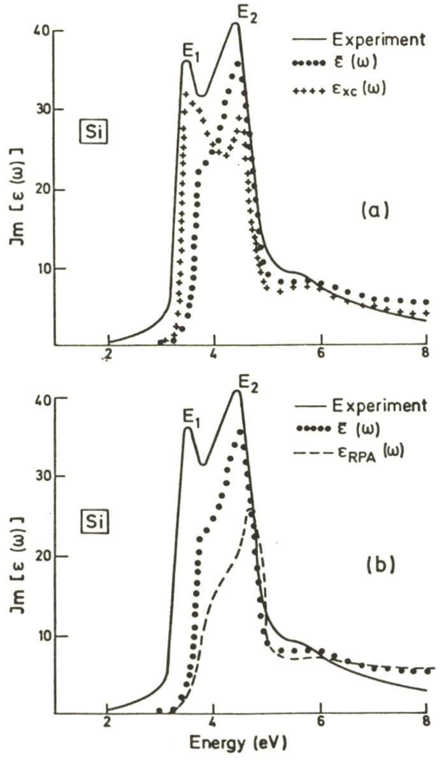

Fig. 5 Absorption spectrum of Si, taken from [23]

parison with experiment. In diamond, the main peak is similarly reduced in strength and shifted to higher energy in both methods of calculation, though again the change is smaller in the plane wave method [22] than in the LCAO [27]. The quantitative difference between the two representations may be attributed to that LCAO yields a more localized wave function for the conduction and valence electron and hence a larger local-field effect than the plane wave representation.

The excitonic effect affects the optical spectrum in just the opposite manner. This effect has only been calculated in LCAO [23,27]. Fig. (5) shows that in Si, it increases the intensity of the E_1 peak, particularly so in the low energy side such that the peak appears shifted to lower energy by about 0.2 eV. Agreement with experiment is thereby improved. The lifting of the E_2 peak is not sufficient to compensate the lowering due to the local-field effect. One might venture a guess that the present LCAO calculation yields too small an E_2 peak in the non-interacting approximation in comparison with the plane wave representation which is larger than the measured E_2 peak. If the larger E_2 peak is taken as the true non-interacting result, then the net result of exciton and local-field effects would reduce it to just about the measured value. In diamond, the excitonic effect raises the main peak to about the experimental value.

IV. A Model Dielectric Function

From the calculations of the excitonic and local-field effects for Si and diamond, we would like to draw the same conclusion for other semiconductors. Now a model is discussed which is not only applicable to the wide class of semiconductors discussed in Sec. II, but also provides a simple physical picture for the many-body effects on the optical spectrum.

In the non-interacting approximation, the dielectric function as a function of frequency ω of the exciting field has the form

$$\bar{\varepsilon}(\omega+io) = 1 - \frac{\omega_p^2}{\omega^2} + \frac{\omega_p^2}{\omega^2}\left\{\Pi(\omega+io) + \Pi(-\omega-io)\right\} \tag{2}$$

where ω_p is the plasma frequency and the sum of the two Π functions in the curly brackets is the current response function, rendered dimensionless, for uncorrelated electron-hole pair excitations. Now we approximate the interaction I in eq. (1) by a separable form of a dimensionless coupling constant V times a current vertex function of the relative momentum of the incoming electron-hole pair and the same function of the relative momentum of the outgoing electron-hole pair. Then eq. (1) may be solved to yield the dielectric function including the interaction effects:

$$\varepsilon(\omega+io) = 1 - \frac{\omega_p^2}{\omega^2} + \frac{\omega_p^2}{\omega^2}\left\{\frac{\Pi(\omega+io)}{1-V\Pi(\omega+io)} + \frac{\Pi(-\omega-io)}{1-V\Pi(-\omega-io)}\right\} \tag{3}$$

Near E_1 peak, ReΠ is positive as can be seen by either the Kramer-Kronig relation or the Reε plot deduced from reflectivity [9]. In RPA, V is negative and the denominator in eq. (3) reduces the strength of the non-interacting peak. V is positive for the electron-hole attraction and, hence, the exciton effect raises the E_1 peak. Near E_2 peak, ReΠ is negative [9]. Even though the net V is positive because the excitonic interaction dominates, the denominator in eq. (3) is larger than unity and, hence, reduces the E_2 peak strength. From the size of Reε and hence ReΠ, which is larger for

Si than for Ge and GaAs [9], one can deduce that the many-body corrections for E_1 and E_2 peaks are larger in Si than in Ge and GaAs.

Cardona's model analysis discussed in Sec. II enables us to evaluate $\Pi(\omega+io)$, which consists of two parts. The E_1 peak part comes from the k·p Hamiltonian around L_{1C} and $L_{3'V}$. The E_2 peak part comes from the two plane-wave mixture connected by the reciprocal lattice vector (220) near the Jones zone surface centered at the X point [28]. The fraction of the Jones zone surface which contributes is determined by the f sum rule. The dotted curve in Fig. (2) is calculated in this approximation for Si. The E_2 peak, which is grossly overestimated, can clearly be improved.

With interaction effects included, each part is screened as in eq. (3). The interaction constants V_1 and V_2 are estimated from phase space averages of the statically screened Coulomb interaction. The dashed curve in Fig. (2) represents the inclusion of the excitonic effect. The bound exciton line below the E_1 edge is smeared by the $\Gamma = 0.01$ eV broadening. The higher E_1 peak is only partly due to the smearing of the bound exciton line and is mostly due to the continuum exciton effect above the E_1 edge. The spike below the E_2 edge shows that $\Gamma = 0.01$ eV is insufficient to smear out the bound exciton line. (A 0.1 eV broadening will suffice). The intensity above the E_2 edge is reduced and that below the E_2 edge is increased, in accordance with the more precise calculation.

The results of this simple model for Si are qualitatively valid for the other semiconductors and may be regarded as providing a simple description of the many-body effects on the optical spectrum of the semiconductor in general.

Acknowledgement

This work is supported by the U. S. National Science Foundation. Without the collaboration of Professor W. Hanke over the last seven years, this review article could not have been written. I wish to thank the Aspen Center for Physics where this manuscript was prepared.

References

1) J. C. Phillips: Solid State Physics 18 (1966) 55.
2) F. Bassani and G. P. Parravacini: "Electronic States and Optical Transitions in Solids," ed. by R. A. Bullinger (Pergamon Press 1975).
3) B. Velicky and J. Sak: Phys. Stat. Solids 16 (1966) 147.
4) E. O. Kane: Phys. Rev. 180 (1969) 852.
5) J. Hermanson: Phys. Rev. 166 (1966) 893.
6) Y. Toyozawa, M. Inoue, T. Inui, M. Okazaki and E. Hanamura: J. Phys. Soc. Jpn. 22 (1967) 1337.
7) W. Hanke: Festkörperprobleme XIX (1979) 1.
8) T. K. Bergstresser and G. W. Rubloff: Phys. Rev. Lett. 30 (1973) 794.
9) H. R. Phillip and H. Ehrenreich: Phys. Rev. 129 (1963) 1550.
10) M. Cardona: Proc. Intern. School of Physics "Enrico Fermi" Course LII, Atomic Structure and Properties of Solids, ed. E. Burstein (Academic Press 1972) p.514.
11) S. G. Louie, J. R. Chelikowsky and M. L. Cohen: Phys. Rev. Lett. 34 (1975) 155.
12) W. D. Grobman, D. E. Eastman and J. L. Freeouf: Phys. Rev. B12 (1975) 4405.
13) F. Herman, R. L. Kortum, C. D. Kuglin and J. L. Shay: "II-VI Compounds," ed. D. G. Thomas, (W. A. Benjamin 1967) p. 503.
14) J. P. Walter and M. L. Cohen: Phys. Rev. 183 (1969) 763.
15) C. W. Higginbotham, F. H. Pollack and M. Cardona: Proc. IX Intern. Conf. Semiconductors (Leningrad 1968) p. 57.
16) J. L. Freeouf: Phys. Rev. B7 (1973) 3810.
17) G. Dresselhaus and M. S. Dresselhaus: Inter. J. Quan. Chem. 1s (1967) 595; Phys. Rev. 160 (1967) 649.
18) K. C. Pandey and J. C. Phillips: Phys. Rev. B13 (1976) 750.
19) R. A. Roberts and W. C. Walker: Phys. Rev. 161 (1967) 730.
20) G. S. Painter, D. E. Ellis and A. R. Lubinsky: Phys. Rev. B4 (1971) 3610.
21) A. R. Lubinsky, D. E. Ellis and G. S. Painter: Phys. Rev. B6 (1972) 3950.
22) J. A. Van Vechten and R. M. Martin: Phys. Rev. Lett. 28 (1972) 446.
23) W. Hanke and L. J. Sham: Phys. Rev. Lett. 43 (1979) 380; Phys. Rev. B21 (1980) 4656.
24) L. J. Sham and T. M. Rice: Phys. Rev. 144 (1966) 708.
25) K. Cho: Proc. Intern. Conf. Semiconductors (Edinburgh 1978) p. 841; Phys. Rev. B14 (1976) 4463.
M. H. Denisov and V. P. Makarov: Phys. Stat. Sol. b56 (1973) 9.
26) R. J. Elliott: Phys. Rev. 108 (1957) 1384.
27) W. Hanke and L. J. Sham: Phys. Rev. B12 (1975) 4501.
28) V. Heine and R. O. Jones: J. Phys. C2 (1969) 719.
29) J. R. Chelikowsky and M. L. Cohen: Phys. Rev. B10 (1974) 5095.