

Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities

John P. Perdew

*Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, and
Department of Physics and Quantum Theory Group, Tulane University, New Orleans, Louisiana 70118*

and

Mel Levy

Department of Chemistry and Quantum Theory Group, Tulane University, New Orleans, Louisiana 70118

(Received 17 June 1983)

The local-density approximation for the exchange-correlation potential underestimates the fundamental band gaps of semiconductors and insulators by about 40%. It is argued here that underestimation of the gap width is also to be expected from the unknown *exact* potential of Kohn-Sham density-functional theory, because of derivative discontinuities of the exchange-correlation energy. The need for an energy-dependent potential in band theory is emphasized. The center of the gap, however, is predicted exactly by the Kohn-Sham band structure.

PACS numbers: 71.45.Gm, 71.25.Rk, 71.25.Tn

Kohn-Sham density-functional theory¹ is the *de facto* foundation of most modern band-structure calculations. The Kohn-Sham self-consistent equations yield, in principle, the exact ground-state density $\tilde{n}(\mathbf{r})$ and total energy E . The Kohn-Sham one-electron orbital energies in solids are usually assigned no precise physical meaning,^{1,2} except for the highest occupied one which, in a metal, equals minus the work function.³ Here we will show that the proper generalization of this familiar theorem to nonmetallic crystals is this: The exact Kohn-Sham band structure predicts exactly the center of the fundamental energy gap relative to the vacuum level. The width of the gap, however, is underestimated, *even though it is a difference of ground-state energies*. The local-density approximation¹ (LDA) to the exact exchange-correlation potential $\delta E_{xc}/\delta n(\mathbf{r})$ is

known⁴ to underestimate the gap width by about 40% in insulators and semiconductors. We suggest that these errors might not be substantially reduced by improved approximations to $\delta E_{xc}/\delta n(\mathbf{r})$, but rather by *energy-dependent* electron self-energies⁵ $\Sigma_{xc}([n]; \epsilon; \mathbf{r})$ which lie beyond the Kohn-Sham formalism.

The Kohn-Sham self-consistent equations for a system with an integer number M of electrons are

$$\left\{-\frac{1}{2}\nabla^2 + v_{\text{eff}}([n]; \mathbf{r})\right\}\psi_i(\mathbf{r}) = \epsilon_i(M)\psi_i(\mathbf{r}), \quad (1)$$

$$n(\mathbf{r}) = \sum_{i=1}^M |\psi_i(\mathbf{r})|^2. \quad (2)$$

The one-electron orbital energies are labeled so that $\epsilon_1(M) \leq \epsilon_2(M) \leq \dots$, and the ground-state total energy is

$$E(M) = \sum_{i=1}^M \epsilon_i(M) - \int d^3r n(\mathbf{r})v_{\text{eff}}([n]; \mathbf{r}) + \int d^3r n(\mathbf{r})v(\mathbf{r}) + \frac{1}{2} \int d^3r \int d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[n], \quad (3)$$

where $E_{xc}[n]$ is the exchange-correlation energy. The effective potential is

$$v_{\text{eff}}([n]; \mathbf{r}) = v(\mathbf{r}) + \int d^3r' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{xc}}{\delta n(\mathbf{r})}, \quad (4)$$

where $v(\mathbf{r})$ is the electron-nuclear attraction. The functional derivative $\delta E_{xc}/\delta n(\mathbf{r})$ is a multiplicative exchange-correlation potential. In LDA, $E_{xc}[n]$ is approximated as $\int d^3r n(\mathbf{r})\epsilon_{xc}(n(\mathbf{r}))$.

Define the ionization potential and electron affinity of the M -electron system as $I(M) = E(M-1) - E(M)$ and $A(M) = E(M) - E(M+1)$, respec-

tively. If the M -electron crystal is charge-neutral, the center and width of its fundamental gap are respectively $-\frac{1}{2}[I(M)+A(M)]$ and $I(M) - A(M)$ (cf. Mott⁶). These total-energy differences are related to the one-electron energies through a theorem of Janak³:

$$I(M) = - \int_0^1 df \epsilon_M(M-1+f) = -\epsilon_M(M-\delta), \quad (5)$$

$$A(M) = - \int_0^1 df \epsilon_{M+1}(M+f) = -\epsilon_{M+1}(M+\delta). \quad (6)$$

The appearance of noninteger electron number in

Eqs. (5) and (6) is justified by the density-functional theory of open systems.^{1,7-9} Furthermore, it is only in an open system that $\delta E_{xc}/\delta n(\vec{r})$ of Eq. (4) is well defined: In a closed system $\delta E_{xc}/\delta n(\vec{r})$ is defined only up to an arbitrary additive constant (cf. Ref. 8). In the limit $M \rightarrow \infty$, the integrals over occupation number of f in Eqs. (5) and (6) may be replaced³ by the integrand at any point in the range $0 < f < 1$; the positive infinitesimal δ brings the electron number N as close to the integer M as possible. Recall that $\epsilon_M(M)$ and $\epsilon_{M+1}(M)$ are, respectively, the highest occupied and lowest unoccupied one-electron levels of the M -electron system.

The addition or removal of an infinitesimal number of electrons makes only an infinitesimal change $\delta n(\vec{r})$ in the time-averaged electron density. One might suppose that this change arises from an infinitesimal change in the effective potential:

$$\delta n(\vec{r}) \leftrightarrow \delta v_{\text{eff}}([n]; \vec{r}). \quad (7)$$

Such a change would not alter the one-electron energies, so Eqs. (5) and (6) would become $I(M) = -\epsilon_M(M)$ and $A(M) = -\epsilon_{M+1}(M)$. Thus in a metal the work function $\Phi = I(M) = A(M)$ is given by $-\epsilon_M(M)$, where $\epsilon_M(M)$ is the greatest occupied one-electron energy.³ [Because of the long range of the electrostatic potential, the surface must be included in the calculation of $\epsilon_M(M)$.] When this argument is applied^{10,11} to a semiconductor, the result is that the physical gap width $I(M) - A(M)$ is given by $\epsilon_{M+1}(M) - \epsilon_M(M)$, the gap width of the Kohn-Sham band structure.

Strictly, however, the infinitesimal density change $\delta n(\vec{r})$ of Eq. (7) could arise from a change of the effective potential which is an infinitesimal *plus* a finite constant C :

$$\delta n(\vec{r}) \leftrightarrow \delta v_{\text{eff}}([n]; \vec{r}) + C. \quad (8)$$

In recent work on the density-functional theory of open systems,^{1,7-9} we have discovered derivative discontinuities⁷ of the exact $E_{xc}[n]$: $\delta E_{xc}/\delta n(\vec{r})$ and hence $v_{\text{eff}}([n]; \vec{r})$ may jump by a positive constant C when the electron number increases through the integer M , i.e.,

$$\left. \frac{\delta E_{xc}}{\delta n(\vec{r})} \right|_{M+\delta} - \left. \frac{\delta E_{xc}}{\delta n(\vec{r})} \right|_{M-\delta} = C. \quad (9)$$

The subscript $N = M \pm \delta$ means the functional derivative is to be evaluated for the N -electron ground-level density. C , a functional of n , is independent of \vec{r} . We have recently shown⁹ that,

for *any* fixed N ,

$$\lim_{|\vec{r}| \rightarrow \infty} \delta E_{xc}/\delta n(\vec{r})|_N = 0.$$

This result is not inconsistent with

$$\lim_{|\vec{r}| \rightarrow \infty} \lim_{\delta \rightarrow 0} \delta E_{xc}/\delta n(\vec{r})|_{M+\delta} = C.$$

As $\delta \rightarrow 0$ from above, the boundary shell, separating the inner region where the exchange-correlation potential is C from the outer region where it is zero, expands out to infinity.

It follows from Eqs. (5), (6), and (9) that

$$I(M) - A(M) = \epsilon_{M+1}(M) - \epsilon_M(M) + C, \quad (10)$$

i.e., the Kohn-Sham band structure underestimates the gap width by an amount equal to the derivative discontinuity C . Recently Sham and Schlüter,¹² independently of us, have also concluded that the gap in the Kohn-Sham band structure is unphysical.

In the generalization of Kohn-Sham theory to an open system¹ characterized by temperature T and chemical potential μ , the time-averaged electron number N becomes a continuous variable. By construction the real system and its auxiliary noninteracting counterpart have the same chemical potential; in particular,

$$-\frac{1}{2}[I(M) + A(M)] = \frac{1}{2}[\epsilon_M(M) + \epsilon_{M+1}(M)] \quad (11)$$

in the limit $T \rightarrow 0$.⁷⁻⁹ Thus the Kohn-Sham band structure does yield the correct gap center relative to the vacuum level. It follows from Eqs. (5), (6), and (11) that

$$\left. \frac{\delta E_{xc}}{\delta n(\vec{r})} \right|_M = \frac{1}{2} \left[\left. \frac{\delta E_{xc}}{\delta n(\vec{r})} \right|_{M+\delta} + \left. \frac{\delta E_{xc}}{\delta n(\vec{r})} \right|_{M-\delta} \right], \quad (12)$$

i.e., the exchange-correlation potential appropriate to an integer number M of electrons in the limit $T \rightarrow 0$ is the average of the left and right functional derivatives of E_{xc} . Again the order of limits is important. The results (11) and (12) are obtained by setting $N = M$ before taking the limit $T \rightarrow 0$. If $T \rightarrow 0$ first, the chemical potential of Eq. (11) will be⁷ either $-I(M)$ or $-A(M)$, depending on whether N approaches M from below or above.

Equations (10) and (11), and indeed all numbered equations of this Letter, are valid for finite systems as well as for an infinite one, because the exact $\epsilon_M(M-1+f)$ and $\epsilon_{M+1}(M+f)$ are in either case independent⁷ of f in the range $0 < f < 1$. This independence is a familiar fact only for infinite systems, which are in any case our main interest here.

Under what conditions will the derivative discontinuity C of Eq. (10) be nonzero? First, C is positive for an isolated atom: For an open-shell atom like hydrogen, $\epsilon_{M+1}(M) = \epsilon_M(M)$ and so $C = I(M) - A(M) > 0$. For a closed-shell atom like helium, von Barth¹¹ has given a convincing argument that $I(M) - A(M) > \epsilon_{M+1}(M) - \epsilon_M(M)$ so that again $C > 0$. Next, when identical atoms are arrayed on a lattice with infinite lattice constant, the resulting crystal will have zero-width Kohn-Sham energy bands at the atomic orbital energies; the open-shell atoms will form Mott insulators^{6,7} in which *all* of the band gap $I(M) - A(M)$ arises from the derivative discontinuity C , while the closed-shell atoms will form conventional insulators in which C contributes *part* of the band gap.

Finally, to discuss C in real crystals, we recall the basic result⁷ of the ground-level density-functional theory for an open system with fluctuating particle number: Regarded as a function of the continuous variable N (the time-averaged electron number), the energy $E(N)$ is a linkage of straight-line segments with possible slope discontinuities at integer values of N . (This is the unique physical continuation^{7,9} of E to noninteger N . By contrast, a differentiable "interpolation" between integer values of N would lead to an improper dissociation limit, i.e., fractionally charged fragments, for a heteronuclear molecule⁷ or an atom desorbed from a metal surface.^{9,13}) The ground level, for N between the integers $M-1$ and M , is a statistical mixture of the $(M-1)$ and M -electron ground states. Thus

$$\frac{\delta E}{\delta n(\vec{r})} = \frac{\partial E(N)}{\partial N} = \begin{cases} -I(M) & (M-1 < N < M) \\ -A(M) & (M < N < M+1). \end{cases} \quad (13)$$

The band gap is then

$$I(M) - A(M) = \left. \frac{\delta E}{\delta n(\vec{r})} \right|_{M+\delta} - \left. \frac{\delta E}{\delta n(\vec{r})} \right|_{M-\delta}. \quad (14)$$

The energy E in Eq. (14) is made up of noninteracting kinetic, electrostatic, and exchange-correlation pieces, as in Eq. (3). The functional derivative of the electrostatic energy, i.e., the electrostatic potential $v(\vec{r}) + \int d^3r' n(\vec{r}')/|\vec{r} - \vec{r}'|$, is manifestly continuous. Thus the band gap $I(M) - A(M)$ arises from discontinuities in the functional derivatives of the noninteracting kinetic and exchange-correlation energies. For a model of noninteracting electrons in a periodic potential, all of the band gap arises from the derivative discontinuity of the kinetic energy, i.e., by Eq. (10), $I(M) - A(M) = \epsilon_{M+1}(M) - \epsilon_M(M)$. But, for interact-

ing electrons, the noninteracting kinetic energy is a clever mathematical construct rather than a physical reality; a situation in which it displays a derivative discontinuity, while the exchange-correlation energy does not, can arise only by the purest accident. Therefore we assert that C , the discontinuity of $\delta E_{xc}/\delta n(\vec{r})$, is nonzero in all real semiconductors or insulators. Whether C is large or small is a separate question, to which we will return later. (Similar comments apply to the spin-density-functional theory,^{1,10,11} which is not otherwise discussed here.)

The origin of the derivative discontinuity of the kinetic energy is of course just the finite change of the highest partially occupied orbital which may occur when the electron number increases through an integer. The derivative discontinuity of the exchange energy E_x has a similar origin, while that of the correlation energy E_c has a more complicated origin. For example, in a Mott insulator^{6,7} the functional derivative of the kinetic energy and $\delta E_x/\delta n(\vec{r})$ are both continuous, and the gap arises from the discontinuity in $\delta E_c/\delta n(\vec{r})$. We have no reason to believe that, in a semiconductor, the discontinuity of $\delta E_c/\delta n(\vec{r})$ would conspire to cancel the nonzero discontinuity of $\delta E_x/\delta n(\vec{r})$.

In a metal, $I(M) = A(M)$ and there is no derivative discontinuity. If open-shell atoms are arrayed on a lattice and the lattice constant is shrunk from infinity toward microscopic values, the gap width $I(M) - A(M)$ eventually goes to zero, while the center of the gap, $-\frac{1}{2}[I(M) + A(M)]$, remains almost invariant. It is a well-known fact¹⁴ that the work function of a metal is approximately the average of the free-atom ionization potential and electron affinity.

Equations (10) and (11), which are exact in the exact Kohn-Sham theory, should be approximately true in approximations like LDA. The LDA makes a continuous approximation to $\delta E_{xc}/\delta n(\vec{r})$ which presumably averages left and right derivatives as in Eq. (12). Typically, in an open-shell [$\epsilon_{M+1}(M) = \epsilon_M(M)$] atom, $\epsilon_M^{\text{LDA}}(M)$ predicts the physical quantity $-\frac{1}{2}[I(M) + A(M)]$ with reasonable accuracy. In Li, for example, where the derivative discontinuity $I - A$ is 4.8 eV, $\epsilon_M^{\text{LDA}}(M) = -2.9$ eV is close to $-\frac{1}{2}(I + A) = -3.0$ eV.

LDA underestimates⁴ the width of the fundamental gap by about 40% in many insulators and semiconductors. How much of this error is intrinsic to the exact Kohn-Sham theory? This question might be answered tentatively by calculations with improved approximations to $\delta E_{xc}/\delta n(\vec{r})$. For

example, a weighted-density calculation by Kerker¹⁵ gave substantial improvement over LDA in the band gap of silicon, but has been questioned¹⁶ for neglect of a potentially important term in $\delta E_{xc}/\delta n(\vec{r})$. A more recent calculation for silicon by von Barth and Car,¹⁷ employing the Langreth-Mehl¹⁸ generalized gradient expansion, gave only a small improvement over LDA for the gap. Similarly, Langreth-Mehl calculations for neon and sodium chloride by Norman and Perdew¹⁹ yield only small improvements over LDA in the fundamental gaps. Atomic calculations¹⁸ indicate that the Langreth-Mehl potential goes a long way toward correcting the small errors of the local-density approximation to $\delta E_{xc}/\delta n(\vec{r})$. We therefore suggest that the error of the LDA band gap (or at least a substantial fraction of it) may be endemic to the Kohn-Sham formalism itself.

It is too much to expect that a local (i.e., multiplicative) potential $\delta E_{xc}/\delta n(\vec{r})$, common to all the one-electron states and generating the exact electron density, can also generate more than one exact feature of the physical band structure. What is needed instead is at least an energy-dependent exchange-correlation potential $\Sigma_{xc}([n]; \epsilon; \vec{r})$. Such a potential is provided by the self-interaction correction^{2,4} (SIC) to the local-density approximation, and SIC band-structure calculations^{19,20} have in fact produced accurate gap widths in rare-gas and alkali halide crystals. More direct approximations to the self-energy⁵ Σ_{xc} have also produced an accurate gap in diamond²¹ and a significant improvement over LDA in silicon.¹⁶ The next fundamental advance in band theory may be the *a priori* construction of simple but accurate energy-dependent potentials.

We are grateful for discussions with U. von Barth, L. J. Bartolotti, W. Kohn, R. G. Parr, L. M. Sander, M. Schlüter, and L. J. Sham. Special thanks go to D. C. Langreth for suggesting the proof of Eq. (11), and for comments on the manuscript. This research was supported in part by the National Science Foundation, Grants

No. DMR80-16117 and No. PHY77-27084, and by the donors of the Petroleum Research Fund administered by the American Chemical Society.

¹W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965); N. D. Mermin, Phys. Rev. **137**, A1441 (1965).

²J. P. Perdew and M. R. Norman, Phys. Rev. B **26**, 5445 (1982).

³J. F. Janak, Phys. Rev. B **18**, 7165 (1978).

⁴J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

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

⁶N. F. Mott, *Metal-Insulator Transitions* (Barnes and Noble, New York, 1974).

⁷J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982), and Bull. Am. Phys. Soc. **28**, 266 (1983), Abstracts AP3 and AP4.

⁸R. G. Parr and L. J. Bartolotti, J. Phys. Chem. **87**, 2810 (1983).

⁹J. P. Perdew and M. Levy, "Many-Body Phenomena at Surfaces," edited by D. C. Langreth and H. Suhl (Academic, New York, to be published); M. Levy, J. P. Perdew, and V. Sahni, unpublished.

¹⁰A. R. Williams and U. von Barth, in *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N. H. March (Plenum, New York, 1983).

¹¹U. von Barth, "Density Functional Theory for Solids," to be published.

¹²L. J. Sham and M. Schlüter, following Letter [Phys. Rev. Lett. **51**, 1888 (1983)].

¹³J. P. Perdew and J. R. Smith, unpublished.

¹⁴E. C. M. Chen, W. E. Wentworth, and J. A. Ayala, J. Chem. Phys. **67**, 2642 (1977).

¹⁵G. P. Kerker, Phys. Rev. B **24**, 3468 (1981).

¹⁶C. S. Wang and W. E. Pickett, Phys. Rev. Lett. **51**, 597 (1983).

¹⁷U. von Barth and R. Car, unpublished calculation referred to in Ref. 11.

¹⁸D. C. Langreth and M. J. Mehl, Phys. Rev. B **28**, 1809 (1983).

¹⁹M. R. Norman and J. P. Perdew, Phys. Rev. B **28**, 2135 (1983).

²⁰R. A. Heaton, J. G. Harrison, and C. C. Lin, Solid State Commun. **41**, 827 (1982).

²¹G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. B **25**, 2867 (1982).