Density-Functional Theory of Excitation Spectra of Semiconductors: Application to Si

C. S. Wang

Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742, and Condensed Matter Physics Branch, Naval Research Laboratory, Washington, D.C. 20375

and

W. E. Pickett

Condensed Matter Physics Branch, Naval Research Laboratory, Washington, D.C. 20375

This Letter presents a general approach to calculation of the quasiparticle excitation energies of semiconductors which includes the energy dependence of the self-energy with a local density-functional approach. Both direct and indirect band gaps in Si are in much better agreement with experiment than are the bands from the ground-state theory. The relatively large corrections arise only after the dielectric screening of the electron gas is modified to account for a gap in the excitation spectrum.

PACS numbers: 71.45.Gm, 71.25.Rk

By demonstrating the one-to-one correspondence between the external potential $v_{ext}(r)$ and the ground-state charge density $n(r)$ in a system of interacting electrons, Hohenberg and Kohn[1] established that all properties of the system may be considered as functionals of n rather than v_{ext}. This result has been utilized primarily in the form of an approximate ground-state energy functional using the local density approximation (LDA). As a by-product of the calculations one-electron-like eigenfunctions and eigenvalues are generated which are used, without formal justification, to interpret single-particle excitations. Although for metals the correspondence between "theory" and experiment is often quite reasonable, LDA typically underestimates band gaps in semiconductors and insulators by 30%–50%. These results have engendered various theoretical attempts to go beyond LDA, such as "self-interaction corrections"[2] and the weighted (non-local) density approximation (WDA).[3] The former approach seems to work well in atoms and in wide-gap (atomiclike) insulators but its adaptation to extended states remains an unsolved problem. Because of numerical complications the WDA approach also has not been given a conclusive test in extended systems (see below). Alternatively, Levine and Louie[4] (LL) have proposed to modify the dielectric function $\epsilon(q, \omega)$ of the homogeneous electron gas to allow for the possibility of a gap in the electronic excitation spectrum. They obtained a one-parameter (energy gap) family of LDA exchange-correlation (XC) potentials μ_{XC}^{LL} and found a small improvement of the band gaps in Si.

Sham and Kohn[5] formulated a density-functional theory of single-particle excitations ["quasiparticles" (QP's)] and suggested a local-density simplification (QPLDA) demonstrably valid in the limit of slowly varying density. At the same time Hedin,[6] in a seminal study of the QP properties of the homogeneous electron gas, argued in favor of the "GW" (Coulomb hole and dynamically screened exchange) approximation for the complicated Coulomb self-energy. The QPLDA has been applied only infrequently in detailed calculations, probably because the corrections are comparatively small in metals.

In this Letter we present a general approach to the calculation of the QP excitation spectrum of semiconductors, with application to Si. Our QPLDA approach is based on the GW approximation and the LL model dielectric function. Our approach has the same objective (the GW QP spectrum) as the recent study of diamond by Strinati, Mattausch, and Hanke.[7] Our theory is suitable for semiconductors, however, while their approach appears to be limited computationally to wide-gap insulators. We find that the calculated QPLDA direct band gaps are widened by 0.5 to 0.8 eV over the LDA values, in close agreement with experiment, and the discrepancy in the indirect gap is decreased by $\frac{2}{3}$. These important corrections result directly from the dielectric screening of the QP's peculiar to systems with an energy gap.

The QP energies are solutions to the following Dyson equation ($\hbar = 1$):

$$[-\nabla^2/2m + V_h(\vec{r})]\psi(\vec{r}) + \int d^3r' \, M(\vec{r}, \vec{r}'; E)\psi(\vec{r}') = E\psi(\vec{r}), \tag{1}$$

where V_h denotes the sum of ionic and Hartree potentials. Sham and Kohn introduced the Wentzel-Kramers-Brillouin–type approximation for the mass operator

$$M(\vec{r}, \vec{r}', E) = \mathfrak{M}_h\big(p(\vec{r}), E - \mu + \mu_h(n(\vec{r})); n(\vec{r})\big)\, \delta(\vec{r} - \vec{r}'), \tag{2}$$

where the chemical potentials, μ and μ_h, were introduced to line up the Fermi energies of the inhomogeneous and the homogeneous system, respectively, and $\mathfrak{M}_h$ is the mass operator of the homogeneous electron gas in momentum space. The local wave vector p is determined by

$$P^2/2m + \mathfrak{M}_h(p, E - \mu + \mu_h(n); n) = E - \mu + \mu_h(n). \tag{3}$$

For a QP at the Fermi level, Eq. (3) is satisfied by $p = k_F$ and hence

$$\mathfrak{M}_h(k_F, \mu_h; n) = \mu_h(n) - k_F^2/2m = \mu_{XC}(n), \tag{4}$$

where the XC part μ_{XC} of the chemical potential is the usual XC potential in LDA. For semiconducting systems we take the chemical potential to lie at midgap for both homogeneous and inhomogeneous systems.

The GW approximation of Hedin for $\mathfrak{M}_h$ is

$$\mathfrak{M}_h(p, E) = \frac{i}{(2\pi)^4} \int \frac{v(p')d^3p'}{\epsilon(p', \omega')}\, G(|\vec{p} - \vec{p}'|, E - \omega')\exp(-i\omega' O^{\dagger})d\omega', \tag{5}$$

where the bare Coulomb interaction $v(\vec{p}')$ is screened by the dielectric function $\epsilon(p', \omega')$. In principal $\mathfrak{M}_h$ should be solved self-consistently because the Green's function G is related to $\mathfrak{M}_h$ by the Dyson equation

$$G(p, E) = [E - E(p) - \mathfrak{M}_h(p, E)]^{-1}. \tag{6}$$

To make the calculation tractable $\mathfrak{M}_h(p, \omega)$ in Eq. (6) is approximated by $\mathfrak{M}_h(k_F, \mu_h) = \mu_{XC}^{LL}$. To model semiconducting behavior we introduce a direct gap E_g in the single-particle spectrum:

$$E(p) = p^2/2m + (E_g/2)\mathrm{sgn}(p - k_F) \tag{7}$$

in Eq. (6) and use the LL model dielectric function $\epsilon^{LL} = \epsilon_1^{LL} + i\epsilon_2^{LL}$:

$$\epsilon_2^{LL}(p, \omega) = \begin{cases} 0, & |\omega| < \lambda\omega_F, \\ \epsilon_2^{RPA}(p, \omega_-), & |\omega| > \lambda\omega_F, \end{cases} \tag{8}$$

where $\omega_- = (\omega^2 - \lambda^2\omega_F^2)^{1/2}\,\mathrm{sgn}\,\omega$, $\omega_F = k_F^2/2m$, and $\epsilon_1(p, \omega)$ can be obtained analytically by Kramers-Kronig transformation; RPA denotes random-phase approximation. The function satisfies the f-sum rules on ϵ and ϵ^{-1} and reduces to the Lindhard function $\epsilon^{RPA}(p, \omega)$ in the metallic limit $\lambda = 0$. Figure 1 illustrates that ϵ^{LL} provides a very realistic model of both the p and ω behavior of screening in Si, for which $\lambda = 0.4$ is chosen[4] for the average valence-electron density ($r_s = 2.0$) to fit the experimental value of $\epsilon(0, 0)$. For E_g in Eq. (7) we use the zone-average direct band gap (3.5 eV) for Si. Integrations in Eq. (5) were performed analogously to the procedure outlined by Hedin.[6]

Since considerable sophistication has gone into determining μ_{XC}^{LL} for the electron gas, we take advantage of Eq. (4) and write

$$\mathfrak{M}_h(p, E; n) = \mu_{XC}^{LL}(n) + [\mathfrak{M}_h(p, E; n) - \mathfrak{M}_h(k_F, \mu_h; n)] = \mu_{XC}^{LL}(n) + \Delta_h(p, E; n) \tag{9}$$

and thereby use our model only for the self-energy Δ_h, which is small compared to μ_{XC}^{LL}. In the present study the imaginary part of Δ_h has been neglected and the charge density is taken from a self-consistent nonlocal pseudopotential[10] calculation. The resulting QPLDA eigenvalues, using plane waves up to 11 Ry in the basis, near the gap at Γ, X, and L are given in Table I. For comparison the corresponding LDA results using Hedin-Lundqvist[15] (HL) and LL exchange-correlation potentials are shown, as well as values inferred from various experiments. The conduction bands are generally improved, often dramatically. The X_4^v valence-band level is, however, shifted somewhat below the accepted experimental value. We have also listed the (directly measured) lowest direct transitions at X and L to illustrate that the QPLDA gives large (0.5–0.8-eV) corrections and essentially correct gaps at these points. Introducing the effective gap into the screening ($\lambda = 0.4$) gives roughly 75% of the cor-
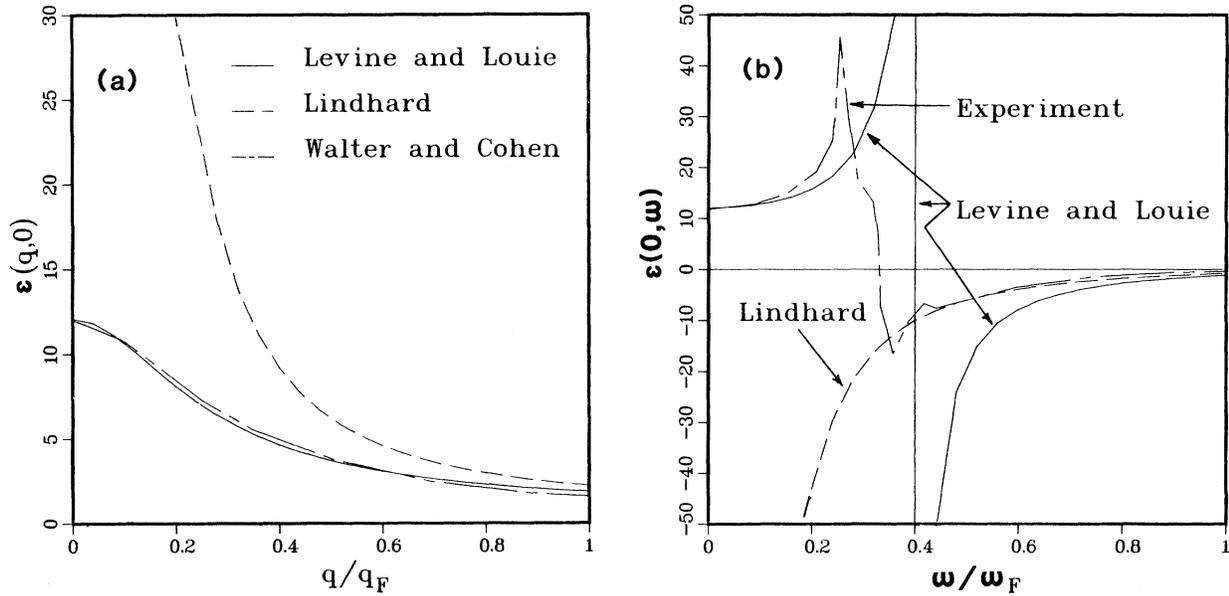
FIG. 1. (a) The wave-vector (q) dependence of the model semiconductor dielectric function $\epsilon^{LL}(q,0)$ of Levine and Louie, compared to the calculated values of Walter and Cohen (Ref. 8) and contrasted with the metallic (Lindhard) counterpart. (b) The real part of $\epsilon^{LL}(0,\omega)$ compared to experiment (Ref. 9) and contrasted with the metallic (Lindhard) counterpart.

rection (see Table I), with the remainder arising from the direct gap E_g in the zero-order QP spectrum [Eq. (7)].

In Fig. 2 we plot the self-energy correction $\Delta E_{n,\vec{k}} = E_{n,\vec{k}}^{QPLDA} - E_{n,\vec{k}}^{LDA}$ vs $E_{n,\vec{k}}^{QPLDA} - \mu$ for several $\vec{k}$ points

TABLE I. Eigenvalues (in electronvolts) for states at Γ, X, and L with energies near the gap. The LDA eigenvalues result from our self-consistent pseudopotential calculations for Hedin-Lundqvist (HL) and Levine-Louie (LL) XC potentials; QPLDA results utilized the charge density using the LL XC potential, except for $\lambda = 0$, $E_g = 0$ which used a HL charge density. Experimental values were taken from Stöhr and Bross (Ref. 11) except where noted.

| | LDA | | QPLDA | | | |
	HL	LL	$\lambda=0.0$ $E_g=0$	$\lambda=0.4$ $E_g=0$	$\lambda=0.4$ $E_g=3.5$ eV	Expt.
$\Gamma_{25'}^{v}$	0.00	0.00	0.00	0.00	0.00	0.00
Γ_{15}^{c}	2.55	2.63	2.57	2.81	3.07	3.3, 3.4
$\Gamma_{2'}^{c}$	3.55	3.59	3.59	3.99	4.10	4.15
X_4^{v}	-2.88	-2.84	-2.89	-3.12	-3.17	-2.9
X_1^{c}	0.65	0.80	0.66	0.92	1.07	1.25^{a}
$L_{3'}^{v}$	-1.22	-1.20	-1.23	-1.37	-1.37	-1.2 ± 0.2
L_1^{c}	1.58	1.66	1.60	1.73	2.01	2.1^{b}
L_3^{c}	3.37	3.47	3.38	3.79	4.00	3.9
Δ_{min}	0.56	0.70	0.71	0.79	0.93	1.17
$L_{3'}^{v}\rightarrow L_1^{c}$	2.80	2.86	2.83	3.24	3.38	3.4^{c}
$L_{3'}^{v}\rightarrow L_3^{c}$	4.59	4.67	4.61	5.20	5.37	5.4, 5.5
$X_4^{v}\rightarrow X_1^{c}$	3.53	3.64	3.55	4.11	4.24	4.25^{c}

[a] Estimated from the low-temperature value of the indirect gap, 1.17 eV (Ref. 12).
[b] Ref. 13.
[c] Ref. 14.
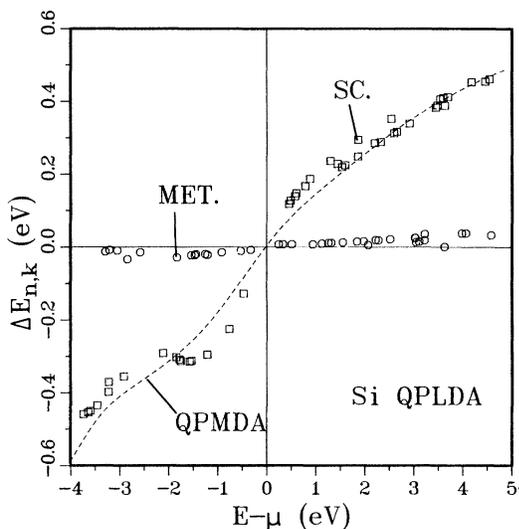
FIG. 2. Scatter plot of the QPLDA correction $\Delta E_{n,k}$ vs $E - \mu$, for both semiconducting (SC., square symbols) and metallic (MET., round symbols) self-energies. Also shown is the corresponding curve (dashed line) in the mean-density approximation (QPMDA).

In summary, we have used a realistic model self-energy appropriate to a semiconductor to demonstrate that including dynamical correlation —*explicit excitation-dependent corrections*—substantially improves the calculated band structure of Si. The theory is easily implementable in existing band-structure calculations and should be useful in understanding the fundamental excitations in many other semiconductors.

for both semiconducting (square symbols) and metallic (round symbols) self-energies. Approximately linear behavior in the region $|E - \mu|$ $\lesssim$ gap suggests that a primarily energy-dependent correction is occurring. To investigate this notion we have examined a mean-density approximation (QPMDA) consisting of evaluating Δ at the density $(r_s^{QPMDA} = 2.06)$ which produces a plasmon in $1/\epsilon$ at the observed energy. The resulting self-energy (see Fig. 2) represents an excitation-energy–dependent constant potential and produced eigenvalues which differ from those of the QPLDA by less than 0.1 eV. Therefore the corrections we find are *dominated by the energy dependence* of the QPLDA. This suggests to us that nonlocal-density (i.e., energy-independent) generalizations alone will not succeed in giving substantially improved (over LDA) QP bands.

This conclusion may seem inconsistent with the WDA calculation of Kerker,[16] who found substantial corrections in Si. To make the numerical calculation tractable, however, Kerker introduced an *Ansatz* which increases the Kohn-Sham exchange $(\alpha = \frac{2}{3})$ to Slater exchange $(\alpha = 1)$ for a homogeneous system and correspondingly increases the correlation potential. This increases the gap in Si but since it is incorrect in the homogeneous limit, it does not appear to represent a fundamental advance in the theory.

[1]P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[2]See J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981), and references therein; R. A. Heaton, J. G. Harrison, and C. C. Lin, Solid State Commun. **41**, 827 (1982).

[3]O. Gunnarsson, M. Jonson, and B. I. Lundqvist, Phys. Rev. B **20**, 3136 (1979); J. A. Alonso and L. A. Girifalco, Phys. Rev. B **17**, 3735 (1978).

[4]Z. H. Levine and S. G. Louie, Phys. Rev. B **25**, 6310 (1982).

[5]L. J. Sham and W. Kohn, Phys. Rev. **145**, 561 (1966).

[6]L. Hedin, Phys. Rev. **139**, A796 (1965).

[7]G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. B **25**, 2867 (1982).

[8]J. P. Walter and M. L. Cohen, Phys. Rev. B **5**, 3101 (1972).

[9]H. R. Phillip and H. Ehrenreich, Phys. Rev. **129**, 1550 (1963).

[10]W. E. Pickett, unpublished, using the ionic pseudopotential of G. Bachelet, M. Schlüter, and D. R. Hamann, Phys. Rev. B **26**, 4199 (1982).

[11]H. Stöhr and H. Bross, Phys. Status Solidi (b) **90**, 497 (1978).

[12]K. L. Shaklee and R. E. Nahory, Phys. Rev. Lett. **24**, 942 (1970).

[13]R. Hulthén and N. G. Nilsson, Solid State Commun. **18**, 1341 (1976).

[14]D. E. Aspnes and A. A. Studna, Phys. Rev. B **27**, 985 (1983).

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

[16]G. P. Kerker, Phys. Rev. B **24**, 3468 (1981).