Exact Kohn-Sham Exchange Potential in Semiconductors

M. Städele, J. A. Majewski, and P. Vogl
Physik-Department und Walter Schottky Institut, Technische Universität München, Am Coulombwall, D-85748 Garching, Germany

A. Görling
Lehrstuhl für Theoretische Chemie, Technische Universität München, D-85748 Garching, Germany

(Received 13 March 1997)

A new Kohn-Sham method that treats exchange interactions within density functional theory exactly is applied to Si, diamond, GaN, and InN. The exact local exchange potential leads to significantly increased band gaps that are in good agreement with experimental data. Generalized gradient approximations yield exchange energies that are much closer to the exact values than those predicted by the local density approximation. The exchange contribution to the derivative discontinuity of the exchange-correlation potential is found to be very large (of the order of 5–10 eV). [S0031-9007(97)04017-9]

PACS numbers: 71.15.Mb, 71.45.Gm, 71.55.Cn, 71.55.Eq

All current electronic structure calculations of solids that are based on the Kohn-Sham (KS) method of density-functional theory [1] approximate both the exchange as well as the correlation potential and the corresponding energies. Most frequently, the local density approximation (LDA) [1] or generalized gradient approximations (GGA) [2–5] are employed. Recently, several schemes have been proposed to calculate the KS exchange potential rigorously [6–9]. The first practical implementations of these schemes for solids gave encouraging results for band structures in semiconductors, but were forced to use shape approximations for the potentials [9,10] or averaged eigenvalue gaps in the one-particle Green function [11,12].

In this Letter, we present an *exact* determination of the KS exchange potential, its discontinuity, and the exchange energy for solids, and discuss results for Si, diamond, GaN, and InN. In addition, the exact treatment of the exchange interaction enables us to assess the quality of conventional approximate KS methods. Our calculations are based on a recently developed general procedure [7,8] that allows one to evaluate the local KS exchange potential $V_x(\mathbf{r})$ rigorously. The potential $V_x(\mathbf{r})$ equals the functional derivative of the exchange energy E_x with respect to the electron density $\rho(\mathbf{r})$. This derivative is not directly accessible, however, since E_x is known only in terms of the one-particle KS states φ_i, and the explicit functional dependence of the orbitals φ_i on the electron density is unknown. The key idea of Ref. [7] is to determine $V_x(\mathbf{r})$ by noting that the exchange potential requires only the knowledge of the *first order* change of E_x with respect to the density. One may employ the chain rule to obtain, in a shorthand matrix notation,

$$\boldsymbol{V}_x = \frac{\delta E_x}{\delta \boldsymbol{\rho}} = \sum_i^{\mathrm{occ}} \left[\frac{\delta E_x}{\delta \boldsymbol{\varphi}_i} \frac{\delta \boldsymbol{\varphi}_i}{\delta \boldsymbol{V}_{\mathrm{KS}}} + \mathrm{c.c.} \right] \frac{\delta \boldsymbol{V}_{\mathrm{KS}}}{\delta \boldsymbol{\rho}} . \quad (1)$$

In Eq. (1), $\boldsymbol{V}_{\mathrm{KS}}$ denotes the total local potential in the single-particle KS Hamiltonian $\mathbf{T} + \boldsymbol{V}_{\mathrm{KS}}$, the eigenstates of which are $\boldsymbol{\varphi}_i$. The three first-order derivatives on the right-hand side of Eq. (1) can be calculated rigorously from E_x, from first order perturbation theory, and by invoking linear response theory [7,8], respectively.

Because of the Hohenberg-Kohn theorem, there exists a one-to-one mapping between the potentials $V_{\mathrm{KS}}(\mathbf{r})$ and densities $\rho(\mathbf{r})$ which guarantees that the functional derivative $\delta V_{\mathrm{KS}}(\mathbf{r})/\delta\rho(\mathbf{r}')$ is defined. Its matrix representation has to be constructed by inverting $\delta\boldsymbol{\rho}/\delta\boldsymbol{V}_{\mathrm{KS}}$ on a restricted function space of $\delta\rho(\mathbf{r})$ and $\delta V_{\mathrm{KS}}(\mathbf{r})$ that excludes constant functions. This can most easily be achieved by representing the KS orbitals and potentials in a plane-wave basis and excluding the zero wave-vector components [8]. In a plane-wave representation, with $\mathbf{G}, \mathbf{G}'$ denoting reciprocal lattice vectors, one obtains

$$V_x(\mathbf{G}) = \sum_{\mathbf{G}'\neq 0} [E(\mathbf{G}') + E^*(-\mathbf{G}')] \tilde{\chi}_0^{-1}(\mathbf{G}, \mathbf{G}'), \quad (2)$$

$$E(\mathbf{G}) = \frac{1}{\Omega} \sum_{vc\mathbf{k}} \langle v\mathbf{k} | V_x^{\mathrm{NL}}(\mathbf{r}, \mathbf{r}') | c\mathbf{k} \rangle \times \rho_{cv\mathbf{k}}(-\mathbf{G}) [\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}]^{-1}, \quad (3)$$

$$\chi_0(\mathbf{G}, \mathbf{G}') = \frac{2}{\Omega} \sum_{vc\mathbf{k}} \rho_{vc\mathbf{k}}^*(\mathbf{G}) \rho_{cv\mathbf{k}}(\mathbf{G}') [\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}]^{-1}, \quad (4)$$

where $\rho_{cv\mathbf{k}}(\mathbf{G}) = \langle c\mathbf{k} | \exp[i\mathbf{G} \cdot \mathbf{r}] | v\mathbf{k} \rangle$, $\varepsilon_{n\mathbf{k}}$ are the valence (v) and conduction (c) state KS eigenvalues of Bloch momentum $\mathbf{k}$, and Ω denotes the crystal volume. The operator V_x^{NL} in Eq. (3) has the same form as the nonlocal Hartree-Fock exchange operator but is constructed with KS one-particle states. The inverse of χ_0 is to be taken in the subspace of nonzero reciprocal lattice vectors and is denoted by $\tilde{\chi}_0^{-1}$. Note that the zero wave-vector component of V_x can be chosen arbitrarily since the KS orbitals are invariant with respect to adding a constant to the KS potential. If correlations are neglected, the solution of the KS equations with the exchange potential of Eq. (1) represents an exact “exchange-only” KS procedure which is physically equivalent to the so-called optimized potential method (OPM) [6] that has only been applied to atoms so far [6,11,13]. However, a straightforward extension of this method to periodic systems leads to divergencies since the integral kernel of the OPM integral equation is not invertible for infinite systems. The present exact exchange formalism, that we abbreviate by EXX, eliminates these divergencies exactly.

Equation (2) gives the exact local KS exchange potential V_{KS} in a form that is well suited for pseudopotential plane-wave calculations. In order to rigorously stay within the density functional framework, the external potential in the KS equations should be local as well. We have therefore employed a local ionic pseudopotential for Si [14] at first. In addition, we have evaluated V_x using standard nonlocal, norm-conserving pseudopotentials [15].

All densities and electronic energies were determined with ten special **k** points. Kinetic energy cutoffs of 25, 65, 45, and 40 Ry, for Si, diamond, GaN, and InN, respectively, were used, and the band summations included all conduction bands. The matrix size of $\tilde{\chi}_0$ lay between 258×258 and 458×458. We have checked very carefully that the presented band gaps are converged to 0.02 eV, LDA and GGA energies to 0.01 eV/atom, and EXX and Hartree-Fock energies [16] to 0.04 eV/atom.

Figure 1(a) depicts the exact and approximate exchange potentials of the valence electrons in silicon along the [111] direction. The LDA potential underestimates the spatial variation of the exchange potential significantly. In the physically most relevant bonding region in between the Si atoms, V_x^{GGA} is seen to be superior to the LDA potential that is less attractive and smoother than the exact potential. This is due to the unphysical, repulsive, self-interaction that is contained in the LDA. Other gradient approximations [2,3,5] show similar trends and yield too large exchange potentials in regions of low density. In accord with these findings, the EXX method yields a more pronounced variation of the valence charge density across the [111] axis than LDA as shown in Fig. 1(b). The GGA methods tend to overestimate this more pronounced inhomogeneity of the charge density. These trends in the exchange potentials and densities are present in all semiconductors that we have studied. The following values (given in parenthesis) are the maximum and minimum valence electron densities, respectively, in Å^{-3} along the [111] axis analogous to Fig. 1. Diamond: LDA (0.095; 2.03), GGA [4] (0.066; 2.13), and EXX (0.078; 2.06). GaN: LDA (0.039; 3.41), GGA [4] (0.022; 3.55), and EXX (0.031; 3.41).

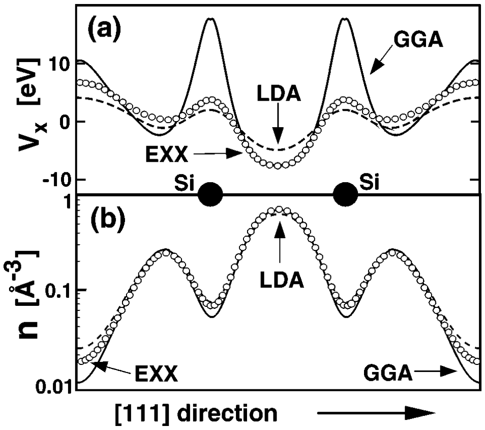

FIG. 1. (a) Comparison of the calculated exact exchange potential (open circles), in eV, along the [111] direction in Si with the approximate LDA (dashed line) and GGA (solid line) [4] exchange potentials. The filled black circles correspond to the positions of the Si atoms. The LDA and GGA potentials were evaluated at the exchange-only EXX density and the local ionic pseudopotential was employed. The mean values of all exchange potentials have been set equal to zero. (b) Self-consistent charge densities computed with the indicated exchange-only functionals.

The accuracy of an approximate density functional $E_x^{\mathrm{app}}[\rho]$ can be assessed in two ways. One may determine the deviation from the exact functional $E_x^{\mathrm{EXX}}[\rho]$ at a given, fixed reference density ρ^0. Alternatively, one may determine the self-consistent density ρ^{app} that corresponds to this functional and compare $E_x^{\mathrm{app}}[\rho^{\mathrm{app}}]$ with the exact value $E_x^{\mathrm{EXX}}[\rho^{\mathrm{EXX}}]$. In Si, we find the exact exchange energy $E_x^{\mathrm{EXX}}[\rho^{\mathrm{EXX}}]$, evaluated at the self-consistent density ρ^{EXX}, to equal -29.40 eV/atom. Here, the local pseudopotential [14] has been employed again. In Fig. 2, this result is compared to those obtained with several approximate functionals [2–5]. The LDA exchange energy is -27.72 eV/atom and deviates from the exact value by a very significant amount of 6% whereas all gradient approximations show much improved values. Thus, gradient approximations that have been originally fitted to atomic data show a remarkable transferability to extended systems. Figure 2 also shows that the agreement with the exact energy deteriorates if one calculates the exchange energies $E_x^{\mathrm{app}}[\rho^{\mathrm{app}}]$ self-consistently. We note

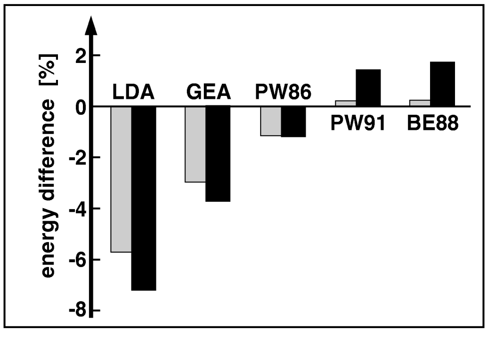

FIG. 2. Exchange energy differences of Si, in percent, as calculated using the LDA [17], GEA [5], PW86 [2], PW91 [3], and BE88 [4] approximate exchange functionals E_x^{app}, relative to the EXX exchange energy. The grey bars have been calculated for the fixed density ρ^{EXX}. The black bars represent results from self-consistent, approximate exchange-only calculations using the charge density ρ^{app}.

that these self-consistent calculations were performed without taking correlation contributions into account.

The self-consistent total energy $E_{\text{tot}}^{\text{EXX}}[\rho^{\text{EXX}}]$ in Si is found to be equal to -104.75 eV/atom. This has to be compared with the exchange-only LDA and GGA [4] values of -103.21 and -104.92 eV/atom, respectively. Thus, the relative error is -1.5% for LDA and only 0.2% for the GGA of Ref. [4]. All these findings show clearly that GGA exchange and total energies are definitely superior to the ones obtained within the LDA. We obtained very similar trends in all the other solids that we have investigated.

The self-consistently determined Hartree-Fock exchange and total energies are -29.61 eV/atom and -104.87 eV/atom, respectively, in Si. Thus, they are very close to the corresponding EXX values. The same holds for the self-consistent Hartree-Fock and EXX densities that deviate by less than $10^{-4}\%$ at the bond center. Similar findings have been reported for atoms before [11]. We point out that the difference between Hartree-Fock and exchange-only EXX is that E_x is evaluated with Hartree-Fock orbitals and Kohn-Sham orbitals, respectively. In the former method, the single particle equations contain a *nonlocal* exchange potential, whereas the EXX method leads to a *local* potential. This locality may be considered as an additional restriction for the wavefunctions, which causes the EXX total energy to be higher than in Hartree-Fock.

In Table I, we summarize the fundamental band gaps in Si, diamond, GaN, and InN as obtained by the LDA and EXX method, respectively. In both cases, we have included LDA correlations [17], even though they are found to raise the gaps by only 0.1–0.2 eV in all materials and in both methods. In the case of Si, the use of local or nonlocal pseudopotentials makes very little difference. For all considered semiconductors, the EXX band gaps are roughly 1 eV larger than the corresponding LDA gaps and in much better agreement with experiment. This systematic improvement can be attributed to the absence of the self-interaction in the EXX formalism. In the LDA method, the self-interaction is well known to raise the more localized valence states relative to the delocalized conduction band states and to diminish the gap.

TABLE I. Energy gaps (in eV) and exchange parts Δ_x of the discontinuity (in eV) for various semiconductors. The minimal gap E_{gap} and higher band gaps, relative to the top of the valence band, are given. The calculated values shown are obtained by the LDA and the present EXX method, respectively. For Si, values computed with local and nonlocal pseudopotentials are shown. The experimental value for InN refers to the wurtzite structure.

		LDA	EXX	Δ_x	Expt.
Si	L	1.54	2.36	5.84	2.4[a]
(local)	Γ	2.79	3.46	6.15	3.05[b]
	E_{gap}	0.52	1.43	5.48	1.17[c]
Si	L	1.45	2.30	5.98	2.4[a]
(nonlocal)	Γ	2.55	3.29	6.31	3.05[b]
	E_{gap}	0.49	1.44	5.62	1.17[c]
C	L	8.42	9.19	10.18	
	Γ	5.57	6.28	9.29	7.3[d]
	E_{gap}	4.16	5.06	8.70	5.47[d]
GaN	L	4.58	5.94	8.40	
	Γ	1.90	3.46	7.63	3.30[e]
	X	3.38	5.09	7.62	
InN	L	3.19	4.55	7.44	
	Γ	-0.18	1.40	6.14	1.95[f]
	X	2.94	4.68	6.76	

[a] Ref. [24].
[b] Ref. [25].
[c] Ref. [26].
[d] Ref. [27].
[e] Ref. [28].
[f] Ref. [26].

In Table II, we compare the presently calculated band gaps for silicon with results obtained by alternative density functional or quasiparticle methods. Both the EXX/ASA [9,10] and the KLI method of Ref. [12] also attempt to compute the exchange potential exactly. The former method employs a shape approximation (atomic sphere approximation, ASA) for the KS potential and density. The KLI method uses an averaged eigenvalue approximation in the one-particle Green function contained implicitly in Eqs. (3) and (4). In order to check the dependence of the results on the type of pseudopotential, we have employed standard LDA [15] as well as relativistic KLI [18] and (consistent) EXX [19] pseudopotentials. As summarized in Table II, the different pseudopotentials have very little influence on the EXX results in Si. By contrast, the KLI approximation for the valence electrons is seen to have a very noticeable effect on band gaps.

The exact KS band gap differs from the true band gap by a finite discontinuity Δ_{xc} of the exchange-correlation potential [20], whose size has not been unambiguously determined until now [21]. Fortunately, the exact exchange potential enables us to calculate the exchange contribution

TABLE II. Comparison of band gaps in Si, calculated with the present EXX method and other schemes that go beyond LDA. In parenthesis beneath the top row, the types of ionic pseudopotentials employed are given.

Si	EXX (LDA)	EXX (EXX)[d]	EXX (KLI)[a]	KLI[a] (KLI)[a]	EXX/ASA[b]	GW[c]
L	2.30	2.35	2.26	1.82	1.98	2.27
Γ	3.29	3.26	3.25	2.87	2.87	3.35
X	1.58	1.50	1.49	0.94	1.24	1.44

[a] Ref. [18].
[b] Ref. [10].
[c] Ref. [29].
[d] Ref. [19].

Δ_x to the discontinuity exactly. It is given by [22,23]

$$\Delta_x = \langle c\mathbf{k}|V_x^{\mathrm{NL}} - V_x|c\mathbf{k}\rangle - \langle v\mathbf{k}'|V_x^{\mathrm{NL}} - V_x|v\mathbf{k}'\rangle , \tag{5}$$

where $|c\mathbf{k}\rangle$ and $|v\mathbf{k}'\rangle$ denote the energetically lowest conduction and highest valence band states, respectively. As can be deduced from Table I, Δ_x amounts to typically twice the band gap. Since the discontinuity Δ_{xc} is guaranteed to be smaller than the true band gap by its definition, the correlation contribution to the discontinuity must cancel a large portion of the exchange part Δ_x. Importantly, the sum of the EXX band gap and Δ_x turns out to be almost identical to the Hartree-Fock band gap; the differences are found to be less than 3% in the semiconductors that we have studied. This can be understood by assuming the difference between the Hartree-Fock and the exchange-only EXX orbitals to be negligible. With this assumption, the matrix element $\langle \varphi_i|V_x^{\mathrm{NL}} - V_x|\varphi_i\rangle$ equals the difference between the ith Hartree-Fock and exchange-only EXX one-particle energy. Thus, Δ_x equals the difference between the Hartree-Fock and the EXX eigenvalue gaps. We note that one may interpret the right-hand side of Eq. (5) alternatively as the exchange contribution to the excitation energy between any pair of valence and conduction band states [22], and we have included several of these values in Table I as well.

In conclusion, we have provided a way to take into account exchange in the KS theory exactly that may be considered a rigorous starting point for treating correlations.

We are indebted to L. Kleinman and D. M. Bylander for making their KLI pseudopotentials available to us prior to publication. Financial support by the Deutsche Forschungsgemeinschaft (DFG) and Bayerische Forschungsverbund FOROPTO II is gratefully acknowledged. A. G. thanks Deutsche Forschungsgemeinschaft for a Heisenberg fellowship.

[1] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer, Berlin, 1990).
[2] J. P. Perdew and W. Yue, Phys. Rev. B **33**, 8800 (1986).
[3] J. P. Perdew, in *Electronic Structure of Solids '91,* edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991).
[4] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[5] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992), and references therein.
[6] J. D. Talman and W. F. Shadwick, Phys. Rev. A **14**, 36 (1976).
[7] A. Görling and M. Levy, Phys. Rev. A **50**, 196 (1994); Int. J. Quant. Chem. Symp. **29**, 93 (1995).
[8] A. Görling, Phys. Rev. B **53**, 7024 (1996).
[9] T. Kotani, Phys. Rev. Lett. **74**, 2989 (1995).
[10] T. Kotani and H. Akai, Phys. Rev. B **54**, 16 502 (1996).
[11] J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **45**, 101 (1992).
[12] D. M. Bylander and L. Kleinman, Phys. Rev. Lett. **74**, 3660 (1995); Phys. Rev. B **52**, 14 566 (1995); *ibid.* **54**, 7891 (1996).
[13] E. Engel and S. H. Vosko, Phys. Rev. A **47**, 2800 (1993).
[14] J. Ihm, M. L. Cohen, and D. J. Chadi, Phys. Rev. B **21**, 4592 (1980).
[15] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991); G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).
[16] In this context, we have used the method of F. Gygi and A. Baldereschi, Phys. Rev. B **34**, 4405 (1986).
[17] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[18] D. M. Bylander and L. Kleinman, Phys. Rev. B **55**, 9432 (1997).
[19] M. Moukara, M. Städele, J. A. Majewski, A. Görling, and P. Vogl, unpublished.
[20] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).
[21] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10 159 (1988); O. Gunnarsson and K. Schönhammer, Phys. Rev. Lett. **56**, 1968 (1986).
[22] A. Görling and M. Levy, Phys. Rev A **52**, 4493 (1995); A. Görling, Phys. Rev A **54**, 3912 (1996).
[23] J. P. Perdew, in *Density Functional Methods in Physics,* edited by R. M. Dreizler and J. da Providencia (Plenum, New York, 1985), p. 265.
[24] D. Straub, L. Ley, and F. J. Himpsel, Phys. Rev. Lett. **54**, 142 (1985).
[25] J. E. Ortega and F. J. Himpsel, Phys. Rev. B **47**, 2130 (1993).
[26] *Numerical Data and Functional Relationships in Science and Technology,* edited by K.-H. Hellwege, Landolt-Börnstein, New Series, Group III, Vol. 17, edited by O. Madelung, M. Schulz, and H. Weiss (Springer, Berlin, 1982).
[27] A. Mainwood, in *Properties and Growth of Diamond,* edited by G. Davies (Electronic Materials Information Service, London, 1994), p. 3.
[28] G. Ramirez-Flores, H. Navarro-Contreras, A. Lastras-Martinez, R. C. Powell, and J. E. Greene, Phys. Rev. B, **50**, 8433 (1994).
[29] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).