Quasiparticle Energies in Semiconductors: Self-Energy Correction to the Local-Density Approximation

F. Gygi[(1),(a)] and A. Baldereschi[(1,2)]

[(1)]*Institut de Physique Appliquée, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland*

[(2)]*Dipartimento di Fisica Teorica e Gruppo Nazionale di Struttura della Materia, Consiglio Nazionale delle Ricerche, Università di Trieste, Trieste, Italy*

(Received 9 May 1988; revised manuscript received 22 December 1988)

A model self-energy correction to the local-density approximation is derived from the *GW* approximation of the self-energy operator. Excitation energies calculated in diamond, Si, Ge, GaAs, and AlAs compare favorably with those obtained with the full *GW* approximation. Results for GaP are close to experimental data. It is shown that the model consists of a "scissor" operator and additional nonrigid corrections.

PACS numbers: 71.45.Gm, 71.10.+x, 71.25.Tn

The local-density approximation (LDA) is a powerful approach for the description of ground-state properties of solids.[1,2] The eigenvalues of the Kohn-Sham equations, however, do not in general compare well with experimental excitation energies; the fundamental gap of nonmetals being typically underestimated by 50% to 100%. Excitation energies of many-electron systems can be derived in the context of many-body perturbation theory from the one-particle Green's function, which can be obtained by solving Dyson's equation. In this spirit, Hybertsen and Louie[3] and Godby, Schlüter, and Sham[4] have recently calculated accurate quasiparticle energies in semiconductors and insulators using the *GW* approximation of Hedin.[5] These quasiparticle energies were obtained by subtracting from the Kohn-Sham eigenvalues the LDA self-energy and replacing it with the expectation value of the *GW* self-energy operator. Considering the complexity of a full *GW* calculation and the simplicity of the LDA, it seems desirable (i) to understand the origin of the difference between the LDA and *GW* approximations in nonmetals, (ii) to derive the correction to be added to the LDA in order to recover the *GW* results, and (iii) to examine whether such a correction can be expressed in a simple form.

In this work, we analyze the *GW* self-energy operator in a semiconductor and separate it into a short-range part, which, following Kohn and Sham[2] can be approximated by a local exchange-correlation potential, and a correction which results from incomplete screening of the Coulomb interaction. This correction is calculated in diamond, Si, Ge, GaAs, and AlAs. When added to LDA eigenvalues, it reproduces the essential features of the quasiparticle energies obtained in the *GW* approximation.[3,4] We also show that the self-energy correction can be decomposed into a "scissor" operator[6] and additional terms giving rise to nonrigid shifts of the eigenvalues.

We start from the *GW* expression of the self-energy operator[5]

$$\Sigma^{GW}(\mathbf{r},\mathbf{r}';E)=i\int G(\mathbf{r},\mathbf{r}';E+E')W(\mathbf{r},\mathbf{r}';E')dE', \quad (1)$$

where $G(\mathbf{r},\mathbf{r}';E)$ is the one-particle Green's function and $W(\mathbf{r},\mathbf{r}';E)$ is the effective interaction potential which includes the effects of local fields and dynamical correlations. In metals, $W(\mathbf{r},\mathbf{r}';E)$ decreases rapidly as $|\mathbf{r}-\mathbf{r}'|>1/k_{\mathrm{TF}}$, where k_{TF} is the Thomas-Fermi wave vector, whereas in semiconductors it decreases as $1/\epsilon_0|\mathbf{r}-\mathbf{r}'|$ for large $|\mathbf{r}-\mathbf{r}'|$ because of the finite value of the high-frequency dielectric constant ϵ_0. In view of this qualitative difference, we write

$$W(\mathbf{r},\mathbf{r}';E)=W^{\mathrm{IEG}}(\mathbf{r},\mathbf{r}';E)+\delta W(\mathbf{r},\mathbf{r}';E),$$

where W^{IEG} is the short-range effective interaction potential of a metallic inhomogeneous electron gas and δW has the same long-range behavior as W, so that

$$\Sigma^{GW}(\mathbf{r},\mathbf{r}';E)=i\int G(\mathbf{r},\mathbf{r}';E+E')W^{\mathrm{IEG}}(\mathbf{r},\mathbf{r}';E')dE' + i\int G(\mathbf{r},\mathbf{r}';E+E')\delta W(\mathbf{r},\mathbf{r}';E')dE'. \quad (2)$$

The first term on the right-hand side (rhs) of Eq. (2) is the self-energy operator of a metallic inhomogeneous electron gas which has been analyzed by Sham and Kohn.[7] They showed that it is short ranged and depends only on the charge density in the vicinity of $\mathbf{r}$ and $\mathbf{r}'$, and can therefore be approximated by a local, energy-dependent potential. Accordingly,

$$\Sigma^{GW}(\mathbf{r},\mathbf{r}';E)=\mu_{\mathrm{XC}}(\mathbf{r},E)\delta(\mathbf{r}-\mathbf{r}') + i\int G(\mathbf{r},\mathbf{r}';E+E')\delta W(\mathbf{r},\mathbf{r}';E')dE'. \quad (3)$$

Calculations of quasiparticle energies including the first term on the rhs of Eq. (3) have been carried out by Wang and Pickett,[8] who found that the inclusion of the energy dependence of the exchange-correlation potential only leads to small improvements over the eigenvalues obtained with an energy-independent LDA potential. In what follows, we will assume that the energy dependence of the local exchange-correlation potential can be neglected for states close to the Fermi level, so that $\mu_{\mathrm{XC}}(\mathbf{r},E)$ can be replaced by its value at the Fermi level,

© 1989 The American Physical Society

$\mu_{XC}^{LDA}(r)$.[7]

We now derive an approximation of the last term on the rhs of Eq. (2), which we denote by $\delta\Sigma$. This term contains a long-range part because of the slow decrease of δW, and cannot be reduced to a local operator. Little is known about $\delta W(\mathbf{r},\mathbf{r}';E)$ apart from its asymptotic behavior as $\mathbf{r}\rightarrow\mathbf{r}'$ and as $|\mathbf{r}-\mathbf{r}'|\rightarrow\infty$. First, we expect δW to be small in the limit $\mathbf{r}\rightarrow\mathbf{r}'$, since the short-distance screening properties of a semiconductor are the same as those of a metal having a similar density distribution. On the other hand, $\delta W\simeq 1/\epsilon_0|\mathbf{r}-\mathbf{r}'|$ as $|\mathbf{r}-\mathbf{r}'|\rightarrow\infty$. We further make the assumption that $\delta W(\mathbf{r},\mathbf{r}';E)$ only depends on $|\mathbf{r}-\mathbf{r}'|$. This is strictly valid only as $|\mathbf{r}-\mathbf{r}'|\rightarrow\infty$, but calculations show that in nonmetals, local-field effects become negligible as $|\mathbf{r}-\mathbf{r}'|$ exceeds the interatomic distance. The energy dependence of δW is assumed to be dominated by a plasmon-pole structure around the plasmon frequency ω_p. Since we consider excitation energies much smaller than ω_p, $\delta\Sigma$ reduces to a static Coulomb-hole plus screened-exchange (COHSEX) contribution. It must be noted that this last assumption on $\delta\Sigma$ is much less restrictive than making the same approximation for the *whole* self-energy operator (COHSEX approximation). The correction is therefore

$$\delta\Sigma(\mathbf{r},\mathbf{r}')=-\rho(\mathbf{r},\mathbf{r}')\delta W(|\mathbf{r}-\mathbf{r}'|)+\delta E_{\mathrm{CH}}\,, \tag{4}$$

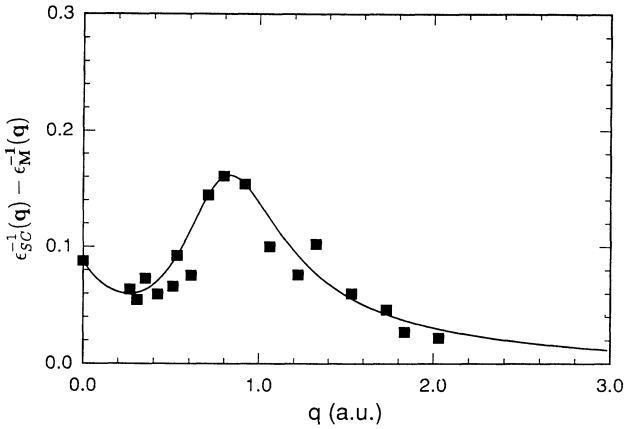

FIG. 1. Difference $\epsilon_{\mathrm{SC}}^{-1}(\mathbf{q},\mathbf{q};0)-\epsilon_M^{-1}(q,0)$ used in the definition of $\delta W(q)$. The solid line shows the numerical fit used in the calculations.

where $\rho(\mathbf{r},\mathbf{r}')$ is the Dirac density matrix and

$$\delta E_{\mathrm{CH}}=\frac{\Omega}{2(2\pi)^3}\int d^3q\,\delta W(\mathbf{q}) \tag{5}$$

is the Coulomb-hole contribution, which in this case shifts all quasiparticle energies by the same amount without affecting the excitation energies.

The expectation value of the correction in a Bloch state $n,\mathbf{k}$ is

$$\delta\Sigma_{n,\mathbf{k}}=\langle n,\mathbf{k}|\,\delta\Sigma\,|n,\mathbf{k}\rangle=-\sum_{m,\mathbf{q}}\sum_{\mathbf{G}}|\langle n,\mathbf{k}|\,e^{i(\mathbf{k}-\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}\,|m,\mathbf{q}\rangle|^2\delta W(\mathbf{k}-\mathbf{q}+\mathbf{G})+\delta E_{\mathrm{CH}}\,, \tag{6}$$

where $m,\mathbf{q}$ denote occupied states, $\mathbf{G}$ is a reciprocal-lattice vector, and

$$\delta W(q)=\frac{4\pi e^2}{\Omega q^2}[\epsilon_{\mathrm{SC}}^{-1}(\mathbf{q},\mathbf{q};\omega=0)-e_M^{-1}(\mathbf{q},\omega=0)]\,, \tag{7}$$

where $\epsilon_{\mathrm{SC}}^{-1}(\mathbf{q},\mathbf{q};0)$ is the diagonal part of the inverse dielectric matrix of the semiconductor calculated in the random-phase approximation, and $\epsilon_M^{-1}(q,0)$ is the inverse of the static Lindhard dielectric function of a homogeneous electron gas. The function $\epsilon_{\mathrm{SC}}^{-1}(\mathbf{q},\mathbf{q};0)$ used in our calculations was obtained using the empirical pseudopotential method, as described in Refs. 9 and 10. Figure 1 shows the difference $\epsilon_{\mathrm{SC}}^{-1}(\mathbf{q},\mathbf{q};\omega=0)-\epsilon_M^{-1}(q,\omega=0)$ calculated for various values of $\mathbf{q}$ as well as a numerical fit used in our calculation of $\delta\Sigma_{n,\mathbf{k}}$. Calculations done with various models of $\epsilon_{\mathrm{SC}}^{-1}(\mathbf{q},\mathbf{q};\omega=0)-\epsilon_M^{-1}(q,\omega=0)$ have shown that the resulting $\delta\Sigma_{n,\mathbf{k}}$ is not sensitive to the details of this function. The Bloch functions and eigenvalues used in the evaluation of Eq. (6) were obtained with the LDA *ab initio* norm-conserving pseudopotential method,[11] using the Ceperley-Alder functional.[12] The wave functions are expanded in plane waves with an energy cutoff of 50 Ry for diamond and 18 Ry for all other semiconductors. The divergence of $\delta W(\mathbf{k}-\mathbf{q}+\mathbf{G})$, appearing in Eq. (6) as $|\mathbf{k}-\mathbf{q}+\mathbf{G}|\rightarrow 0$, is treated with the technique described in Ref. 13, and two special points are used in the Brillouin-zone integration.

In Table I, we give the resulting self-energy corrections ($\delta\Sigma$) for all semiconductors for which *GW* results are available—i.e., diamond, Si, Ge, GaAs, and AlAs (Ref. 14)—and also for GaP. The differences between *GW* quasiparticle energies and LDA eigenvalues ($\epsilon^{GW}-\epsilon^{\mathrm{LDA}}$) are also given in Table I.

In general, the self-energy corrections $\delta\Sigma$ and the differences $\epsilon^{GW}-\epsilon^{\mathrm{LDA}}$ agree within 0.2 eV. Larger deviations (0.3–0.4 eV) occur for the $X_{4v}\rightarrow X_{1c}$ gap in diamond and the $\Gamma_{25'v}\rightarrow\Gamma_{2'c}$ gaps in Ge. We also give the LDA eigenvalues (ϵ^{LDA}), the experimental excitation energies (ϵ^{expt}), as well as the corrected LDA eigenvalues ($\epsilon^{\mathrm{LDA}}+\delta\Sigma$) for comparison. It should be noted that these latter values depend on the LDA eigenvalues, which can vary up to 0.2 eV depending on the details of the LDA computation (choice of atomic pseudopotentials, of energy cutoffs, and of exchange-correlation functionals). For example, the $\Gamma_{2'c}$ and L_{1c} states in Ge, as well as the Γ_{1c} state in GaAs, are particularly sensitive to the energy cutoff used in the plane-wave expansion.[16] A detailed comparison of the self-energy correction with the difference $\epsilon^{GW}-\epsilon^{\mathrm{LDA}}$ (Ref. 3) for silicon is given in Fig. 2. The correction proposed in this work follows closely the difference between *GW* and LDA results, thus

TABLE I. Self-energy corrections in various semiconductors as obtained with Eq. (6) ($\delta\Sigma$), compared to the difference between GW quasiparticle energies and LDA eigenvalues ($\epsilon^{GW}-\epsilon^{\mathrm{LDA}}$). The LDA eigenvalues ($\epsilon^{\mathrm{LDA}}$), the *corrected* LDA eigenvalues ($\epsilon^{\mathrm{LDA}}+\delta\Sigma$), and the experimental excitation energies are also given for comparison. All values are given in eV and measured with respect to the valence-band maximum. Experimental values are taken from Refs. 3, 4, and 15 with spin-orbit effects removed.

	$\delta\Sigma$	$\epsilon^{GW}-\epsilon^{\mathrm{LDA}}$	ϵ^{LDA}	$\epsilon^{\mathrm{LDA}}+\delta\Sigma$	ϵ^{expt}
Diamond					
$\Gamma_{15,c}$	2.01	2.0	5.50	7.51	7.3
$X_{4,v}\rightarrow X_{1,c}$	2.29	1.87	10.44	12.73	12.5
$L_{1,c}$	2.10		8.00	10.10	
E_g	2.06	1.7	3.48	5.54	5.48
Si					
$\Gamma_{15,c}$	0.66	0.78	2.55	3.21	3.4
$X_{1,c}$	0.69	0.74	0.64	1.33	1.3
$L_{1,c}$	0.64	0.76	1.46	2.10	2.1,2.4 $\pm$ 0.15
E_g	0.68	0.77	0.50	1.18	1.17
Ge					
$\Gamma_{2',c}$	0.46	0.78	−0.09	0.37	0.99
$X_{1,c}$	0.65		0.69	1.34	1.4 $\pm$ 0.2
$L_{1,c}$	0.55	0.68	0.10	0.65	0.84
GaAs					
$\Gamma_{1,c}$	0.74	0.91	0.48	1.22	1.63
$X_{1,c}$	0.93	0.70	1.40	2.33	2.09,2.12
$L_{1,c}$	0.81	0.79	0.97	1.78	1.92,1.94
AlAs					
$\Gamma_{1,c}$	0.92	0.97	1.87	2.79	3.11
$X_{1,c}$	1.03	0.81	1.35	2.38	2.24
$L_{1,c}$	0.92	0.90	2.03	2.95	2.49,2.54
GaP					
$\Gamma_{1,c}$	1.07		1.94	3.01	2.89
$X_{1,c}$	1.17		1.63	2.80	2.35
$L_{1,c}$	1.08		1.72	2.80	2.67,2.87

indicating that expression (4) is an essential contribution to the self-energy of nonmetals which is missing in the LDA. The Coulomb-hole contribution δE_{CH}, which must be added to both occupied and empty states, is positive and cancels approximately the effect of the screened exchange term in the valence bands. Consequently, the net effect of the self-energy correction is to shift the conduction bands upwards and leave the valence bands practically unchanged.

In order to clarify further the nature of the self-energy correction, we now decompose it into a scissor operator, similar to that proposed on empirical grounds by Baraff and Schlüter,[6] plus nonrigid corrections to the LDA eigenvalues. Considering Eq. (6), we see that the dominant terms in the sums are those with $\mathbf{G}=0$ and $|\mathbf{k}-\mathbf{q}|$ small. Retaining only these terms, which is valid when δW has only small-q components, we get the scissor operator

$$\delta\Sigma_{n,\mathbf{k}}\simeq\begin{cases}\delta E_{\mathrm{CH}}, & n\in\text{ conduction band},\\ \displaystyle\sum_{\mathbf{q}}\delta W(q)+\delta E_{\mathrm{CH}}\simeq\frac{\Omega}{(2\pi)^3}\int_0^{k_{\mathrm{BZ}}}4\pi q^2\,\delta W(q)\,dq+\delta E_{\mathrm{CH}}, & n\in\text{ valence band},\end{cases}\tag{8}$$

where we have replaced the summation over $\mathbf{q}$ by an integration over a sphere of radius k_{BZ} having the same volume as the Brillouin zone. The width of the scissor operator (8) calculated in this way is 0.8 eV for silicon. This value is comparable to the average of the self-energy corrections to the smallest gaps in this material (see Table I).

To summarize, we have proposed a self-energy correction to the LDA which takes into account the effects of incomplete screening in nonmetals. This self-energy correction calculated in various semiconductors follows closely the

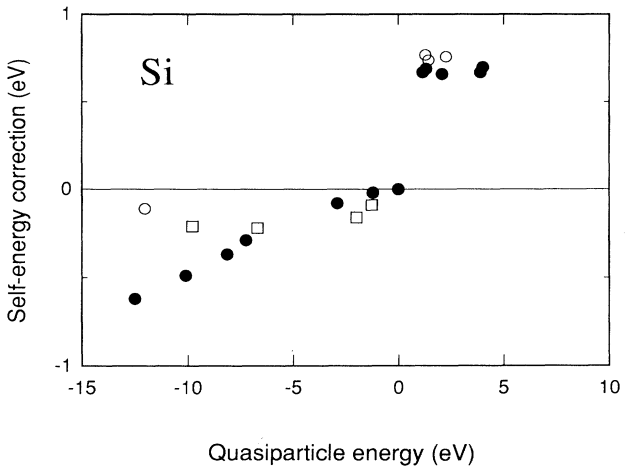

FIG. 2. Self-energy corrections $\delta\Sigma$ obtained in this work for silicon (dots), compared to the difference $\epsilon^{GW}-\epsilon^{\mathrm{LDA}}$ (open symbols). Open circles: Both ϵ^{GW} and ϵ^{LDA} taken from Ref. 3. Open squares: ϵ^{GW} from Ref. 3 and ϵ^{LDA} from our calculations, when Ref. 3 does not give an LDA value. All values are in eV and refer to the valence-band maximum.

difference between *GW* quasiparticle energies and LDA eigenvalues. The accuracy of the model is limited by the neglect of local-field effects and dynamic screening. The self-energy correction also contains as a dominant term a scissor operator whose width can be estimated without adjustable parameters, and can be useful as a first step in band-structure calculations.

[a]Present address: AT&T Bell Laboratories, Murray Hill, NJ 07974.

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

[2]W. Kohn and L. J. Sham, Phys. Rev. **140**, A1113 (1965).

[3]M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985); Phys. Rev. B **34**, 5390 (1986).

[4]R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986); Phys. Rev. B **35**, 4170 (1987); **37**, 10159 (1988).

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

[6]G. Baraff and M. Schlüter, Phys. Rev. B **30**, 3460 (1984).

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

[8]C. S. Wang and W. E. Pickett, Phys. Rev. Lett. **51**, 600 (1984); Phys. Rev. B **30**, 4719 (1985).

[9]A. Baldereschi and E. Tosatti, Phys. Rev. B **17**, 4710 (1978).

[10]R. Resta and A. Baldereschi, Phys. Rev. B **23**, 6615 (1981).

[11]G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

[12]D. Ceperley and B. Alder, Phys. Rev. Lett. **45**, 566 (1980); J. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[13]F. Gygi and A. Baldereschi, Phys. Rev. B **34**, 4405 (1986).

[14]For a discussion of the experimental value of the L_{1c} level in AlAs, see Ref. 4.

[15]*Semiconductors,* edited by O. Madelung, M. Schulz, and H. Weiss, Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology Vol. 17 (Springer-Verlag, Berlin, 1982), Group III, Part a.

[16]G. B. Bachelet and N. E. Christensen, Phys. Rev. B **31**, 879 (1985).