Generalized generalized gradient approximation:
An improved density-functional theory for accurate orbital eigenvalues

Xinlei Hua, Xiaojie Chen, and W. A. Goddard III*

*Materials and Process Simulation Center, Beckman Institute (139-74), Division of Chemistry and Chemical Engineering,
California Institute of Technology, Pasadena, California 91125*

The generalized gradient approximation (GGA) for the exchange functional in conjunction with accurate expressions for the correlation functional have led to numerous applications in which density-functional theory (DFT) provides structures, bond energies, and reaction activation energies in excellent agreement with the most accurate *ab initio* calculations and with the experiment. However, the orbital energies that arise from the Kohn-Sham auxiliary equations of DFT may differ by a factor of 2 from the ionization potentials, indicating that excitation energies and properties involving sums over excited states (nonlinear-optical properties, van der Waals attraction) may be in serious error. We propose herein a generalization of the GGA in which the changes in the functionals due to virtual changes in the orbitals are allowed to differ from the functional used to map the exact density onto the exact energy. Using the simplest version of this generalized GGA we show that orbital energies are within $\sim 5\%$ of the correct values and the long-range behavior has the correct form.
[S0163-1829(97)04120-9]

I. INTRODUCTION

The generalized gradient approximation[1,2] (GGA) for the exchange functional in density-functional theory[3] (DFT) in conjunction with accurate expressions for the correlation functional [e.g., Lee, Yang, and Paar[4] (LYP) or Perdew-Zunger[5] (PZ)] have led to numerous applications[6] in which DFT compares quite well with experiment and with the most accurate *ab initio* calculations for properties such as structure, bond energy, and reaction activation energies. Consequently, DFT is now in routine use for a number of fundamental properties of chemical and physical systems. Despite these successes there remain problems. The orbital energies from the Kohn-Sham auxiliary equations used with DFT differ from the ionization potentials by a factor of about 2. This suggests that electronic excitation energies, and properties depending on a sum over excited electron states (such as nonlinear-optical properties, chemical shielding, and London dispersion), may also be significantly in error.

In this paper we consider a generalization of the GGA in which the functional relationship between density and energy for the virtual changes in the orbitals are allowed to differ from the relationship between the exact density and the exact energy. Using the simplest such generalization of GGA (denoted GGGA), we find that the orbital energies are within $\sim 5\%$ of the exact values and that the long-range potentials of the Kohn-Sham equations have the correct form.

In Sec. II we develop the GGGA function and in Sec. III we apply it to several systems. Section IV discusses various aspects of the GGGA.

II. THE FUNCTIONAL OF DFT

A. DFT review

Hohenberg and Kohn[3] showed that there is a one-to-one mapping between the ground-state density ρ_0 and the ground-state energy E_0 of any system. We shall denote as O_0 this mapping (1) of the exact ground state density of any system onto the exact ground-state energy,

$$\rho_0 \xrightarrow{\mathbf{O}_0} E_0. \tag{1}$$

In order to calculate the correct density ρ_0 of a system, one uses the variational principle to consider all changes in the density $\rho_0 \rightarrow \rho$ allowed by the Pauli principle and select the one leading to the lowest energy. Thus it is useful to define the mapping $\mathbf{O}$ (2) that converts any arbitrary density ρ with the corresponding energy

$$\rho \xrightarrow{\mathbf{O}} E. \tag{2}$$

This density need not correspond to any physical system but it must be allowed by the Pauli principle. In current DFT calculations the functional $\mathbf{O}$ plays the double role of deriving the Euler-Lagrangian equation for calculating the orbitals that lead to ρ_0 and calculating E_0 from ρ_0 once ρ_0 is known.

Using the mapping $\mathbf{O}_x$ for the exchange energy of a uniform electron gas, Kohn and co-worker Sham[3] provided a prescription for including the Pauli principle (N is the representability of the density) in calculating the ground-state properties. Namely, the density is derived from one-particle orbitals that are solutions of a set of Euler-Lagrangian one-particle equations derived from the variational principle. Following convention, we shall use the word energy functional to denote both the mapping $\mathbf{O}$ and the energy E. In all previous forms of DFT, from Thomas and Fermi[7] to Becke[2] and Perdew *et al.*,[1] it has been assumed that $\mathbf{O} \equiv \mathbf{O}_0$.

The GGA (Refs. 1 and 2) extends the description of exchange to include the effect of inhomogeneity in the density gradient, $\gamma = |\nabla_\rho|$. This new mapping for the exchange term $\mathbf{O}_x$ involves first getting the gradient γ and then mapping (ρ, γ) to E. By modeling the ground-state electron density of

atoms, Perdew *et al.*[2] and Becke[2] found an exchange energy functional $E_0^x[\rho]$ that is reasonably accurate for a variety of systems

$$E_0^x[\rho_0]:(\rho_0,\gamma_0)\xrightarrow{F}E_0^x. \qquad (3)$$

Indeed the use of Eq. (3) leads to significant improvements in the total energy calculations. For the variations used to obtain the Euler-Lagrangian equations for ρ_0, GGA assumes $E^x[\rho]\equiv E_0^x[\rho_0]$, as usual. While improving E_{tot}, the GGA does *not* improve the eigenvalues.

B. Kohn-Sham equations

In Kohn-Sham (KS) theory, the energy functional is defined as

$$E_{\text{tot}}[\rho]=T_0[\rho]+E_{\text{Coul}}[\rho]+E_x[\rho]+E_c[\rho]+E_{\text{ext}}[\rho], \qquad (4)$$

which is assumed to hold for arbitrary ρ. $T_0[\rho]$ is the kinetic energy for the antisymmetrized product wave function (independent electrons), $E_{\text{Coul}}[\rho]$ is the classical Coulomb energy term, $E_{\text{ext}}[\rho]$ is potential energy between electrons and ions, $E_x[\rho]$ is the exchange energy, and $E_c[\rho]$ includes all corrections in these quantities due to electron correlation.

The local-density approximation (LDA) uses the expression for a free-electron gas

$$E_x[\rho]=-C_x\int\rho^{4/3}d\mathbf{r}, \qquad (5)$$

and assumes Eq. (5) to be valid even when the density is not constant.

Applying the variational principle to Eq. (4),

$$\left.\frac{\delta E[\rho]}{\delta\rho}\right|_{\rho_0}=0,$$

leads to the Kohn-Sham Euler-Lagrange equation for the optimum orbitals of the ground-state density ρ_0

$$\{-\tfrac{1}{2}\nabla^2+v_{\text{Coul}}(\mathbf{r})+v_x(\mathbf{r})+v_c(\mathbf{r})+v_{\text{ext}}(\mathbf{r})\}\phi_i(\mathbf{r})=\epsilon_i\phi_i(\mathbf{r}), \qquad (6)$$

where

$$\nabla^2v_{\text{Coul}}(\mathbf{r})=-4\pi\rho(\mathbf{r}), \qquad (7)$$

$$\rho(\mathbf{r})=\sum_i f_i|\phi_i(\mathbf{r})|^2, \qquad (8)$$

and

$$v_x(\mathbf{r})=\frac{\delta E_x[\rho]}{\delta\rho(\mathbf{r})}, \qquad (9a)$$

$$v_c(\mathbf{r})=\frac{\delta E_c[\rho]}{\delta\rho(\mathbf{r})}. \qquad (9b)$$

Neglecting electron correlation, $E_c[\rho]=0$ leads to the exchange-only theory, which for LDA is

$$v_x(\mathbf{r})=\frac{\delta E_x[\rho]}{\delta\rho(\mathbf{r})}=-\frac{4}{3}C_x\rho^{1/3}. \qquad (10)$$

LDA predicts fairly accurately the structure for molecules and crystals, but leads to cohesive energies much too large.

C. Generalized gradient approximation

To account for inhomogeneities, the GGA considers the exchange-energy functional to have the form

$$E_x^{\text{GGA}}[\rho]=-C_x\int\rho^{4/3}F(s)d\mathbf{r}, \qquad (11)$$

where

$$s=\frac{|\nabla\rho|}{2k_F\rho} \qquad (12)$$

is the measure of inhomogeneity. Here $k_F=(3\pi^2)^{1/3}\rho^{1/3}$ is the Fermi wave vector for a homogeneous electron gas with density ρ. For the homogeneous electron system, $s=0$, $F(0)=1$, and $E_x^{\text{GGA}}[\rho]=E_x^{\text{LDA}}[\rho]$.

This GGA energy functional leads to

$$v_x^{\text{GGA}}(r)=-\frac{4}{3}C_x\rho(\mathbf{r})^{1/3}F[s(\mathbf{r})]$$

$$-C_x\int\rho^{4/3}F'(s)\frac{\delta s(\mathbf{r}')}{\delta\rho(\mathbf{r})}d\mathbf{r}'. \qquad (13)$$

To determine $F(s)$ we calculated $F(\mathbf{r})=e^{\text{HF}}(r)/e^{\text{LDA}}(\mathbf{r})$ as a function of $s(\mathbf{r})$ for several atoms and ions. Figure 1 compares these results to the $F(s)$ from Becke.[2] Figure 1(a) shows the $(1s)$ (Ref. 2) isoelectronic series, where the E_x only involves the self-interaction energy. We see that the factor $F(s)$ falls nearly onto a single curve. However $F(s)\to0.8$ as $s\to0$, whereas GGA assumes that $F(s)\to1.0$ as $s\to0$. This discrepancy arises because here $s\to0$ corresponds to $r\to0$ where high densities and near constant gradients lead to small s. However, the volume of this region is small, contributing little to E_x.

Figure 1(b) shows the $F(s)$ for H_2 where we see two major branches near $s=0$. One corresponds to the bond midpoint where $\nabla\rho=0$ ($F\to1$) and the other is from regions near the nucleus ($F\to0.85$).

Figure 1(c) shows that carbon atom leads to increased scatter in $F(s)$, but the $F(s)$ fall more or less on one curve.

Figure 1(d) collects together the data for various atoms [including Figs. 1(a), 1(b), and 1(c)] and compares with the analytical form (14) proposed as a universal function $F(s)$ by Becke[2]

$$F(s)=1+\frac{b}{2^{1/3}C_x}\frac{x^2}{[1+6bx\ \ln(x+\sqrt{1+x^2})]} \qquad (14)$$

with

$$x\equiv\frac{|\nabla\rho_\sigma|}{\rho_\sigma^{4/3}}=(48\pi^2)^{1/3}s.$$

Here $\rho_\sigma=0.5\rho$ is the density of spin component σ and $b=0.0042$ a.u. comes from fitting to various atoms and ions. Despite the scatter we see that the Becke formula (14) cap-
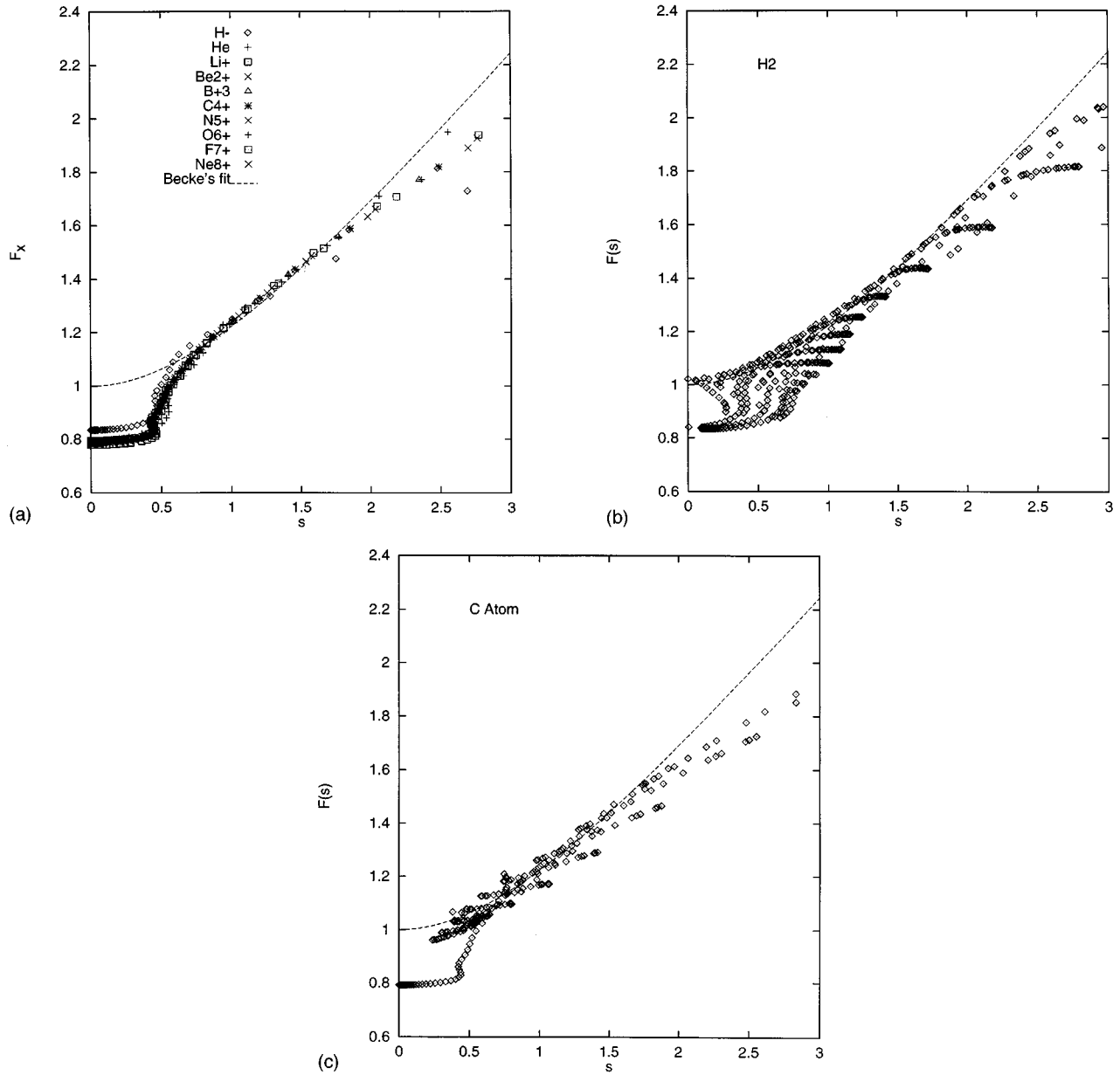
FIG. 1. The scale factor $F(s) = e^{HF}(\mathbf{r})/e^{LDA}(\mathbf{r})$ for $E_x^{GGA}[\rho]$ as a function of $s(\mathbf{r})$. The Becke analytical fit (14) is shown with dashed lines. (a) $1s^2$ configurations H$^-$, He, Li$^+$, Be^{2+}, etc. (b) H$_2$ molecule. (c) Carbon atom.

tures the overall form of the exact results and provides a good description in the region of $s = 0.5$ to 2.0 where ρ is significant.

Equation (14) reproduces the asymptotic behavior for energy density $\epsilon_x(r)$,

$$E_x(r \rightarrow \infty) = -\frac{1}{2} \int d\mathbf{r} \, \frac{\rho(\mathbf{r})}{r}, \tag{15}$$

for atoms,[2] as can be checked by substituting Eq. (16) into Eq. (14),

$$\rho = \rho_0 e^{-\alpha r}. \tag{16}$$

On the other hand, for an atom the potential $v_x(\mathbf{r})$ should satisfy the asymptotic behavior[8]

$$v_x \rightarrow -\frac{1}{r}. \tag{17}$$

However, using Eq. (14) in Eq. (13) with an expontially decaying density, Eq. (16) does not lead to Eq. (17) for $r \rightarrow \infty$. That is, *in GGA v_x does not satisfy the asymptotic behavior for atoms*. In fact using Eq. (16) in Eqs. (14), (11), and (9) and examining the asymptotic behavior, we find $v_x \rightarrow -2.89/\alpha r^2$ at large r.

GGA improves upon LDA significantly for total energy calculations of atoms and molecules.[7] However, Table I shows that the orbital energies (Koopman theorem IP) are low by about 50%. To eliminate any ambiguity caused by possible flaws in the analytical fit (14), we also used directly the $F[s(\mathbf{r})]$ and dF/ds from Hartree-Fock (HF) in Eq. (13), again finding that the eigenvalues[9] are not improved. This

TABLE I. Total energies (Hartrees) and HOMO orbital energies (Hartrees) calculated with GGA (exchange only) and HF. The 6-31G* basis set is used.

	Total energies		HOMO energy		GGA % error
	GGA	HF	GGA	HF	
He	$-2.853\,999$	$-2.855\,160$	$-0.539\,726$	-0.9149	41.0
Be	$-14.560\,105$	$-14.566\,76$	$-0.170\,738$	-0.3013	43.3
C	$-37.588\,045$	$-37.585\,673$	$-0.138\,214$	-0.3420	59.6
O	$-74.672\,631$	$-74.656\,607$	$-0.251\,667$	-0.5758	56.3
Ne	$-128.496\,452$	$-128.474\,402$	$-0.396\,733$	-0.8306	52.2
Ar	$-526.745\,126$	$-526.773\,735$	$-0.330\,767$	-0.5900	43.9
Ave.					49.4

demonstrates that the failure to reproduce the correct eigenvalues is an intrinsic flaw of the exchange functional (11) used with GGA. Since the orbital energies are the basis of describing electronic properties, it is important to correct these errors in calculated eigenvalues.

D. The new functional

The foundation of DFT is Eq. (1); there exists a universal functional that maps the exact density ρ_0 of any system onto the exact energy $E[\rho_0]$. However, to derive the one-particle equation (6), we need to consider *all possible changes in the orbitals for the wave function*, Eq. (2), that may involve nonphysical changes in ρ_0. Thus it is not necessarily the case that $\mathbf{O}=\mathbf{O}_0$.

We propose to generalize the GGA approach by choosing the $\mathbf{O}$ in such a way that $\mathbf{O}\rightarrow\mathbf{O_O}$ as $\rho\rightarrow\rho_0$ and the correct asymptotic behavior is obtained as $\mathbf{r}\rightarrow\infty$ for both the energy density $\epsilon_x(\mathbf{r})$ *and* the one-particle potential $v_x(\mathbf{r})$.

As a simple form for this generalized GGA (GGGA) for the exchange-energy functional, we propose

$$E_x^{\text{GGGA}}[\rho]=-C_x\int\rho^{4/3}F(s_0)e^{\beta(\rho-\rho_0)/\rho_0}d\mathbf{r}. \quad (18)$$

As $\rho\rightarrow\rho_0$, Eq. (18) leads to

$$E_x^{\text{GGGA}}[\rho_0]=-C_x\int\rho_0^{4/3}F(s_0)d\mathbf{r}=E_x^{\text{GGA}}[\rho_0] \quad (19)$$

and

$$v_x^{\text{GGGA}}[\rho_0]=-C_x\rho_0^{1/3}F(s_0)(\tfrac{4}{3}+\beta)=(\tfrac{4}{3}+\beta)\epsilon_x, \quad (20)$$

where

$$\epsilon_x=-C_x\rho_0^{1/3}F(s_0).$$

Thus far we have not specified the spatial dependence of β; it could depend on $\rho_0(\mathbf{r})$, $\nabla\rho_0(\mathbf{r})$, etc.

Since the GGGA functional requires as a generalized parameter the ground-state density ρ_0, which is unknown, the evaluation of the $E_x[\rho]$ is done iteratively. Our procedure is to set $\rho=\rho_0$ *after* taking the variation

$$v_x^{\text{GGGA}}[\rho_0(\mathbf{r})]=\frac{\delta E_x^{\text{GGGA}}[\rho]}{\delta\rho(\mathbf{r})}\bigg|_{\rho_0},$$

and to solve the resulting partial differential equation, (6), self-consistently. This is similar to the procedure for finding the minimum of a function in ordinary calculus; the energy functional is to the KS equation as the function $f(x)$ is to the algebraic equation $f'(x_0)=0$. Thus the KS equation is an ''algebraic equation'' in density space.

From the asymptotic behavior of v_x in Eq. (16) and ϵ_x in Eq. (15), we want to have

$$v_x=2\epsilon_x \quad \text{as} \quad r\rightarrow\infty \quad (21)$$

for atoms. From Eq. (20) this requires that

$$\beta\rightarrow\tfrac{2}{3} \quad \text{as} \quad r\rightarrow\infty. \quad (22)$$

For simplicity, we will in this paper assume that β is a constant,

$$\beta=\tfrac{2}{3}, \quad (23)$$

independent of density ρ_0 or the gradient $\nabla\rho_0$.

The proposed GGGA functional for the exchange energy then becomes

$$E_x[\rho]=-C_x\int\rho^{4/3}F(s_0)e^{2(\rho-\rho_0)/3\rho_0}d\mathbf{r}. \quad (24)$$

Other definitions, e.g.,

$$E_x[\rho]=-C_x\int\rho_0^{4/3}F(s_0)e^{\beta(\rho-\rho_0)/\rho_0}d\mathbf{r} \quad (25)$$

would lead to the same Euler-Lagrangian equation but with different definitions of β. Variation on the $E_x[\rho]$ in Eq. (24) or (25) leads to the relationship

$$v_x(\mathbf{r})=2\epsilon_x(\mathbf{r}). \quad (26)$$

Thus in GGGA both v_x and ϵ_x satisfy the asymptotic relationships (15) and (16).

The key feature of the GGGA functional is that *the gradient term s does not participate in the variation* $\delta E_x/\delta\rho$. This is analogous to $E_{\text{Coul}}=\int\int d\mathbf{r}d\mathbf{r}'\rho(\mathbf{r})\rho(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'|$ where the $1/|\mathbf{r}-\mathbf{r}'|$ term does not participate in variations of ρ.

In the case of a homogeneous electron gas, we have $F(s_0)=1$, leading to

$$\epsilon_x=-C_x\rho^{1/3}$$

and

TABLE II. Total energies (Hartrees) and HOMO orbital energy (Hartrees) using GGGA (exchange only) and HF. The 6-31G* basis set is used.

| | Total energy | | HOMO energy | | |
	GGGA	HF	GGGA	HF	% error
He	−2.849 1	−2.855 16	−0.889 7	−0.914 9	7.8
Be	−14.557 0	−14.566 76	−0.323 8	−0.301 3	7.5
C	−37.559 4	−37.585 67	−0.358 9	−0.342 0	4.9
Ne	−128.440 6	−128.474 40	−0.822 1	−0.830 6	1.0
H_2O	−76.011 392	−76.049 40	−0.489 160	−0.497 470	1.6
C_2H_6	−79.194 130	−79.208 62	−0.520 626	−0.483 115	7.8
glysine	−282.819 449	−282.844 442	−0.414 989	−0.397 651	4.4
Ave.					4.3

$$v_x = -2C_x\rho^{1/3} = -\frac{3}{2}\left(\frac{3}{\pi}\right)^{1/3}\rho^{1/3}. \tag{27}$$

Equation (27) is exactly the exchange potential of a homogeneous electron gas averaged over **k**, the orbital quantum number for the uniform gas, (see the Appendix) obtained originally by Slater.[9,10] Kohn and Sham[3] showed that applying the variational principle to Eq. (4) using Eq. (5) leads to Eq. (6) with Eq. (10) rather than with Eq. (27). That is, KS obtained an exchange function that is $\frac{2}{3}$ of the Slater value, $v_x^{KS} = \alpha v_x^{Slater}$ where $\alpha = \frac{2}{3}$. This led to the X_α version of the LDA in which α was considered as a variable ($\alpha = \frac{2}{3}$ for KS, $\alpha = 1$ for Slater). Both approaches employ the uniform electron-gas expression but in different ways (total energy versus the potential). The GGGA functional unifies these two approaches and resolves this paradox by showing that $\alpha = \frac{2}{3}$ is correct for the total energy (as shown by Kohn and Sham) while $\alpha = 1$ is correct in the one-particle equation (as shown by Slater). For a nonhomogeneous electron system, the gradient factor $F(s_0)$ in GGGA plays the role of setting the scale.

III. RESULTS

We have tested the simplest GGGA functional on several atoms and molecules. The results are listed in Tables II, III, and IV. All structures were fixed at the equilibrium geometry calculated from DFT using the Becke LYP (BLYP) exchange and correlation functional.[2,4]

A. Orbital energies

Table II lists the energies of the highest-occupied molecular orbitals (HOMO), along with the total energy for several atoms and molecules. With GGGA we see that the average error in the orbital energy is about 5% whereas with GGA and LDA it is about 50%. All orbital eigenvalues are improved significantly as shown in Table III. The GGGA functional (24) leads to the same accuracy and sometime better than the more elaborate procedure based on orbital representation of exchange operators, e.g., that of Sahni[8] or Talman and Shadwick.[11]

B. Excitation energy

To estimate the energy gap (i.e., the lowest excitation energy) we used $E_g = \epsilon_{LUMO} - \epsilon_{HOMO}$ where ϵ denotes the orbital energy and LUMO denotes the lowest unoccupied MO. In Table IV, we compare the results calculated with HF, GGGA without correlation, BLYP, and GGGA with correlation. For GGGA with correlation, we used the prescription, $v_c = 2\epsilon_c$, similar to the case of $v_x(\mathbf{r})$. However, the results are not sensitive to this choice. We found that both ϵ_{HOMO} and ϵ_{LUMO} shift down, leading to gaps similar to that by BLYP.

IV. DISCUSSION

A. Self-interaction corrections

Because $E_{Coul}[\rho]$ in Eq. (4) is the energy for the density interacting with itself, it includes the energy of each electron

TABLE III. Eigenvalues (Hartrees) calculated with the LDA, GGA, GGGA, and HF methods. All calculations used the 6-31G* basis set with spherical averaging.

| | | LDA | | GGA | | GGGA | | |
		ϵ_i	% error	ϵ_i	% error	ϵ_i	% error	HF
C	1s	−9.8685	13.0	−10.0190	11.7	−11.0990	2.2	−11.3461
	2s	−0.4298	39.5	−0.4415	37.8	−0.6969	1.8	−0.7099
	2p	−0.1415	58.6	−0.1382	59.6	−0.3590	5.0	−0.3422
Ne	1s	−30.1933	7.8	−30.4461	7.1	−32.1177	2.0	−32.7628
	2s	−1.2105	36.7	−1.2319	35.6	−1.6708	12.6	−1.9120
	2p	−0.3827	53.9	−0.3891	53.2	−0.8221	1.0	−0.8306
Ave.			34.9		33.3		4.1	

TABLE IV. Energy gaps calculated using $\epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}}$. The structure is for the GGA-LYP minimum energy structure. The basis set is 6-31G*.

	HF	GGA-LYP	GGGA-LYP
N_2	21.8	8.6	8.7
C_2H_4	14.9	5.8	5.7
Benzene	12.8	5.1	5.0

interacting with itself [i.e., the self-interaction (SI) energy]. Thus for the hydrogen atom we have

$$E_{\text{Coul}} = \frac{1}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{|\phi_0(r_1)|^2 |\phi_0(r_2)|^2}{r_{12}}, \qquad (28)$$

even though there is only one electron. Consequently $E_x[\rho]$ must be such as to include the SI in order to exactly cancel the SI from the Coulomb term. This requirement, while satisfied by HF, is not met by most DFT functionals. We can see why by comparing the total energy expression for HF and DFT,

$$E_{\text{tot}}^{\text{HF}} = \sum_i^{\text{occ}} f_i \epsilon_i^{\text{HF}} - \frac{1}{2} \sum_{ij}^{\text{occ}} \left(\left\langle \phi_i \phi_j \left| \frac{1}{r_{12}} \right| \phi_i \phi_j \right\rangle \right.$$
$$\left. - \left\langle \phi_i \phi_j \left| \frac{1}{r_{12}} \right| \phi_j \phi_i \right\rangle \right), \qquad (29)$$

$$E_{\text{tot}}^{\text{DFT}} = \sum_i^{\text{occ}} f_i \epsilon_i^{\text{DFT}} - \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 + E_{\text{xc}}[\rho]$$
$$- \int v_{\text{xc}}(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r}, \qquad (30)$$

where f_i is the occupation number for each spin orbital. Consider the He atom where all of $E_x[\rho]$ corresponds to self-interaction (since the orbitals have opposite spin). Ignoring electron correlation we want to have $E_x^{\text{DFT}} = E_x^{\text{HF}}$. Thus requiring $\epsilon_i^{\text{DFT}} = \epsilon_i^{\text{HF}}$ leads to

$$E_x^{\text{DFT}}[\rho_0] - \int d\mathbf{r} \, v_x(\mathbf{r})\rho(\mathbf{r}) = -E_x^{\text{HF}}[\rho_0],$$

which leads to

$$\int d\mathbf{r} \, v_x(\mathbf{r})\rho_0(\mathbf{r}) = 2E_x^{\text{DFT}}[\rho]. \qquad (31)$$

But for LDA we always have

$$\int d\mathbf{r} \, v_x(\mathbf{r})\rho_0(\mathbf{r}) = \tfrac{4}{3} E_x^{\text{LDA}}[\rho] \neq 2E_x^{\text{LDA}}[\rho]. \qquad (32)$$

Thus LDA does not satisfy the requirement. Similar arguments show that GGA also fails to satisfy Eq. (31). On the other hand, by Eq. (26) the GGGA functional does satisfy Eq. (31).
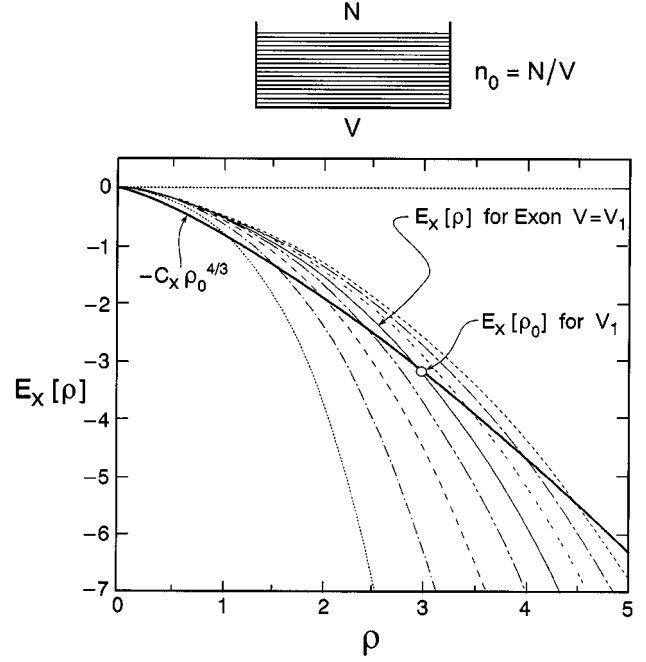

FIG. 2. Illustration of the basis for GGGA. Each dashed curve shows how $E_x[\rho]$ depends on ρ for a fixed box containing a homogeneous electron gas with fixed V. This corresponds to the mapping **O** in Eq. (2), leading to the Slater result (27). For a particular V, say V_1, the optimum wave function leads to a ground-state density ρ_0 and energy E_0 corresponding to a point (ρ_0, E_0^k) on the solid curve. The solid curve associates with the optimum ρ_0 for each physical box V of the exact energy. This corresponds to the mapping $\mathbf{O}_x$ in Eq. (1), leading to the Kohn-Slater result.[10]

B. Energy functional

For the GGGA functionals, the mapping **O** is totally different from $\mathbf{O}_O$. The former is $\sim \rho^{4/3} e^{\beta\rho}$ while the latter is $\rho^{4/3} F(s)$ with $s \sim |\nabla\rho|/\rho^{4/3}$. No foundation has been given for assuming the constraint $\mathbf{O} \equiv \mathbf{O}_O$. Thus, contrary to common assumptions, there need *not* be a universal $E[\rho]$ uniquely giving **O** (though Levy[12] has given one definition of such an **O**).

This is illustrated in Fig. 2 where Eq. (24) is applied to a box containing a uniform electron gas. The ground-state density ρ_0 at each V leads to a unique energy $E_0^x[V]$. These $E_0^x[V]$ for various V lead to the solid curve for how E_0^x depends parametrically on $\rho_0 = N/V$. While this parametric dependance on ρ_0 is the popular one used by LDA [called O_0 in Eq. (1)], it is the bundles of energies as a function of density [called O in Eq. (2)] that should be used in deriving the Euler-Lagrangian equations for various external constraints V. This leads to a different mapping of ρ to E, as indicated by dashed lines.

The exact dynamical equation to give the correct ρ_0 is still elusive. By relaxing the constraint of $\mathbf{O} \equiv \mathbf{O}_o$ we have proposed a class of GGGA functionals and illustrated it with the simplest version. This gives additional degrees of freedom in searching for the right dynamical equation (for ρ_0) and the best total energy functional (for E_0). For example, we can improve the potential by studying the one-electron properties and at the same time improve the total energy

functional by studying the total energy-related properties. This should help the search for the exact density functional that simultaneously leads to the proper dynamical equations for solving for ρ_0 and the total energy functional for E_{tot}. Such an energy functional would provide the best single-electron representation of a many-electron problem.

Summarizing, based on a reexamination of the foundation for DFT we have proposed the generalized GGA functional. Even in its simplest form the GGGA functional leads to a significant improvement in the single-particle eigenvalues. It has several attractive features: (1) it leads to the correct asymptotic behavior, $v_x(r) = 2\epsilon_x(r) \to -1/r$ as $r \to \infty$ for atoms, (2) it satisfies the global constraints for the self-interaction correction, and (3) it resolves the discrepancy on whether the uniform electron gas expression should model the potential (Slater) or the total energy (Kohn-Sham). We anticipate that this GGGA type of functional can be further improved to provide more accurate results for DFT.

of these calculations were carried out on the JPL Cray and the San Diego Supercomputer.

APPENDIX: A DERIVATION OF THE SLATER EXCHANGE POTENTIAL (REF. 10)

For a free-electron gas

$$\omega_{\mathbf{k}\sigma} = \frac{k^2}{2m} - \frac{1}{\Omega} \sum_{\mathbf{q}} \frac{4\pi}{q^2} n_{\mathbf{k}+\mathbf{q},\sigma}. \tag{A1}$$

Summing over $\mathbf{q}$ gives

$$\omega_{\mathbf{k}\sigma} = \frac{k^2}{2m} - \frac{k_F}{2\pi}\left(2 + \frac{k_F^2 - k^2}{kk_F}\ln\left|\frac{k+k_F}{k-k_F}\right|\right) = \frac{k^2}{2m} - \frac{k_F}{2\pi}F\left(\frac{k}{k_F}\right)$$

$$= \frac{k^2}{2m} - v_x(\mathbf{k}). \tag{A2}$$

Averaging over all occupied states, i.e., the Fermi sphere, leads to Eq. (27),

$$v_x = -\frac{k_F}{2\pi}\overline{F\left(\frac{k}{k_F}\right)} = -\frac{3}{2}\left(\frac{3}{\pi}\right)^{1/3}\rho^{1/3} \tag{A3}$$

where we used

$$\overline{F\left(\frac{k}{k_F}\right)} = \frac{3}{4\pi k_F^3}\int_{k<k_F} d\mathbf{k}\, F\left(\frac{k}{k_F}\right) = 3. \tag{A4}$$

*Author to whom correspondence should be addressed.

[1] 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).

[2] A. D. Becke, J. Chem. Phys. **96**, 2155 (1992); **97**, 9173 (1992).

[3] P. Hohenberg and W. Kohn, Phys. Rev. **136**, 864 (1964); W. Kohn and L. J. Sham, *ibid*. **140**, 1133 (1965).

[4] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

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

[6] B. G. Johnson, P. M. W. Gill, and J. A. Pople, J. Chem. Phys. **98**, 5612 (1993).

[7] L. H. Thomas, Proc. Camb. Philos. Soc. **23**, 542 (1927); E. Fermi, Z. Phys. **48**, 73 (1928).

[8] C.-O. Almbladh and U. von Barth, Phys. Rev. B **31**, 3231 (1985); L. J. Sham, *ibid*. **32**, 3876 (1985).

[9] M. K. Harbola and V. Sahni, Phys. Rev. Lett. **62**, 489 (1989); V. Sahni, Y. Li, and M. Harbola, Phys. Rev. A **45**, 1434 (1992).

[10] J. C. Slater, Phys. Rev. **81**, 385 (1951).

[11] J. D. Talman and W. F. Shadwick, Phys. Rev. A **14**, 36 (1976).

[12] M. Levy, Proc. Natl. Acad. Sci. U.S.A. **76**, 6062 (1979).