

Accurate Exchange-Correlation Potential for Silicon and Its Discontinuity on Addition of an Electron

R. W. Godby and M. Schlüter

AT&T Bell Laboratories, Murray Hill, New Jersey 07974

and

L. J. Sham

Department of Physics, University of California at San Diego, La Jolla, California 92093

(Received 10 February 1986)

We obtain an accurate density-functional exchange-correlation potential, $V_{xc}(\mathbf{r})$, for silicon, from calculations of the self-energy $\Sigma(\mathbf{r}, \mathbf{r}', \omega)$. No local-density approximation (LDA) is used for V_{xc} . The band structure with this V_{xc} is in remarkably close agreement with that obtained with the LDA, while both differ significantly from the quasiparticle spectrum of Σ . The 50% band-gap error found in LDA calculations is therefore not caused by the LDA but by the discontinuity, Δ , in the exact V_{xc} on addition of an electron.

PACS numbers: 71.45.Gm, 71.25.Rk

Density-functional theory^{1,2} (DFT) has allowed much progress to be made in calculating structural properties of solids from first principles.³ In DFT, the huge problem of calculating the ground state of the true system of 10^{23} interacting electrons is transformed rigorously into that of finding the ground state of a much simpler system of noninteracting electrons, moving in an effective potential $V(\mathbf{r})$. A part of the total energy functional, the *exchange-correlation energy* E_{xc} , and the corresponding potential $V_{xc}(\mathbf{r})$, are not known exactly. Nevertheless, the *local-density approximation*^{2,4} (LDA), in which the contribution to E_{xc} from each volume element is taken to be the same as it would be in a *homogeneous* electron gas, has been found to be remarkably successful. We present here calculations of V_{xc} for silicon that do not invoke the LDA, and indeed find it to be virtually indistinguishable from V_{xc}^{LDA} .

Electronic *excitations* are not, in general, described by DFT, but by the *quasiparticles* of many-body theory, whose energy $E(\mathbf{k})$ (the "energy-band structure") is the energy required to add an electron with given momentum to the system. Because of the absence of other techniques, it has often been assumed that the quasiparticle energy-band structure is similar to the band structure of the fictitious noninteracting electrons that appear in DFT, $E_{DFT}(\mathbf{k})$. This has been justified by the formal similarity of the Schrödinger-type equations that the DFT electrons and the quasiparticles obey, and by the exact result² that at zero temperature, the Fermi energy, E_F is given correctly by the highest occupied DFT eigenvalue.

It is found, however, that even the best-calculated LDA band structures fail to reproduce certain key features of experimental (quasiparticle) band structures; worst of all, the band gaps of insulators and

semiconductors are much too small. For example, the band gaps of rare-gas solids are 40% too small,⁵ and those of silicon and germanium 50% and 100% too small.⁶ Because of the importance of a detailed knowledge of the quasiparticle energies in the neighborhood of the Fermi energy for the understanding of defect states, surface states, heterojunctions, and transport properties, these shortcomings are severe.

Until recently it was thought that this error occurred either because DFT was simply incapable of describing excited states or because of the use of the LDA rather than exact DFT. However, Sham and Schlüter⁷ and Perdew and Levy⁸ have shown that the difference between the highest occupied and lowest unoccupied exact DFT eigenvalues in a large N -electron system (the DFT band gap, E_g^{DFT}) is *not* the exact quasiparticle band gap E_g , but differs from it by Δ , the discontinuity in the exchange-correlation potential when an electron is added to the insulating system:

$$\Delta \equiv E_g - E_g^{DFT} = V_{xc}^{(N+1)}(\mathbf{r}) - V_{xc}^{(N)}(\mathbf{r}) \quad (1)$$

(independent of \mathbf{r}), where $V_{xc}^{(N)}$ is the exact DFT exchange-correlation potential for the N -electron system. Δ has been shown⁹ to be a significant fraction of the energy gap for a one-dimensional model semiconductor within a two-plane-wave basis set, but its size for a real semiconductor was unknown until now (although there has been some discussion of the corresponding quantity for atoms). If Δ is small, exact DFT will give accurate band structures, and attention should be paid to going beyond the LDA in density-functional theory; if Δ is large, no attempt to add non-local-density corrections to the LDA will correct the calculated DFT band structures, and one must either go outside DFT, or take the discontinuity in V_{xc} into account explicitly.

In order to answer these questions we consider the true¹⁰ V_{xc} , and thence Δ , by using an exact relationship between V_{xc} and the self-energy Σ , as given by Sham and Schlüter⁷:

$$\text{Im} \int_{-\infty}^{E_F} [G_{\text{DFT}}(\Sigma - V_{xc})G]_{\mathbf{r},\mathbf{r}} d\omega = 0, \quad (2)$$

where the subscript \mathbf{r},\mathbf{r} denotes matrix multiplication inside the brackets with the $\mathbf{r}=\mathbf{r}'$ element then taken. G is the exact "interacting" one-particle Green's function and G_{DFT} the exact DFT Green's function. Equation (2) can be written as a set of linear equations for the elements of $V_{xc}(\mathbf{r})$, which, together with the fact that the highest occupied DFT eigenvalue equals the highest occupied quasiparticle energy, enable us to calculate $V_{xc}(\mathbf{r})$ without using the LDA. The discontinuity, Δ , is the difference between the resulting V_{xc} 's for the N - and $(N+1)$ -particle systems.

We now discuss our approximations. We begin by calculating the dielectric function for silicon, $\epsilon(\mathbf{r}, \mathbf{r}', \omega) = 1 - V\chi$, where χ is the random-phase approximation (RPA) susceptibility (including local field effects) and V the Coulomb interaction $1/|\mathbf{r}-\mathbf{r}'|$. We use well-converged LDA wave functions and eigenvalues from a nonlocal pseudopotential calculation to construct χ . This indirect use of the LDA quantities is justified by the tested insensitivity of our results to the LDA's band-gap error, and *a posteriori* by the closeness of the LDA, accurate DFT, and quasiparticle wave functions. Here and in the rest of the calculation we use a 169-plane-wave basis set, 65 energy bands, and either six or ten \mathbf{k} points in the irreducible wedge of the Brillouin zone. The various frequency integrals are performed on the imaginary axis. Further details of our methods will be included in a later paper.¹¹

The screened Coulomb interaction $W(\omega) = \epsilon^{-1}(\omega) \times V$ is formed straightforwardly, and used in Hedin's^{12,13} "GW" expression for the self-energy, which neglects only the vertex corrections in the expansion of G , and which they find to be very accurate for the homogeneous electron gas:

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{4\pi} \int_{-\infty}^{\infty} W(\mathbf{r}, \mathbf{r}', \omega') \times G(\mathbf{r}, \mathbf{r}', \omega + \omega') d\omega'. \quad (3)$$

We replace G (and, later, G_{DFT}) by $G_{\text{DFT}}^{\text{LDA}}$, justified as above. Hybertsen and Louie¹⁴ have recently shown a similar approximation to be adequate by comparing calculated quasiparticle energies with the experimental band gaps. We also calculate the bare (Hartree-Fock) and statically screened exchange self-energies (without the Coulomb hole) in which $W(\omega')$ is replaced by V and $W(\omega'=0)$, respectively. The true V_{xc} is then calculated by use of (2).

In Fig. 1 we compare our calculated $V_{xc}(\mathbf{r})$ with $V_{xc}^{\text{LDA}}(\mathbf{r})$, where we have chosen a LDA derived from

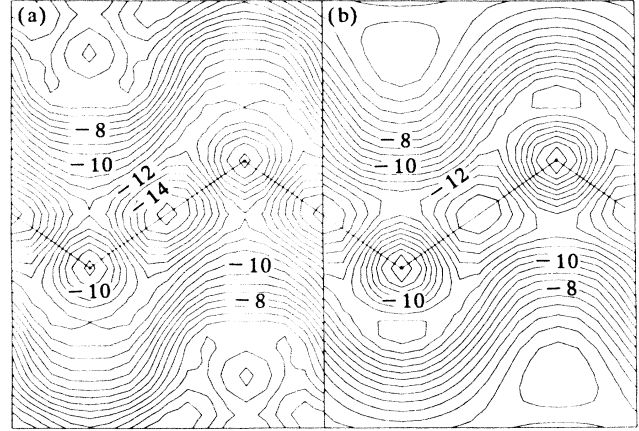


FIG. 1. Contour plots in the (110) plane containing the zigzag bond chain of (a) the true DFT exchange-correlation potential V_{xc} and (b) the RPA LDA potential (Ref. 14) V_{xc}^{LDA} , in electronvolts. The close similarity illustrates the high quality of the LDA for the creation of the ground-state electron density $n(\mathbf{r})$. In each case 169 plane waves were included in the summation.

RPA calculations of a homogeneous electron gas, which is also almost exactly equal to the more accurate Ceperley-Alder LDA potential minus a constant 0.49 eV. The agreement between these quite differently constructed potentials is surprising.¹⁵

Having obtained the true V_{xc} , we come to our main objective: to compare the true DFT band structure with that of the LDA and with the quasiparticle energies. We calculate these quantities by performing second-order perturbation theory in $\Sigma(E) - V_{xc}^{\text{LDA}}$ and $V_{xc} - V_{xc}^{\text{LDA}}$, and find all the second-order terms to be already less than 0.02 eV, illustrating the similarity of the LDA, quasiparticle, and exact DFT wave functions, and justifying our retention of the LDA Hartree potential.

We give the resulting energies in Table I. Remarkably, the true DFT eigenvalues are extremely close to the LDA (RPA) eigenvalues, even though the true V_{xc} was calculated from Σ , whose quasiparticle energies agree with the experimental band gaps. The rms deviation of the true DFT valence-band eigenvalues (relative to $\Gamma_{25'v}$) from the LDA values is only 0.02 eV. The desired discontinuity in V_{xc} , Δ , which is equal to the difference between the quasiparticle and DFT minimum band gaps, is 0.58 eV. Thus Δ is responsible for over 80% of the LDA gap error, 0.72 eV, in contrast to the results of recent model calculations.¹⁹ Approximations to V_{xc} that go beyond the LDA are therefore of rather limited value in the calculation of ground-state properties of silicon; neither will they yield the correct quasiparticle energies.

Interestingly, silicon's quasiparticle and DFT band structures differ mainly in a *rigid* upward shift in the

TABLE I. Ground-state DFT eigenvalues $E(V_{xc})$, LDA eigenvalues $E(V_{xc}^{LDA})$, and quasiparticle energies (bare exchange, statically screened exchange without Coulomb hole, and “ GW ”) $E(\Sigma)$ in electronvolts. The DFT eigenvalues (and especially the band gaps) are close to those of the LDA, and differ from the quasiparticle energies by an approximately rigid shift of the conduction bands. The quasiparticle energies $E(\Sigma_{GW})$ agree well with experiment.

	Experiment	$E(V_{xc}^{LDA})^a$	Bare	$E(\Sigma)$ Static	GW	$E(V_{xc})^b$	$E(\Sigma_{GW}) - E(V_{xc})$
$\Gamma_{25'\nu}$	0.00 ^c	-0.07	-0.55	7.88	0.00	0.00	0.00
$L_{3'\nu}$	-1.2 ± 0.2^d , -1.5^e	-1.29	-2.54	6.34	-1.19	-1.21	0.02
c-band min.	1.17 ^d	0.45	4.50	8.04	1.24	0.66	0.58
Γ_{15c}	3.40 ^d	2.50	7.39	10.91	3.30	2.68	0.62
$\Gamma_{2'c}$	4.19 ^d	3.49	8.54	11.95	4.27	3.66	0.61
L_{1c}	2.1 ^f , 2.4 ^e	1.46	6.05	9.61	2.30	1.62	0.68
L_{3c}	4.3 ± 0.2^d , $4.0^{d,e}$	3.30	8.35	11.24	4.11	3.49	0.62
Min. gap	1.17 ^d	0.52	5.05	0.16	1.24	0.66	0.58 ^g

^aHedin LDA (RPA) (Ref. 12); each eigenvalue is merely increased by 0.49 eV if the more sophisticated Ceperley-Alder LDA (Ref. 4) is used.

^b V_{xc} is the calculated true exchange-correlation potential for the N -particle system, $V_{xc}^{(N)}$.

^cAligned with the quasiparticle valence-band maximum.

^dReference 16.

^eReference 17.

^fReference 18.

^gEqual to Δ , the discontinuity in V_{xc} .

conduction bands, which is therefore also equal to Δ . This justifies the “scissor operator” approach²⁰ to estimating silicon quasiparticle energies from DFT band structures. However, comparison of LDA band structures and experimental quasiparticle energies in other semiconductors shows that if the LDA remains as good an approximation to V_{xc} as we have shown it to be here, a rigid shift cannot occur universally.

As is well known, the bare exchange (Hartree-Fock) self-energy causes the bands and band gaps to be far too wide ($E_g \approx 5.1$ eV). Screening this exchange with a static dielectric function (adjacent column) decreases E_g to only 0.2 eV and raises all energies. The inclusion of dynamic screening and the Coulomb hole (GW column) opens the gap to the experimental value and lowers the overall energies by about 7 eV.

The exact quasiparticle spectrum coincides with the exact DFT spectrum at the top of the valence band.² Table I shows that this condition is satisfied (to within 0.07 ± 0.10 eV) if we compare our Σ , calculated in the GW RPA-like approximation, with a V_{xc}^{LDA} derived from RPA calculations for a homogeneous electron gas. In contrast, the LDA of Ceperley and Alder,⁴ which goes beyond RPA, increases *all* the eigenvalues by 0.49 eV, leading to an apparent violation of the condition. The accuracy of the quasiparticle band gaps calculated here and the requirement that the tops of the valence bands be aligned therefore suggest that corrections to the RPA-like self-energy will merely

lead to a *constant* increase in the quasiparticle energies and V_{xc} of about 0.5 eV.

To gain insight into the physics that causes the spectrum of Σ to be very nearly that of V_{xc} and V_{xc}^{LDA} , but with a constant shift of the unoccupied bands, we consider the effects of the energy dependence and the nonlocality of Σ . We find that Σ 's energy dependence is crucial for obtaining the correct dispersion *within* each band, but its contribution to the gap is actually negative, decreasing it by 0.3 eV. The importance of nonlocality is demonstrated by Fig. 2, which shows $\Sigma(\mathbf{r}, \mathbf{r}')$ with \mathbf{r} fixed. The shape and range (approximately 0.5 bond length) of this function, though clearly not the amplitude, are fairly independent of the choice of \mathbf{r} . More distant extrema in Σ , reflecting the polarization of bonds, are very weak ($\leq 10^{-2}$ times the weight of the main “hole” around $\mathbf{r}=\mathbf{r}'$). The fact that the range of Σ and the length scale of the spatial variation of the wave functions are comparable, especially in the conduction bands where oscillations imposed by orthogonality are more prevalent (see Fig. 2), demonstrates that the nonlocality is crucial in obtaining the correct matrix elements of Σ , quasiparticle energies and band gap. Therefore Σ cannot be represented by a single local potential; effectively, the more rapidly oscillating conduction-band wave functions see a weaker (less negative) potential, so that the band gap is widened.

In summary, we have shown that nearly all the error

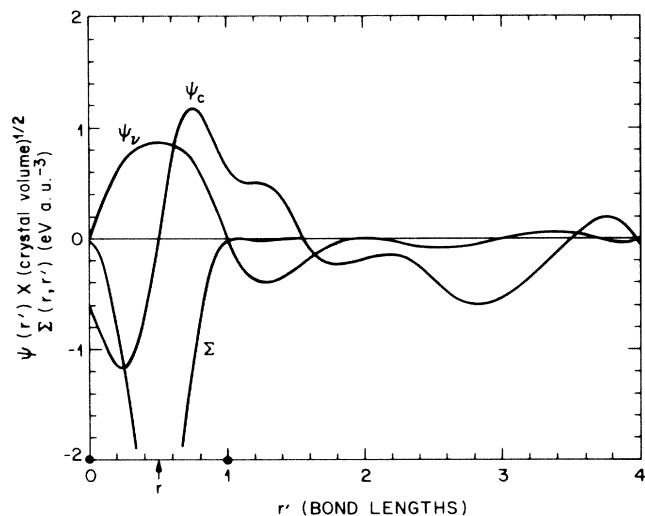


FIG. 2. Σ : the self-energy $\Sigma(\mathbf{r}, \mathbf{r}', \omega = \text{mid-gap})$ for \mathbf{r} at a bond center and \mathbf{r}' along [111]. ψ_v : a valence band wave function $\text{Re}\psi_{\text{DFP}}^{\text{LDA}}(\mathbf{r}')$ in the highest valence band, near Γ . ψ_c : a conduction-band wave function in the lowest conduction band, near Γ . Atomic positions are indicated by closed circles. The oscillations of the conduction-band wave functions over the range of Σ lead to less negative matrix elements and therefore a larger band gap than if Σ were local.

in the silicon band gap that occurs in LDA calculations also occurs with an accurate exchange-correlation potential, and is therefore inherent in the ground-state DFT band structure. The LDA is, however, an excellent approximation for obtaining ground-state properties of silicon. The correct gap may be obtained from DFT if the discontinuity in V_{xc} is taken into account.

We thank S. G. Louie and M. Hybertsen for many valuable discussions, and G. A. Baraff, P. Fulde, and D. R. Hamann for reading the manuscript critically. One of us (L.J.S.) was supported in part by the National Science Foundation through DMR Grant No. 8514159.

¹P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).

²W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).

³See, for example, *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N. H. March (Plenum, New York, 1983), and references therein.

⁴D. M. Ceperley and B. I. Alder, Phys. Rev. Lett. **45**, 566 (1980); parametrized in J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

⁵S. B. Trickey, F. R. Green, Jr., and Frank W. Averill, Phys. Rev. B **8**, 4822 (1973).

⁶D. R. Hamann, Phys. Rev. Lett. **42**, 662 (1979); G. B. Bachelet and N. E. Christensen, Phys. Rev. B **31**, 879 (1985).

⁷L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983); L. J. Sham and M. Schlüter, Phys. Rev. B **32**, 3883 (1985).

⁸J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).

⁹M. Lannoo, M. Schlüter, and L. J. Sham, Phys. Rev. B **32**, 3890 (1985).

¹⁰We use "true V_{xc} " to mean the V_{xc} that corresponds to the calculated Σ through this exact relationship; if Σ were exact, this would be the exact V_{xc} .

¹¹R. W. Godby, M. Schlüter, and L. J. Sham, to be published.

¹²L. Hedin, Phys. Rev. A **139**, 796 (1965).

¹³L. Hedin and S. Lundqvist, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1969), Vol. 23, p. 1.

¹⁴M. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985).

¹⁵The most prominent difference, in the interstitial region, is weak and sensitive to numerical errors in solving the integral equation (2) and therefore may be spurious. The eigenvalues, however, are not sensitive to such errors.

¹⁶D. Von Bimberg *et al.*, in *Physics of Group IV Elements and III-V Compounds*, edited by O. Madelung, M. Schulz, and H. Weiss, Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology, New Series Group 3, Vol. 17, Part a (Springer, New York, 1982).

¹⁷D. Straub, L. Ley, and F. J. Himpsel, Phys. Rev. Lett. **54**, 142 (1985).

¹⁸R. Hulthen and N. G. Nilsson, Solid State Commun. **18**, 1341 (1976).

¹⁹O. Gunnarsson and K. Schönhammer, Phys. Rev. Lett. **56**, 1968 (1986).

²⁰G. A. Baraff and M. Schlüter, Phys. Rev. B **30**, 3460 (1984).