Nonlocality of the density functional for exchange and correlation: Physical origins and chemical consequences

John P. Perdew and Matthias Ernzerhof
Department of Physics and Quantum Theory Group, Tulane University, New Orleans, Louisiana 70118

Aleš Zupan
Department of Physical and Organic Chemistry, Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia

Kieron Burke
Department of Chemistry, Rutgers University, Camden, New Jersey 08102

(Received 17 June 1997; accepted 14 October 1997)

Gradient corrections to the local spin density approximation for the exchange-correlation energy $E_{\rm xc}$ are increasingly useful in quantum chemistry and solid state physics. We present elementary physical arguments which explain the qualitative dependencies of the exchange and correlation energies upon the local density, local spin polarization, and reduced density gradient. The nearly local behavior of the generalized gradient approximation for $E_{\rm xc}$ at valence-electron densities, due to strong cancellation between the nonlocalities of exchange and correlation, is shared by the exact linear response of the uniform electron gas. We further test and develop our rationale for the chemical and solid-state consequences of gradient corrections. We also partially explain the ''conjointness'' between the exchange energy and the noninteracting kinetic energy, whose generalized gradient approximation is tested here. An appendix presents the full expression for the gradient-corrected correlation potential. © *1998 American Institute of Physics.*
[S0021-9606(98)00504-2]

I. NONLOCALITY AND ITS PHYSICAL ORIGINS

Kohn–Sham spin density functional theory[1] routinely provides useful first-principles predictions for the ground-state structure of many-electron systems.[2] For practical applications, the exchange-correlation energy $E_{\rm xc}[n_\uparrow,n_\downarrow]$ must be approximated. The local spin density approximation (LSD) (Refs. 1,3) has been used successfully for 30 years in solid-state physics. Recently, generalized gradient approximations (GGA's) (Refs. 4–11) have improved upon the accuracy of LSD, and GGA has become a standard method of quantum chemistry. Hybrids of GGA's with traditional wave function methods such as exact exchange[12–16] or configuration interaction[17] are tantalizingly close to the elusive goal of chemical accuracy from first principles. Here we stress that nonempirical GGA's are neither black boxes nor stabs in the dark; they can be understood and intuited. Moreover, LSD can be understood and intuited in the same way, as a special case of GGA.

In the Kohn–Sham approach, the ground-state energy is written as

$$E_v[n_\uparrow,n_\downarrow]=T_s[n_\uparrow,n_\downarrow]+\int d^3r\ n(\mathbf{r})v(\mathbf{r})+U[n]+E_{\rm xc}[n_\uparrow,n_\downarrow], \tag{1}$$

where

$$U[n]=\frac{1}{2}\int d^3r d^3r'\ \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{2}$$

We use atomic units in which $\hbar=e^2=m=1$. T_s denotes the noninteracting kinetic energy of the Kohn–Sham determinant, $v(\mathbf{r})$ is the external potential, $n=n_\uparrow+n_\downarrow$ the ground-state density corresponding to $v(\mathbf{r})$, and $E_{\rm xc}$ accounts for exchange and correlation effects. Generalized gradient approximations to the universal functional $E_{\rm xc}$ are of the form

$$E_{\rm xc}^{\rm GGA}[n_\uparrow,n_\downarrow]=\int d^3r\ f(n_\uparrow,n_\downarrow,\nabla n_\uparrow,\nabla n_\downarrow). \tag{3}$$

In this article we focus on the Perdew–Burke–Ernzerhof (PBE) GGA.[9] This functional is a nonempirical extension of the local spin-density approximation, to which it reduces in the limit where $\nabla n_\uparrow$ and $\nabla n_\downarrow$ both vanish. The principal aim of this article is to explain the physical origin of the nonlocality or density-gradient dependence of the GGA. Our arguments provide a check on the derivations of GGA's, and insights into their structure. This is particularly useful because, while the construction of the PBE is very simple, the inputs to this construction are not all so easily checked. We also discuss the characteristic chemical effects of gradient corrections (Sec. II), and present numerical tests and comparisons of GGA's for the noninteracting kinetic, exchange, and exchange-correlation energies (Sec. III).

A GGA may be represented approximately as

$$E_{\rm xc}^{\rm GGA}[n_\uparrow,n_\downarrow]=\int d^3r\ n\left[\left(-\frac{c}{r_s}\right)F_{\rm xc}(r_s,\zeta,s)\right], \tag{4}$$

where $c=3\alpha/4\pi=0.458\,165$ and $\alpha=(9\pi/4)^{1/3}=1.919\,16$. Here

$$r_s(\mathbf{r})=\left(\frac{3}{4\pi n}\right)^{1/3} \tag{5}$$

is the local density parameter or Seitz radius (roughly the mean interelectronic spacing),

$$\zeta(\mathbf{r}) = \frac{n_\uparrow - n_\downarrow}{n} \tag{6}$$

is the local relative spin polarization, and

$$s(\mathbf{r}) = \frac{|\nabla n|}{2k_F n} = \frac{3}{2\alpha}\,|\nabla r_s| \tag{7}$$

is the local inhomogeneity parameter or reduced density gradient, which measures how fast and how much the density varies on the scale of the local Fermi wavelength $2\pi/k_F = 2\pi r_s/\alpha$. The factor $(-c/r_s)$ in the integrand of Eq. (4) is the exchange energy per particle $\epsilon_x(r_s,\zeta=0)$ of a spin-unpolarized uniform electron gas. The enhancement factor $F_{xc} \geqslant 1$ accounts for the effects of inhomogeneity, spin-polarization, and correlation. It is the analog of $3\alpha/2$ in Slater's $X\alpha$ method,[18] and, like $3\alpha/2$, its variation is bounded and plottable. $F_{xc}(r_s,\zeta,s=0)$ recovers the local spin-density (LSD) approximation.[1] We decompose F_{xc} into an exchange and a correlation contribution

$$F_{xc} = F_x(\zeta,s) + F_c(r_s,\zeta,s), \tag{8}$$

where $F_x > 0$, $F_c > 0$, and $F_c \to 0$ as $r_s \to 0$. Clearly $F_x(\zeta = 0, s=0) = 1$. GGA's typically also include $\nabla\zeta$ contributions to the exchange energy, which vanish for an unpolarized ($\zeta=0$) or fully-polarized ($\zeta=1$) system and are neglected in Eq. (4) and the rest of this section to simplify our qualitative discussion.

The full nonlocality of $E_{xc}[n_\uparrow, n_\downarrow]$ cannot be captured by Eqs. (3) or (4). Because of its simple semilocal form, GGA (like LSD) can be at best a controlled extrapolation away from the limit of slowly-varying density. But this restricted form greatly simplifies the derivation, characterization, and physical understanding of nonlocality and its characteristic effects, and eases the labor required for practical computation. The LSD enhancement $F_{xc}(r_s,\zeta,s=0)$ is unique in principle, since there is a possible system (the uniform electron gas) in which r_s and ζ are constant and thus for which LSD is exact. At least in this sense, there is no unique GGA, since r_s and $s = (3/2\pi)^{1/3}|\nabla r_s|$ cannot both be constant, except in the limit $s \to 0$. Our conservative philosophy of GGA construction is to retain all known correct formal properties of LSD, while adding others.[19,20] Among the welter of possible conditions which could be imposed to construct a GGA, the most natural and important are those already satisfied by LSD, or by the real-space construction of PW91.

The Perdew–Wang 1991 (PW91) (Refs. 7,8,10,11) and Perdew–Burke–Ernzerhof (PBE) (Refs. 8,9) GGA's construct the enhancement factor $F_{xc}(r_s,\zeta,s)$ nonempirically. Both begin from accurate Quantum Monte Carlo calculations for the energy of the uniform electron gas. PW91 relies further upon the density-gradient expansion for the exchange-correlation hole and conserving cutoffs of its spurious long-range parts, while PBE relies upon the gradient coefficient $C_c(r_s=0)$ for the correlation energy and a few exact limits. In the ''GGA made simple'' of PBE, all parameters other than those in LSD are fundamental constants. Figures 1 and 2 display the PBE result for $F_{xc}(r_s,\zeta,s)$ in the physical

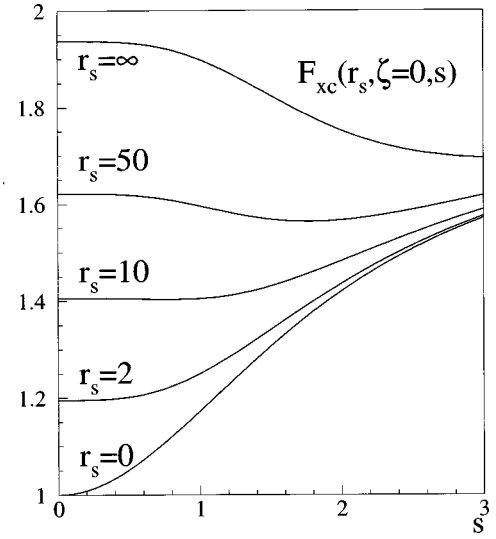

FIG. 1. The $\zeta=0$ enhancement factor of Eq. (4), showing GGA nonlocality or s-dependence for the PBE functional of Ref. 9. $s = |\nabla n|/2k_F n$ is the reduced density gradient or inhomogeneity parameter. The $s \to \infty$ asymptote of all curves is $1+\kappa = 1.804$. The LSD curves would be horizontal straight lines coinciding with the GGA curves at $s=0$.

range $0 \leqslant s \lesssim 3$.[21] (Core electrons have $0 \lesssim s \lesssim 1$; s diverges in the evanescent tail of the density for a finite system.) PW91 is almost indistinguishable from PBE on the scale of Fig. 1, as shown in Ref. 9. The PBE gradient corrections significantly reduce[9,22] the LSD overbinding of molecules, correctly describe the hydrogen bond in ice,[23] and even yield realistic binding energy curves for the rare-gas dimers,[22,24] which other popular GGA's do not bind.

The exchange-correlation energy[25] $E_{xc} = E_x + E_c$ is the sum of two negative contributions. The first is the exchange energy E_x, evaluated from the expectation value of the electron–electron interaction using a Slater determinant of Kohn–Sham orbitals. The second is the correlation energy E_c, which arises from correlations within the interacting wave function and thus is of higher order in the electron–electron interaction. Clearly $F_{xc} \geqslant 0$ in Figs. 1 and 2, as expected. We seek an intuitive understanding of these figures. The corresponding figures for other popular GGA's are presented in Ref. 20.

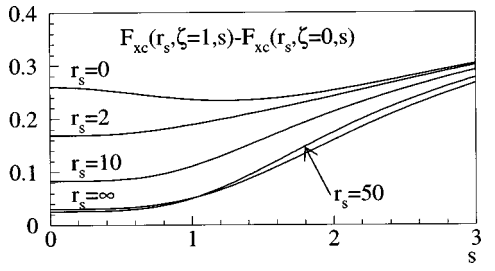

FIG. 2. Same as Fig. 1 for the spin-polarization energy enhancement factor. The value at $r_s=0$ and $s=0$ is $2^{1/3}-1 = 0.260$, from the spin-scaling relation for the exchange energy (Ref. 28). For $0<|\zeta|<1$, $F_{xc}(r_s,\zeta,s) \approx F_{xc}(r_s,0,s) + \zeta^2[F_{xc}(r_s,1,s) - F_{xc}(r_s,0,s)]$. The $s \to \infty$ asymptote of all curves is $(2^{1/3}-1)(1+\kappa) = 0.469$. The LSD curves would be horizontal straight lines coinciding with the GGA curves at $s=0$.

First we consider the dependence of the individual components of the ground-state energy upon density scaling. Temporarily suppressing spin indices, we replace the density n in Eq. (1) by the uniformly-scaled density n_γ,

$$n(\mathbf{r}) \to n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma \mathbf{r}), \tag{9}$$

so that $\zeta(\mathbf{r}) \to \zeta(\gamma\mathbf{r})$, $s(\mathbf{r}) \to s(\gamma\mathbf{r})$, and $r_s(\mathbf{r}) \to \gamma^{-1} r_s(\gamma\mathbf{r})$. Increasing γ makes the density distribution more compact and increases the value of the average density, while conserving the number of electrons. The expectation value of the sum of the kinetic and electron–electron repulsion operators, as a functional of n_γ, can be written as

$$\begin{aligned} T[n_\gamma] + V_{ee}[n_\gamma] &= T_s[n_\gamma] + U[n_\gamma] + E_x[n_\gamma] + E_c[n_\gamma] \\ &= \gamma^2 T_s[n] + \gamma(U[n] + E_x[n]) \\ &\quad + \gamma^2 E_c^{1/\gamma}[n]. \end{aligned} \tag{10}$$

The latter identity follows from known scaling relations[26,27] for each individual term (e.g., $U[n_\gamma] = \gamma U[n]$). $E_c^{1/\gamma}[n]$ is the correlation energy of a system with density n and with electron-electron interaction scaled by γ^{-1}. As $\gamma \to \infty$, the leading term in $E_c^{1/\gamma}$ goes as $1/\gamma^2$. The noninteracting kinetic energy T_s and the exchange energy E_x always scale like γ^2 and γ^1, respectively, while E_c scales in the high-density limit like γ^0. For any density n, the high-density ($\gamma \to \infty$) limit is one in which E_x dominates E_c, and T_s dominates E_x. The form of Eqs. (4) and (8) for the exchange energy at $\zeta = 0$ is consistent with the exchange scaling of Eq. (10).

The exchange energy for the spin-polarized case can then be constructed from that for the unpolarized case via the exact spin-scaling relation,[28]

$$E_x[n_\uparrow, n_\downarrow] = \frac{1}{2} E_x[n_\uparrow, n_\uparrow] + \frac{1}{2} E_x[n_\downarrow, n_\downarrow]. \tag{11}$$

Neglecting $\nabla\zeta$, Eqs. (4) and (11) imply

$$\begin{aligned} F_x(\zeta, s) &= \frac{1}{2}(1+\zeta)^{4/3} F_x(\zeta = 0, s/(1+\zeta)^{1/3}) \\ &\quad + \frac{1}{2}(1-\zeta)^{4/3} F_x(\zeta = 0, s/(1-\zeta)^{1/3}). \end{aligned} \tag{12}$$

The gradient-corrected functional for E_c typically does not include $\nabla\zeta$ terms. The analog of Eq. (11) for T_s is[28]

$$T_s[n_\uparrow, n_\downarrow] = \frac{1}{2} T_s[n_\uparrow, n_\uparrow] + \frac{1}{2} T_s[n_\downarrow, n_\downarrow]. \tag{13}$$

(In Ref. 28, $T_s[n_\sigma, n_\sigma]$ was denoted $T_s[2n_\sigma]$, with a similar notation for E_x.)

Now we consider spin densities $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$ that are not too far from the slowly-varying ($s \ll 1$, etc.) limit,[1] and number-conserving variations of these densities which change the local density parameters r_s, ζ, s. How do E_x and E_c change under such variations? Let us first discuss the noninteracting kinetic energy T_s of the Kohn–Sham determinant. Its second-order gradient expansion approximation[28] (GEA),

$$T_0[n_\uparrow, n_\downarrow] + T_2[n_\uparrow, n_\downarrow] = \int d^3r\, n \left[\frac{3\alpha^2}{10 r_s^2}\right] G(\zeta, s), \tag{14}$$

$$G(\zeta, s) = [(1+\zeta)^{5/3} + (1-\zeta)^{5/3}]/2 + 5s^2/27, \tag{15}$$

scales as in Eqs. (10) and (13), and is arguably its own GGA.[29] Clearly the kinetic energy per particle can be increased by decreasing r_s, increasing $|\zeta|$, or increasing s, as demanded by the Pauli exclusion and uncertainty principles and by scaling.

$|E_x|$ is intimately related to T_s, since both are constructed from the one-electron density matrix of the Kohn–Sham determinant.[30,31] Indeed, these two density functionals have been called ''conjoint''.[32] Let us consider a density variation in which $dr_s \leqslant 0$, $d|\zeta| \geqslant 0$, and $ds \geqslant 0$ everywhere, whence $dT_s \geqslant 0$. With more kinetic energy, the occupied orbitals can dig a more localized and deeper exchange hole, as the uncertainty principle suggests. The first effect of such an increase in T_s is an increase in $|E_x|$, i.e., exchange turns on more strongly when r_s decreases, s increases, or $|\zeta|$ increases, which can also be seen from Eq. (4) and Figs. 1 and 2.

The second effect of any such increase in T_s is to strengthen the unperturbed or noninteracting potential (i.e., the Kohn–Sham potential which holds noninteracting electrons at the spin densities $n_\uparrow$ and $n_\downarrow$). The soft-core, long-range electron-electron repulsion becomes a relatively weaker perturbation,[27,33] thus reducing the ratio E_c/E_x, which tends to zero when T_s tends to infinity. This effect is clearly visible in Figs. 1 and 2: When $r_s \to 0$, $F_{xc}(r_s, \zeta, s) \to F_x(\zeta, s)$, the enhancement factor for exchange alone; when $s \to \infty$, again $F_{xc}(r_s, \zeta, s) \to F_x(\zeta, s)$. Thus correlation turns off relative to exchange in these limits, and also weakens relative to exchange when $|\zeta|$ increases from 0 to 1. Of course, the effects of increasing $|\zeta|$ follow directly from the Pauli exclusion principle, which keeps the electrons apart rather efficiently in the spin-polarized case, thus reducing additional correlation effects. We shall confirm these expectations about E_c and E_x in Eqs. (19) and (20) below. Moreover, $F_{xc}(r_s, \zeta, s)$ decreases with decreasing r_s, so the curves of Fig 1 never cross.

Thus far, we have explained: (1) why the $r_s = 0$ curve in Fig. 1 rises with increasing s, (2) why the curves for successively bigger r_s start higher, rise less steeply, and eventually join the $r_s = 0$ curve as $s \to \infty$, and (3) why these curves never cross. We have also explained (4) why the small-r_s curves of Fig. 2 are positive, and lie above the large-r_s curves.

To explain the other qualitative features of Figs. 1 and 2, especially the $r_s \to \infty$ behaviors, we note that

$$E_{xc} = N \int_0^\infty du\; 4\pi u^2 \frac{\langle n_{xc}(u) \rangle}{2u}, \tag{16}$$

where $N = \int d^3r\, n(\mathbf{r})$ is the electron number and

$$\langle n_{xc}(u) \rangle = \frac{1}{N} \int d^3r\, n(\mathbf{r}) \int \frac{d\Omega_\mathbf{u}}{4\pi} n_{xc}(\mathbf{r}, \mathbf{r}+\mathbf{u}) \tag{17}$$

is the system- and spherical-average of the density $n_{\rm xc}(\mathbf{r},\mathbf{r}+\mathbf{u})$ at $\mathbf{r}+\mathbf{u}$ of the exchange-correlation hole around an electron at $\mathbf{r}$,[8,25] which obeys the sum rule

$$\int_0^\infty du\ 4\pi u^2\langle n_{\rm xc}(u)\rangle = -1. \tag{18}$$

Equation (18) asserts that an electron at $\mathbf{r}$ is missing from the rest of the system. As a consequence of the constraint (18) and of the extra factor of $1/u$ in Eq. (16), the exchange-correlation energy $E_{\rm xc}$ of Eq. (16) typically becomes more negative when the system-averaged hole $\langle n_{\rm xc}(u)\rangle$ becomes deeper at small interelectronic separations u, and less negative when this hole becomes shallower.

The small-u behavior of $n_{\rm xc}(\mathbf{r},\mathbf{r}+\mathbf{u})$ has a quasiuniversal dependence upon the local spin densities at $\mathbf{r}$ and their gradients.[8] In the low-density ($r_s\to\infty$) or strong-coupling limit, the system is maximally correlated in the sense that all other electrons are excluded from the vicinity of a given electron, i.e., we expect that $n_{\rm xc}(\mathbf{r},\mathbf{r}+\mathbf{u})=-n(\mathbf{r}+\mathbf{u})$ for small u. Taylor expansion of this hole to order u^2 and integration of Eq. (17) by parts on $\mathbf{r}$ then yields the low-density limit for the system-averaged hole,

$$\langle n_{\rm xc}(u)\rangle = \frac{1}{N}\int d^3r\ n\ \frac{3}{4\pi r_s^3}\left[-1+\left(\frac{\alpha u}{r_s}\right)^2\frac{2}{3}\,s^2+\cdots\right]. \tag{19}$$

While the on-top ($u=0$) hole is negative, the leading inhomogeneity correction to it for small u is positive, i.e, inhomogeneity makes the hole shallower at small u. This suggests that, in the $r_s\to\infty$ limit, $|E_{\rm xc}|$ decreases with increasing inhomogeneity s. This expectation is confirmed in Fig. 1, where the $r_s\to\infty$ curve actually bends downward with increasing s reflecting the dominance of the correlationlike nonlocality in the low-density limit. Moreover, the result of Eq. (19) is independent of ζ, suggesting that the $r_s\to\infty$ curve of Fig. 2 should lie near zero, as it does for $s\lesssim 1$.

We have already explained why $F_{\rm x}$, the high-density limit of $F_{\rm xc}$, depends as it does upon ζ and s. A more detailed version of this explanation requires the high-density analog of Eq. (19). The gradient expansion for the exchange hole density,[5] expanded to order u^2 and spin-scaled as in Eq. (12), is

$$\langle n_{\rm x}(u)\rangle = \frac{1}{N}\int d^3r\ n\ \frac{3}{4\pi r_s^3}\left\{-\frac{1+\zeta^2}{2}+\left(\frac{\alpha u}{r_s}\right)^2 \times\left[\frac{(1+\zeta)^{8/3}+(1-\zeta)^{8/3}}{20}-\frac{s^2}{27}\,(1+\zeta^2)\right]+\cdots\right\}. \tag{20}$$

Equation (20) can also be found by combining the exact Eqs. (14) and (15) of Ref. 30 for the exchange hole with the gradient expansion of the kinetic-energy density of Ref. 30, integrating by parts, and dropping $\nabla\zeta$ terms.

In the $r_s\to 0$ limit (and also in the $r_s\to\infty$ and $|\zeta|\to 1$ limits, but not otherwise[34]), the LSD on-top or $u=0$ hole is exact. Reducing r_s or increasing $|\zeta|$ deepens the on-top hole and so increases $|E_{\rm xc}|$. For $\zeta=0$, the on-top density varies from $-\frac{1}{2}(3/4\pi r_s^3)$ at $r_s=0$ to $-(3/4\pi r_s^3)$ at $r_s=\infty$, consistent with the increase of $F_{\rm xc}$ with r_s shown in Fig. 1; as the hole deepens on the scale of $(-3/4\pi r_s^3)$, it localizes on the scale of r_s and $E_{\rm xc}$ becomes more local. The inhomogeneity or s^2 term of Eq. (20) is negative, unlike that of Eq. (19). Thus, for sufficiently small r_s, increasing s deepens the hole close to the electron and so increases $|E_{\rm xc}|$, as shown in Fig. 1.

Now restore the dependence of $E_{\rm xc}$ upon fundamental constants, including $a_0=\hbar^2/me^2$, and expand in powers of e^2: $c\to ce^2$ in Eq. (4) and

$$F_{\rm xc}\left(\frac{r_s}{a_0},\zeta,s\right)=F_{\rm x}(\zeta,s)+\left(\frac{r_s}{ca_0}\right)F_{\rm c,2}(\zeta,s)+\cdots, \tag{21}$$

where the terms in $F_{\rm xc}$ of zeroth and first order in e^2 are shown explicitly. The function $F_{\rm c,2}(\zeta,s)$ is well-defined in PBE, and is given by Eq. (9) of Ref. 9. As $s\to\infty$, $F_{\rm c,2}\to 0$ as expected. As $s\to\infty$, $F_{\rm c,2}$ diverges like-ln s, leading to the nonanalytic dependence upon e^2 of the correlation energy for the uniform gas.

Under the uniform scaling [Eq. (9)] of a finite system to the high-density limit ($r_s\to 0$), we find that $E_{\rm c}$ tends to a negative constant,

$$E_{\rm c}\to -\frac{me^4}{\hbar^2}\int d^3r\ nF_{\rm c,2}(\zeta,s), \tag{22}$$

as required by a theorem of Levy [Eq. (4) of Ref. 35]. Moreover, the constant of Eq. (22) seems to be correct for N-electron atoms with $N=2$, 3, 9, 10, and 11 when the nuclear charge $Z\to\infty$.[36]

In the opposite or low-density limit ($r_s\to\infty$), the Thomas–Fermi screening length $\propto(a_0r_s)^{1/2}$ makes an explicit appearance in the uniform-gas correlation energy, and

$$F_{\rm xc}\left(\frac{r_s}{a_0},\zeta,s\right)\to F_{\rm xc}(\infty,\zeta,s)-\sqrt{\frac{a_0}{r_s}}\,K(\zeta,s)+\cdots. \tag{23}$$

At $s=0$, the first and second terms of Eq. (23) are analogs of the classical and zero-point-vibrational energies of a Wigner crystal. Indeed Eq. (18) of Ref. 29 shows that none of the first but half of the second is kinetic energy of correlation. Equation (23) explains why the curves in Fig. 1 approach their $r_s=\infty$ asymptote slowly. The value of the asymptote $F_{\rm xc}(r_s=\infty,\zeta,s=0)$ can be estimated as $0.9/c=1.96$, which accounts for the factor-of-2 variation of the curves in Fig. 1. To make the estimate, replace the strong correlation of the low-density uniform electron gas by the perfect correlation of a close-packed Wigner crystal, and approximate the energy per electron as $-9e^2/10r_s$, the electrostatic energy of a neutral spherical unit cell of radius r_s. The latter is the sum of the the self-repulsion $(3e^2/5r_s)$ of the sphere of uniform positive background and its interaction $(-3e^2/2r_s)$ with the electron at its center.

We have argued that an exchangelike nonlocality dominates as $r_s\to 0$, while an opposite correlationlike nonlocality dominates as $r_s\to\infty$. For small reduced gradients ($s\lesssim 1$), the cancellation between these nonlocalities for valence electron densities ($2\lesssim r_s\lesssim 10$), as shown in Figs. 1 and 2, is

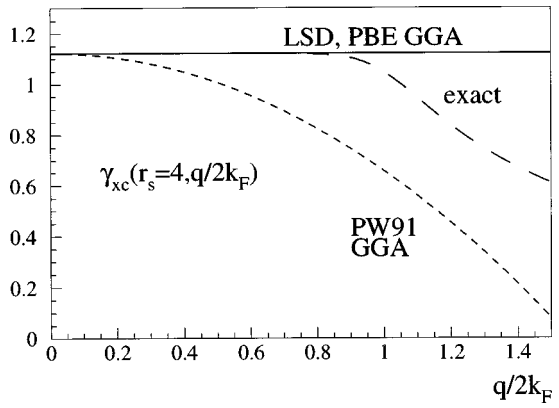

FIG. 3. The linear response function of Eq. (24) for the uniform electron gas ($r_s=4,\zeta=0$), evaluated in LSD and in two different GGA's, and the nearly-exact result of Ref. 39. The curves for $r_s=2$ or 5 are similar. The PW91 curve is also that for the second-order gradient expansion approximation (GEA) for the exchange-correlation energy, which should yield the exact linear response for $q \ll 2k_F$.

striking. The same cancellation occurs in the linear response of the spin-unpolarized uniform electron gas, where the response $\delta v_{xc}(\mathbf{r})=\delta v_{xc}(\mathbf{q})\cos(\mathbf{q}\cdot\mathbf{r})$ of the exchange-correlation potential $v_{xc}(\mathbf{r})=\delta E_{xc}/\delta n(\mathbf{r})$ to a density wave $\delta n(\mathbf{r})=\delta n(\mathbf{q})\cos(\mathbf{q}\cdot\mathbf{r})$ of small amplitude $\delta n(\mathbf{q})$ and wave vector $\mathbf{q}$ is

$$\delta v_{xc}(\mathbf{q})=-\frac{\pi}{k_F^2}\,\gamma_{xc}(r_s,q/2k_F)\,\delta n(\mathbf{q}), \tag{24}$$

where we have returned to atomic units. The coefficient of $\delta n(\mathbf{q})$ in Eq. (24) is simply the Fourier transform with respect to $\mathbf{r}-\mathbf{r}'$ of the second functional derivative $\delta^2 E_{xc}/\delta n(\mathbf{r})\delta n(\mathbf{r}')$. We note that s is small either for a slow density variation ($q\to 0$) of arbitrary amplitude or for a small-amplitude density variation [$\delta n(\mathbf{q})\to 0$] of arbitrary rapidity. The restricted GGA form cannot describe both these limits perfectly. The linear-response limit is physically more important than the slowly-varying limit; for example, the response function γ_{xc} contributes a local field factor $\gamma_{xc}(q/2k_F)^2$ to the screening of a local pseudopotential in a bulk simple metal.[37,38] Figure 3 compares γ_{xc} in the LSD,

$$\gamma_{xc}^{LSD}(r_s)=\gamma_{xc}(r_s,q=0)$$
$$=1+\left(\frac{4\alpha^2}{27}\right)r_s^2\left(2\frac{d\epsilon_c}{dr_s}-r_s\frac{d^2\epsilon_c}{dr_s^2}\right) \tag{25}$$

with $\epsilon_c=(-c/r_s)F_c(r_s,0,0)$, and in the GGA,

$$\gamma_{xc}^{GGA}(r_s,q)=\gamma_{xc}^{LSD}(r_s)-24\pi(3\pi^2)^{1/3}C_{xc}(r_s)(q/2k_F)^2, \tag{26}$$

to the essentially exact γ_{xc} from a Quantum Monte Carlo simulation.[39] In Eq. (26), $C_{xc}(r_s)$ is the coefficient of $|\nabla n|^2/n^{4/3}$ in the second-order gradient expansion of $E_{xc}[n/2,n/2]$. Small $\delta n(\mathbf{q})$ implies small s, so every GGA reduces to its own second-order gradient expansion in this linear-response limit. The PBE GGA was constructed so that its gradient coefficient C_x for exchange cancels that for correlation, and thus its linear response reduces to LSD, which is nearly exact for the uniform gas. Simple hole models based upon Eqs. (18) and (20), such as that of Ref. 30, effectively also yield the PBE C_x, which is 1.78 times bigger than the exact C_x appropriate to the slowly-varying limit.[40] The PW91 GGA was constrained to have what is believed[41] to be the correct gradient coefficient $C_{xc}(r_s)$ (Ref. 42 for exchange and correlation, but provides a less realistic linear response than the PBE GGA. In fact, the QMC simulation shows no trace of a nonzero C_{xc} for $\zeta=0$ and $2\leqslant r_s\leqslant 5$. While the gradient expansion of the hole density $\langle n_{xc}(u)\rangle$ at small u yields physically-useful information, the gradient expansion for the energy often does not, as a consequence of the long range of the Coulomb interaction.[41,43]

At the exchange-only level, the exact γ_x is known and has the small-q expansion[44]

$$\gamma_x=1+\frac{5}{9}\left(\frac{q}{2k_F}\right)^2+\frac{73}{225}\left(\frac{q}{2k_F}\right)^4+\cdots. \tag{27}$$

The GGA form cannot produce a q^4 term, but the PBE GGA emulates the exact γ_x for $q/2k_F\lesssim 1$ through enhancement of the q^2 term by the factor 1.78.

The second-order gradient contribution to the correlation energy scales like exchange, i.e., for $s\ll 1$, $F_c=F_c(r_s,\zeta,0)-\mu\phi s^2$, where $\mu=0.219\,51$ (Ref. 45) and (in the PBE of Ref. 9) $F_x=F_x(\zeta,0)+\mu\phi s^2$, where

$$\phi=[(1+\zeta)^{2/3}+(1-\zeta)^{2/3}]/2. \tag{28}$$

Thus these gradient coefficients cancel for all ζ and r_s. Because exchange must dominate correlation (i.e., $F_c\to 0$) as $r_s\to 0$, and also because the correlation energy must be negative ($F_c>0$), the second-order gradient expansion for the correlation energy [i.e., the formula for F_c shown above Eq. (28)] must be generalized. In the PBE of Ref. 9 this is achieved by writing

$$E_c^{GGA}[n_\uparrow,n_\downarrow]=\int d^3r\; n[\epsilon_c(r_s,\zeta)+H(r_s,\zeta,t)]$$
$$=\int d^3r\; n\left[\left(-\frac{c}{r_s}\right)F_c(r_s,\zeta,s)\right], \tag{29}$$

where $\epsilon_c(r_s,\zeta)$ is the correlation energy per particle of a uniform electron gas[3] and

$$t=(\pi\alpha/4)^{1/2}s/\phi r_s^{1/2}. \tag{30}$$

Figure 1 shows that, in the important range $r_s\leqslant 10$, the nonlocalities of exchange and correlation cancel for $t\leqslant 1/3$. As $t\to 0$, $H\to(3/\pi^2)\mu\phi^3 t^2$, but as $t\to\infty$, $H\to-\epsilon_c(r_s,\zeta)$.

Finally, because $\langle n_{xc}(u)\rangle$ cannot be much deeper on the scale of $\langle n\rangle$ than it is in the lowdensity uniform gas (Lieb–Oxford bound[7,19,46]), the gradient expansion for the exchange energy [i.e., the formula for F_x shown above Eq. (28)] must also be generalized. The PBE generalization is

$$F_x(\zeta=0,s)=1+\kappa-\kappa/(1+\mu s^2/\kappa), \tag{31}$$

which tends to $1+\kappa$ as $s\to\infty$. The parameter $\kappa=0.804$ is the largest value consistent with $F_{xc}(r_s,\zeta,s)\leqslant 2.273$ for all r_s, ζ, and s, a condition which ensures satisfaction of the Lieb–Oxford bound on the exchange-correlation energy for all

densities. The LSD $F_{xc}(r_s,\zeta,0)$ automatically satisfies this bound. Note that this is *not* a bound on the exchange-correlation energy density. Because of an integration by parts, GGA is not expected to predict the energy density correctly.[47,48]

The conjointness[32] of T_s and $|E_x|$ is most quantitatively evident in the limit $s\to 0$, where Eq. (31) becomes $1+0.21951s^2$ and $G(\zeta=0,s)$ of Eq. (15) becomes $1+0.18519s^2$.

Gradient corrections to LSD capture much of the nonlocality of $E_{xc}[n_\uparrow,n_\downarrow]$. The residual nonlocality can manifest as a diffuse large-u component of the exact exchange-correlation hole which has no universal dependence upon $n_\uparrow(\mathbf{r})$, $n_\downarrow(\mathbf{r})$, $\nabla n_\uparrow(\mathbf{r})$, and $\nabla n_\downarrow(\mathbf{r})$, and is significant in the metal surface energy[43] and in the valence-electron energy of a multiplybonded molecule, as discussed in Sec. III and Ref. 49. Under this manifestation (which is inauspicious for both LSD and GGA), and because of the approximate cancellation between the nonlocalities of E_x^{GGA} and E_c^{GGA} for valence-electron densities, it can happen that E_{xc}^{LSD} is more accurate than E_{xc}^{GGA}. When an electron is shared by two or more isolated subsystems,[50–52] the diffuse component can have infinite range. This diffuse component is missed by PW91 (due to the sharpness of its real-space cutoffs), and also by PBE (which assumes that the Lieb–Oxford bound is close in the low-density limit), but could be emulated by reduction of the parameter κ in Eq. (31).

There is another way in which LSD and GGA can fail. When the Kohn–Sham noninteracting ground-state wavefunction is not a single Slater determinant but a linear combination of several degenerate ones, the LSD and GGA exchange-correlation holes can fail even close to the electron ($u\to 0$), a situation which is sometimes rectified by symmetry breaking.[53] Near-degeneracies which are very different from those in the uniform electron gas can also cause this problem.[54]

II. CHEMICAL CONSEQUENCES OF NONLOCALITY

We pass now to the consequences of nonlocality. The near locality of $E_{xc}[n_\uparrow,n_\downarrow]$ for $s\lesssim 1$ and $2\lesssim r_s\lesssim 10$ partly explains the remarkable success of LSD for many solid state applications (and its less impressive performance for molecules and atoms, where there are regions with $s\gg 1$). Because of the strong cancellation between the GGA nonlocalities of exchange and correlation in solid-state applications, the residue of this cancellation is not uniformly better than LSD. The spin polarization energy of Fig. 2 is also rather local, i.e., only weakly s-dependent for $s\lesssim 1$ (and for $s\lesssim 3$ at high densities), explaining the success of LSD for magnetism. Where the gradient corrections have an important influence on magnetism [e.g., in solid Fe (Ref. 55)], it is often an indirect influence via the crystal structure or lattice constant.

There are few physical consequences of the correlation-like nonlocality of F_{xc}, which is seen in Fig. 1 only for $r_s\gtrsim 10$. One theoretical consequence is that gradient corrections should oppose the charge-density-wave and Wigner-crystal instabilities of the low-density uniform electron gas.

Chemical consequences of the exchangelike nonlocality of F_{xc} for $r_s\lesssim 10$ have been discussed in Ref. 21. Gradient corrections to LSD lower atomization energies[10] of molecules and solids, favor open structures relative to close-packed ones,[56–59] raise barriers to the formation of highly-bonded transition-state complexes in chemical reactions,[60] and usually stretch equilibrium bond lengths. As a numerical example, consider the hydrogen exchange reaction $H+H_2\to H_3\to H_2+H$. The transition state H_3 has more bonds than the reactants. The barrier heights are 0.8 eV at the Hartree–Fock level, -0.11 eV in LSD, 0.15 eV in the PBE GGA, and 0.42 eV experimentally, as discussed in Ref. 21. The negative LSD barrier height means that H_3 is incorrectly predicted to be stable by LSD. The GGA value given above is found with the exact spin scaling of Eq. (11); the approximate spin scaling of Eq. (12) has little effect (yielding a barrier of 0.21 eV).

To understand these effects, we[21] have defined and studied distributions and energy-weighted averages $\langle r_s\rangle$, $\langle|\zeta|\rangle$, and $\langle s\rangle$ of the electron density parameters, and found that: (1) $\langle r_s\rangle$, and $\langle s\rangle$ are larger for the separated atoms than for the molecule or solid, and larger for the separated reactants than for the highly-bonded transition-state complex. (2) $\langle r_s\rangle$ and $\langle s\rangle$ increase when we stretch a bond or lattice constant. (3) Gradient corrections favor both density inhomogeneity (greater $\langle s\rangle$) and density contraction (smaller $\langle r_s\rangle$), driving an infinitesimal process forward if a certain inequality is satisfied,

$$\frac{d\langle s\rangle}{\langle s\rangle}\geqslant P\,\frac{d\langle r_s\rangle}{2\langle r_s\rangle}+Qd\langle|\zeta|\rangle, \tag{32}$$

where $P\approx 1$ and $Q\approx 0$. In typical processes, $d\langle r_s\rangle$ and $d\langle s\rangle$ have the same sign and so compete—a second reason why LSD works as well as it does. In some cases (e.g., H_2) gradient corrections actually shorten bond lengths and reduce $\langle s\rangle$, when by doing so they can produce a sufficient decrease of $\langle r_s\rangle$.

It is principally the *sign* of the gradient corrections to E_{xc} which determines whether they will drive a process forward. For example, a hypothetical $F_{xc}=1+\nu s^2$ for any positive ν implies Eq. (32) with $P=1$ and $Q=0$. However, it is the *size* of these corrections (or in this example the size of ν) which determines how strongly the process will be favored. The strong nonlocality of the exchange energy (the $r_s=0$ curve in Fig. 1) helps to explain why the exchange-only or Hartree–Fock approximation underbinds atoms in molecules, overly favors open structures relative to close-packed ones, and overestimates barriers to the formation of highly-bonded transition-state complexes. While LSD likes bonds too much,[60] the Hartree–Fock approximation likes them too little.

In Ref. 21, we tested and confirmed the inequality Eq. (32) for several physical processes including atomizations, bond stretchings, and transition-state fragmentations, using Hartree–Fock and LSD densities. In the present study, we have further confirmed this inequality for atomization of the

TABLE I. Average density parameters $\langle r_s\rangle$, $\langle s\rangle$, and $\langle|\zeta|\rangle$ (as defined in Ref. 21) for molecules at the experimental geometry and in the separated-atom limit, as calculated self-consistently within the PBE GGA. N is the total number of electrons.

		Molecule			Atoms		
System	N	$\langle r_s\rangle$	$\langle s\rangle$	$\langle\|\zeta\|\rangle$	$\langle r_s\rangle$	$\langle s\rangle$	$\langle\|\zeta\|\rangle$
H_2	2	1.618	0.892	0.000	2.186	1.092	1.000
B_2	10	0.688	0.821	0.166	0.702	0.851	0.180
LiH	4	0.995	0.889	0.000	1.064	0.928	0.437
CH_4	10	0.777	0.758	0.000	0.874	0.862	0.491
NH_3	10	0.663	0.739	0.000	0.720	0.813	0.463
OH	9	0.537	0.738	0.141	0.551	0.762	0.295
H_2O	10	0.566	0.724	0.000	0.596	0.771	0.340
HF	10	0.484	0.710	0.000	0.496	0.732	0.212
Li_2	6	0.896	0.891	0.000	0.908	0.905	0.242
LiF	12	0.509	0.733	0.000	0.522	0.750	0.158
C_2H_2	14	0.649	0.762	0.000	0.696	0.830	0.347
HCN	14	0.590	0.753	0.000	0.620	0.803	0.341
CO	14	0.533	0.743	0.000	0.550	0.777	0.249
N_2	14	0.540	0.746	0.000	0.559	0.781	0.337
NO	15	0.516	0.736	0.080	0.529	0.765	0.286
O_2	16	0.495	0.727	0.123	0.504	0.752	0.237
F_2	18	0.454	0.711	0.000	0.457	0.723	0.138

seventeen molecules of Table I. This table also shows the averages $\langle r_s\rangle$, $\langle s\rangle$, and $\langle|\zeta|\rangle$ for these molecules at the experimental geometry[9,49] and in the separated-atom limit. These calculations were made self-consistently within the PBE GGA (whose correlation potential is presented in Appendix A), using a Gaussian orbital basis and computer program described in Appendix B. All open-shell systems (including the separated atoms) were described by the single determinant with maximum z-component of spin, and all other symmetries were allowed to break. The exact spin-scaling of Eq. (11) was used instead of the approximate spin scaling of Eq. (12).

While the changes $d\langle r_s\rangle$, $d\langle|\zeta|\rangle$, and $d\langle s\rangle$ of Eq. (32) are due to the valence electrons, the averages $\langle r_s\rangle$, $\langle|\zeta|\rangle$, and $\langle s\rangle$ in Table I were computed for all electrons, including the cores. As the atomic number Z increases, $\langle r_s\rangle$ tends to shrink faster than $\langle s\rangle$ as a consequence of the increasing effect of the core. Thus, one might want to remove the core contribution, as in pseudopotential theory, before testing the inequality Eq. (32). However, for the elements with $Z\leqslant 9$ in Table I, we have not found this to be necessary.

In keeping with our analysis above, the self-consistent effect of gradient corrections upon the all-electron density is typically to increase $\langle s\rangle$ and reduce $\langle r_s\rangle$ relative to LSD. We have found this to be the case for all the systems of Table I. In particular there is a GGA core contraction which helps to explain why pseudopotentials constructed for LSD should be reconstructed for GGA.[23] Analysis of x-ray diffraction[61] on silicon shows that the core density is more realistic in GGA than in LSD.

The electron density parameters r_s, ζ, and s are closely related to standard chemical concepts. For example, $s(r)$ shows the shell structure of an atom, rising as one passes outward through an electronic shell and falling as one passes outward through an intershell region.[62] Also $s(\mathbf{r})$ vanishes at the bond center of a molecule or solid. The changes in the averages upon atomization also have a chemical meaning which can be gleaned from Table I: We observe that $4Nd\langle r_s\rangle$ and $4Nd\langle s\rangle$ are roughly equal to the number of strong bonds broken (with a weak bond counting as a fraction), and $Nd\langle|\zeta|\rangle$ is the number of unpaired spins created by the atomization. Here N is the total number of electrons in the molecule.

III. NUMERICAL STUDY OF SEMILOCAL APPROXIMATIONS

Generalized gradient approximations are sometimes called semilocal approximations. The form of the right-hand side of Eq. (3) is clearly too restricted to represent the Hartree electrostatic energy of Eq. (2), which is fully nonlocal, but Eq. (3) is more appropriate for the noninteracting kinetic (T_s), exchange (E_x), and exchange-correlation (E_{xc}) energies. Here we shall examine just how accurate these approximations are for molecules and their atomization energies.

The kinetic energy T_s is best treated exactly by the Kohn–Sham method.[1] Its second-order gradient expansion T_0+T_2 of Eqs. (14) and (15) is in a sense[29] its own GGA. Arguments[32] have also been given for a conjoint functional $T_s^{\rm conj}$, in which $G(\zeta=0,s)$ of Eq. (14) is replaced by $F_x(\zeta=0,s)\equiv F_{xc}(r_s=0,\zeta=0,s)$. We have not found much evidence to favor one choice over the other as the preferred GGA for the kinetic energy. (For a numerical study of the sixth-order gradient expansion $T_0+T_2+T_4+T_6$ in solids, see Ref. 63.)

Tables II and III show our numerical results, using the PBE $F_{xc}(r_s,\zeta,s)$, for molecules at experimental equilibrium geometries and their atomization energies. All densities have been evaluated self-consistently within the Kohn–Sham implementation of the PBE GGA. Instead of the approximate spin scalings of T_s and E_x implied by Eqs. (15) and (12), we have used the exact spin scalings of Eqs. (13) and (11), as usual.

The ''exact'' T_s and E_x of Tables II and III were evaluated from the occupied Kohn–Sham orbitals by summing their individual kinetic energies and by evaluation of the Fock integral, respectively. The ''exact'' E_{xc} for each molecule or atom was found by starting from the exact nonrelativistic total energy E and then subtracting $E^{\rm PBE}-E_{xc}^{\rm PBE}$. The exact nonrelativistic total energies for the atoms were taken from Refs. 64,65, and those for the molecules were found from those for the atoms by subtraction of the experimental atomization energies, with the zero-point vibrational energy removed. For references and a detailed discussion (including LSD and PW91 energies), see Ref. 49.

There have been previous studies of T_s for molecules[66] and its change upon atomization,[67] but we have found that accurate calculations of T_s and dT_s require much larger basis sets (Appendix B) than are required for accurate calculations of E, E_x, or E_{xc}. We believe that our values of T_s have a basis-set error of 0.01 hartree or less.

TABLE II. Numerical test of generalized gradient approximations for the noninteracting kinetic (T_s), exchange (E_x), and exchange-correlation (E_{xc}) energies of molecules. For T_s, two GGA's are presented: the second-order gradient expansion T_0+T_2 and the conjoint functional T_s^{conj}. Experimental geometries and self-consistent PBE densities have been employed. (All energies in hartrees.)

Molecule	T_0+T_2	T_S^{conj}	T_s	E_x^{PBE}	E_x	E_{xc}^{PBE}	E_{xc}
H_2	1.125	1.107	1.140	−0.648	−0.657	−0.691	−0.698
B_2	49.010	48.784	49.172	−7.539	−7.525	−7.795	−7.872
LiH	8.003	7.934	7.978	−2.105	−2.125	−2.188	−2.212
CH_4	40.141	40.050	40.276	−6.536	−6.576	−6.836	−6.883
NH_3	55.911	55.846	56.301	−7.634	−7.647	−7.948	−7.996
OH	74.779	74.734	75.470	−8.518	−8.525	−8.797	−8.846
H_2O	75.477	75.462	76.150	−8.917	−8.910	−9.241	−9.292
HF	99.242	99.300	100.137	−10.385	−10.378	−10.720	−10.779
Li_2	14.922	14.797	14.857	−3.508	−3.542	−3.631	−3.671
LiF	106.189	106.206	107.051	−11.917	−11.910	−12.293	−12.370
C_2H_2	76.572	76.418	76.951	−10.961	−10.971	−11.391	−11.470
HCN	92.390	92.250	93.025	−12.048	−12.030	−12.488	−12.567
CO	112.013	111.910	112.897	−13.313	−13.289	−13.762	−13.847
N_2	108.242	108.115	109.115	−13.128	−13.094	−13.580	−13.665
NO	128.198	128.107	129.434	−14.732	−14.680	−15.222	−15.300
O_2	148.369	148.312	149.843	−16.358	−16.290	−16.887	−16.958
F_2	196.729	196.832	198.892	−19.951	−19.872	−20.564	−20.661

Table III shows that the performance of the GGA for atomization energies tends to improve as one passes from T_s to E_x, and from E_x to E_{xc}, as we had expected. The sum rule[8,25] and on-top hole[34,54,68] arguments provide especially powerful constraints upon the exact and GGA exchange-correlation energies when the exchange-correlation hole is well-localized around its electron, as it is in these systems.[49]

Upon atomization (Table III), the exact kinetic energy becomes less positive and the exact exchange-correlation energy becomes less negative. The exact exchange energy typically also becomes less negative, but not for the multiply-bonded molecules N_2, NO, O_2, and F_2. For these molecules, the exact exchange hole has a rather complicated shape and a long range not properly imitated by GGA; as a result, GGA gives rather poor results for dT_s and dE_x, and this error is partially propagated into dE_{xc}. These are the molecules whose atomization energies benefit the most from exact-exchange mixing.[12–16]

TABLE III. Numerical test of generalized gradient approximations for the changes upon atomization of the noninteracting kinetic (dT_s), exchange (dE_x), and exchange-correlation (dE_{xc}) energies. See caption of Table II. (All energies in hartrees.)

Molecule	$d(T_0+T_2)$	dT_s^{conj}	dT_s	dE_x^{PBE}	dE_x	dE_{xc}^{PBE}	dE_{xc}
H_2	−0.114	−0.107	−0.149	0.044	0.042	0.075	0.083
B_2	−0.101	−0.144	−0.137	0.092	0.019	0.123	0.114
LiH	−0.038	−0.046	−0.063	0.050	0.046	0.077	0.084
CH_4	−0.600	−0.681	−0.622	0.312	0.298	0.445	0.445
NH_3	−0.106	−0.204	−0.432	0.197	0.153	0.316	0.309
OH	−0.003	−0.046	−0.165	0.071	0.044	0.112	0.107
H_2O	−0.195	−0.273	−0.350	0.168	0.121	0.248	0.245
HF	−0.236	−0.280	−0.221	0.112	0.079	0.152	0.150
Li_2	−0.004	−0.021	−0.019	0.002	−0.002	0.024	0.031
LiF	−0.229	−0.297	−0.211	0.194	0.146	0.229	0.229
C_2H_2	−0.526	−0.678	−0.614	0.326	0.258	0.456	0.441
HCN	−0.080	−0.237	−0.473	0.199	0.101	0.311	0.288
CO	−0.226	−0.351	−0.414	0.153	0.066	0.225	0.209
N_2	0.332	0.171	−0.348	0.067	−0.051	0.161	0.138
NO	0.359	0.226	−0.240	0.056	−0.066	0.135	0.104
O_2	0.172	0.065	−0.223	0.069	−0.058	0.131	0.094
F_2	0.273	0.209	−0.050	0.009	−0.111	0.043	0.019

While the error of the PBE for E_{xc} of molecules is only about 0.7%, it would be about seven times smaller for the Perdew–Wang 1991 (Refs. 7,10,11) GGA. However, the smaller total-energy error of the more complex PW91 does not buy any improvement[9] in the change dE_{xc} upon atomization, in comparison with the PBE ''GGA made simple'' of Ref. 9.

IV. FINAL REMARKS

Density functional theory is a simplification of many-particle quantum mechanics. Its approximations ought to be and are derivable and understandable from the fundamental principles of quantum mechanics.

A universal functional $E_{xc}[n_\uparrow, n_\downarrow]$, which glues one atom to another, links seemingly diverse systems (such as atoms, molecules, and solids) with different bonding characters (covalent, ionic, metallic, hydrogen, and van der Waals). Just as nature seamlessly links atoms to solids via clusters, etc., our density functional description ought to be seamless. Even the uniform electron gas is almost represented in nature's data set: The success of the stabilized jellium model[69] reaffirms that the exchange-correlation energy of the valence electrons in a simple metallic crystal like Na or Al is essentially that of a uniform gas.

Density functionals illuminate our understanding of real systems, and the study of real systems can reflect back upon our understanding of the functionals. We believe that this process would work most effectively if calculated properties were routinely reported at several levels of density functional theory, including the LSD and nonempirical GGA levels.

ACKNOWLEDGMENTS

One of us (J.P.P.) thanks Cyrus Umrigar, Peter Feibelman, Don Hamann, and Frank J. Tipler for asking some of the questions which this article attempts to answer, and Jingsong He for helpful discussions. This work was supported in part by the National Science Foundation under Grant No. DMR95-21353 (J.P.P., K.B., and M.E.), the Deutsche Forschungsgemeinschaft (M.E.), and the Slovenian Ministry of Science and Technology under Grant No. J1-8665 (A.Z.).

APPENDIX A: GGA CORRELATION POTENTIAL

Self-consistent implementations of density functionals $E_{\rm xc}[n_\uparrow,n_\downarrow]$ typically require the exchange-correlation potential or functional derivative $\delta E_{\rm xc}/\delta n_\sigma(\mathbf{r})$. Complete expressions for the GGA exchange potential can be found in Eqs. (16) and (24)–(26) of Ref. 5. Expressions for the GGA correlation potential in the form needed for PW91 or PBE were given in Eq. (33) of Ref. 7, but this reference is not widely available and a term that belongs in its Eq. (33) was inadvertently dropped[70] from that presentation.

The functional derivative of Eq. (3) is

$$\delta E_{\rm xc}^{\rm GGA}/\delta n_\sigma(\mathbf{r}) = v_{\rm xc}^{\sigma,{\rm GGA}}([n_\uparrow,n_\downarrow];\mathbf{r}) = \frac{\partial f}{\partial n_\sigma} - \nabla\cdot\left(\frac{\partial f}{\partial \nabla n_\sigma}\right). \tag{A1}$$

Applying Eq. (A1) to Eq. (29) for any function H, we find the GGA correlation potential

$$\begin{aligned} v_{\rm c}^{\sigma,{\rm GGA}} = {} & \epsilon_{\rm c} - \frac{r_s}{3}\frac{\partial \epsilon_{\rm c}}{\partial r_s} - (\zeta - {\rm sgn}[\sigma])\frac{\partial \epsilon_{\rm c}}{\partial \zeta} + H - \frac{r_s}{3}\frac{\partial H}{\partial r_s} + \frac{r_s}{3}\,t\,\frac{\partial^2 H}{\partial r_s \partial t} + \frac{t}{6}\frac{\partial H}{\partial t} + \frac{7t^3}{6}\frac{\partial}{\partial t}\left(t^{-1}\frac{\partial H}{\partial t}\right) \\ & - (\zeta - {\rm sgn}[\sigma])\left(\frac{\partial H}{\partial \zeta} - \frac{\phi' t}{\phi}\frac{\partial H}{\partial t}\right) - \frac{\nabla n\cdot\nabla\zeta}{(2\phi k_s)^2 n}\left\{\left(t^{-1}\frac{\partial^2 H}{\partial \zeta \partial t}\right) - \frac{\phi'}{\phi}\left[2\left(t^{-1}\frac{\partial H}{\partial t}\right) + t\frac{\partial}{\partial t}\left(t^{-1}\frac{\partial H}{\partial t}\right)\right]\right\} \\ & - \frac{\nabla n\cdot\nabla|\nabla n|}{(2\phi k_s)^3 n^2}\frac{\partial}{\partial t}\left(t^{-1}\frac{\partial H}{\partial t}\right) - \frac{\nabla^2 n}{(2\phi k_s)^2 n}\left(t^{-1}\frac{\partial H}{\partial t}\right), \end{aligned} \tag{A2}$$

where sgn $[\sigma]$ is $+1$ for $\sigma=\uparrow$ and -1 for $\sigma=\downarrow$. At the nucleus, $\nabla^2 n$ diverges like $(2/r)dn/dr$, and so do the GGA exchange and correlation potentials.

Subroutines which evaluate the PBE exchange-correlation ''energy density'' and potential are available by e-mail from perdew@mailhost.tcs.tulane.edu or on the Web at http://www.phy.tulane.edu/~kieron/dft.html

APPENDIX B: BASIS SETS AND OTHER COMPUTATIONAL FACTORS

The calculations reported in this article were done with a modified version of the CADPAC program.[71]. This program uses Gaussian basis sets, and the integrations over the density are performed numerically. The exponents of most of the basis sets in the CADPAC basis set library are optimized for Hartree–Fock calculations on atoms. Additional polarization functions are added to provide enough flexibility for calculations on molecules. These basis sets are usually contracted, and the contraction coefficients are obtained from Hartree–Fock calculations. In the course of this work we found that, at the level of triple zeta plus two sets of polarization functions, these basis sets in general are flexible enough to give atomization energies with an accuracy of the order of 0.001 hartree compared to the basis set limit. However, we also found that the basis sets described above are often not appropriate for the accurate calculation of the kinetic energy. We found it necessary to decontract the basis sets, in order to obtain accurate values for the kinetic energy in DFT calculations. The final basis sets used in this work have been checked by optimizing the geometries of the molecules in LSD exchange-only calculations and by checking the virial relation $T_s=-E$. In all cases we found that this relation is satisfied with an error of 0.01 hartree or less.

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[2] *Modern Density Functional Theory: A Tool for Chemistry*, edited by J. M. Seminario and P. Politzer (Elsevier, Amsterdam, 1995).
[3] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
[4] D. C. Langreth and M. J. Mehl, Phys. Rev. B **28**, 1809 (1983).
[5] J. P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).
[6] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[7] J. P. Perdew, in *Electronic Structure of Solids' 91*, edited by P. Ziesche and H. Eschrig (Akademie, Berlin, 1991).
[8] J. P. Perdew, K. Burke, and Y. Wang, Phys. Rev. B **54**, 16533 (1996).
[9] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); **78**, 1396 (1997).
[10] 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); **48**, 4978 (1993).
[11] K. Burke, J. P. Perdew, and Y. Wang, in *Electronic Density Functional Theory: Recent Progress and New Directions*, edited by J. F. Dobson, G. Vignale, and M. P. Das (Plenum, New York, 1997).
[12] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[13] M. Ernzerhof, Chem. Phys. Lett. **263**, 499 (1996).
[14] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
[15] K. Burke, M. Ernzerhof, and J. P. Perdew, Chem. Phys. Lett. **265**, 115 (1997).
[16] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **106**, 1063 (1997).
[17] M. Noland, E. L. Coitiño, and D. G. Truhlar, J. Phys. Chem. **101**, 1193 (1997).
[18] J. C. Slater, *The Self-Consistent Field for Molecules and Solids* (McGraw–Hill, New York, 1974).
[19] M. Levy and J. P. Perdew, Phys. Rev. B **48**, 11638 (1993).
[20] J. P. Perdew and K. Burke, Int. J. Quantum Chem. **57**, 309 (1996).

[21] A. Zupan, K. Burke, M. Ernzerhof, and J. P. Perdew, J. Chem. Phys. **106**, 10184 (1997).
[22] D. C. Patton, D. V. Porezag, and M. R. Pederson, Phys. Rev. B **55**, 7454 (1997); D. C. Patton and M. R. Pederson, Phys. Rev. A **56**, R2495 (1997).
[23] D. R. Hamann, Phys. Rev. B **55**, R10157 (1997), and references therein.
[24] Y. Zhang, W. Pan, and W. Yang, J. Chem. Phys. **107**, 7921 (1997).
[25] O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
[26] M. Levy and J. P. Perdew, Phys. Rev. A **32**, 2010 (1985).
[27] A. Görling and M. Levy, Phys. Rev. B **47**, 13105 (1993).
[28] G. L. Oliver and J. P. Perdew, Phys. Rev. A **20**, 397 (1979).
[29] J. P. Perdew, Phys. Lett. A **165**, 79 (1992).
[30] A. D. Becke, Int. J. Quantum Chem. **23**, 1915 (1983).
[31] J. A. Alonso and L. A. Girifalco, Phys. Rev. B **17**, 3735 (1978).
[32] H. Lee, C. Lee, and R. G. Parr, Phys. Rev. A **44**, 768 (1991).
[33] A. Görling and M. Levy, Phys. Rev. A **50**, 196 (1994).
[34] K. Burke, J. P. Perdew, and D. C. Langreth, Phys. Rev. Lett. **73**, 1283 (1994).
[35] M. Levy, Int. J. Quantum Chem. **23**, 617 (1989).
[36] S. Ivanov and M. Levy (unpublished).
[37] C. Fiolhais, J. P. Perdew, S. Q. Armster, J. M. MacLaren, and M. Brajczewska, Phys. Rev. B **51**, 14001 (1995); **53**, 13193 (1996).
[38] L. Pollack, J. P. Perdew, J. He, F. Nogueira, and C. Fiolhais, Phys. Rev. B **55**, 15544 (1997).
[39] S. Moroni, D. M. Ceperley, and G. Senatore, Phys. Rev. Lett. **75**, 689 (1995).
[40] P. S. Svendsen and U. von Barth, Phys. Rev. B **54**, 17402 (1996); M. Springer, P. S. Svendsen, and U. von Barth, **54**, 17392 (1996).
[41] D. C. Langreth and S. H. Vosko, Phys. Rev. Lett. **59**, 497 (1987); **60**, 1984 (1988).
[42] M. Rasolt and D. J. W. Geldart, Phys. Rev. B **34**, 1325 (1986).
[43] D. C. Langreth and J. P. Perdew, Phys. Rev. B **21**, 5469 (1980).
[44] E. Engel and S. H. Vosko, Phys. Rev. B **42**, 4940 (1990).
[45] S.-K. Ma and K. A. Brueckner, Phys. Rev. **165**, 18 (1968).
[46] E. H. Lieb and S. Oxford, Int. J. Quantum Chem. **19**, 427 (1981).
[47] K. Burke, F. G. Cruz, and K. C. Lam (unpublished).
[48] K. C. Lam, F. G. Cruz, and K. Burke (unpublished).
[49] M. Ernzerhof, J. P. Perdew, and K. Burke, Int. J. Quantum Chem. **64**, 285 (1997).
[50] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982); J. P. Perdew, in *Density Functional Methods in Physics*, edited by R. M. Dreizler and J. da Providencia (Plenum, New York, 1985).
[51] R. Merkle, A. Savin, and H. Preuss, J. Chem. Phys. **97**, 9216 (1992).
[52] J. P. Perdew and M. Ernzerhof, in *Electronic Density Functional Theory: Recent Progress and New Directions*, edited by J. F. Dobson, G. Vignale, and M. P. Das (Plenum, New York, 1997).
[53] J. P. Perdew, A. Savin, and K. Burke, Phys. Rev. A **51**, 4531 (1995).
[54] J. P. Perdew, M. Ernzerhof, K. Burke, and A. Savin, Int. J. Quantum Chem. **61**, 197 (1997).
[55] P. Bagno, O. Jepsen, and O. Gunnarsson, Phys. Rev. B **40**, 1997 (1989).
[56] N. Moll, M. Bockstedte, M. Fuchs, E. Pehlke, and M. Scheffler, Phys. Rev. B **52**, 2550 (1995).
[57] D. R. Hamann, Phys. Rev. Lett. **76**, 660 (1996).
[58] R. O. Jones and G. Seifert, Phys. Rev. Lett. **79**, 443 (1997).
[59] A. Zupan, P. Blaha, K. Schwarz, and J. P. Perdew (unpublished).
[60] L. Deng, T. Ziegler, and L. Fan, J. Chem. Phys. **99**, 3823 (1993).
[61] J. M. Zuo, P. Blaha, and K. Schwarz, J. Phys.: Condens. Matter **9**, 7541 (1997).
[62] A. Zupan, J. P. Perdew, K. Burke, and M. Causà, Int. J. Quantum Chem. **61**, 287 (1997).
[63] Z. Yan, J. P. Perdew, T. Korhonen, and P. Ziesche, Phys. Rev. A **55**, 4601 (1997).
[64] E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, and C. F. Fischer, Phys. Rev. A **44**, 7071 (1991).
[65] S. J. Chakravorty, S. R. Gwaltney, and E. R. Davidson, Phys. Rev. A **47**, 3694 (1993).
[66] A. J. Thakkar, Phys. Rev. A **46**, 6920 (1992).
[67] J. P. Perdew, M. Levy, G. S. Painter, S. Wei, and J. B. Lagowski, Phys. Rev. B **37**, 838 (1988).
[68] K. Burke, J. P. Perdew, and M. Ernzerhof (in preparation).
[69] J. P. Perdew, H. Q. Tran, and E. D. Smith, Phys. Rev. B **42**, 11627 (1990).
[70] P. Söderlind (private communication).
[71] CADPAC6: The Cambridge Analytical Derivatives Package Issue 6.0 Cambridge, 1995. A suite for quantum chemistry programs developed by R. D. Amos, with contributions from I. L. Alberts, J. S. Andrews, S. M. Colwell, N. C. Handy, D. Jayatilaka, P. J. Knowles, R. Kobayashi, G. J. Laming, A. M. Lee, P. E. Maslen, C. W. Murray, P. Palmieri, J. E. Rice, J. Sanz, E. D. Simandiras, A. J. Stone, M.-D. Su, and D. J. Tozer.