

Spurious fractional charge on dissociated atoms: Pervasive and resilient self-interaction error of common density functionals

Adrienn Ruzsinszky and John P. Perdew

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

Gábor I. Csonka^{a)}

Department of Inorganic Chemistry, Budapest University of Technology and Economics, H-1521 Budapest, Hungary

Oleg A. Vydrov and Gustavo E. Scuseria

Department of Chemistry, Rice University, Houston, Texas 77005

(Received 1 September 2006; accepted 11 October 2006; published online 20 November 2006)

Semilocal density functional approximations for the exchange-correlation energy can improperly dissociate a neutral molecule XY ($Y \neq X$) to fractionally charged fragments $X^{+q} \cdots Y^{-q}$ with an energy significantly lower than $X^0 \cdots Y^0$. For example, NaCl can dissociate to $\text{Na}^{+0.4} \cdots \text{Cl}^{-0.4}$. Generally, q is positive when the lowest-unoccupied orbital energy of atom Y^0 lies below the highest-occupied orbital energy of atom X^0 . The first 24 open sp -shell atoms of the Periodic Table can form 276 distinct unlike pairs XY , and in the local spin density approximation 174 of these display fractional-charge dissociation. Finding these lowest-energy solutions with standard quantum chemistry codes, however, requires special care. Self-interaction-corrected (SIC) semilocal approximations are exact for one-electron systems and also reduce the spurious fractional charge q . The original SIC of Perdew and Zunger typically reduces q to 0. A scaled-down SIC with better equilibrium properties sometimes fails to reduce q all the way to 0. The desideratum of “many-electron self-interaction freedom” is introduced as a generalization of the one-electron concept. © 2006 American Institute of Physics. [DOI: 10.1063/1.2387954]

I. INTRODUCTION

The Kohn-Sham density functional theory^{1,2} reduces the many-electron ground-state problem to a tractable self-consistent one-electron form. The many-electron exchange-correlation energy

$$E_{\text{xc}}[n_{\uparrow}, n_{\downarrow}] = \int d\mathbf{r} \quad g([n_{\uparrow}, n_{\downarrow}]; \mathbf{r}) \quad (1)$$

is often approximated by semilocal density functionals, in which the energy density g at position \mathbf{r} depends upon the electron spin densities $n_{\uparrow}(\mathbf{r}')$ and $n_{\downarrow}(\mathbf{r}')$ (and perhaps on the occupied orbitals) only in an infinitesimal neighborhood of \mathbf{r} . A ladder of nonempirical semilocal approximations shows increasing usefulness for chemistry. For example,³ the mean absolute total error of 223 molecular atomization energies is 121.8 kcal/mol in the local spin density approximation (LSDA),^{1,4,5} 22.2 kcal/mol in the generalized gradient approximation (GGA) of Perdew, Burke, and Ernzerhof (PBE),⁶ and 5.8 kcal/mol in the meta-GGA of Tao, Perdew, Staroverov, and Scuseria (TPSS).⁷ For comparison, the mean atomization energy is 1180 kcal/mol, so the relative errors are very small. These approximations also predict satisfactory equilibrium geometries and vibration frequencies for most molecules,³ satisfactory dipole moments at equi-

lium,⁸ and satisfactory binding energy curves for many molecules X_2 .⁹

The atomization energies above are computed with respect to separated neutral atoms. It is not widely realized that some binding energy curves or energy surfaces can show much bigger errors, because the semilocal approximations can dissociate a neutral molecule XY ($Y \neq X$), to fragments $X^{+q} \cdots Y^{-q}$ with spurious fractional charge $q > 0$, and with an energy below that of $X^0 \cdots Y^0$. For example, in LSDA NaCl can dissociate to $\text{Na}^{+0.4} \cdots \text{Cl}^{-0.4}$ with an energy lowering of about 25 kcal/mol relative to $\text{Na}^0 \cdots \text{Cl}^0$. A similar effect will occur in a polyatomic molecule built in part from atoms X and Y .

This effect was predicted long ago for the $X\alpha$ method by Slater¹⁰ and for the LSDA by Perdew *et al.*¹¹ and others.^{12–14} Since the total energy of $X^{+q'} \cdots Y^{-q'}$ in the dissociation limit is just the sum of the energies of the atomic fragments $X^{+q'}$ and $Y^{-q'}$, the energy-minimizing q can be found from purely atomic calculations. The earlier numerical work was done with the atomic programs of that time, in which the densities of p and d orbitals were replaced by their spherical averages and fractional occupation numbers were allowed. In these calculations, the noninteracting kinetic energy and density of each atom are written as $\sum_i f_i \langle \psi_i | -\nabla^2/2 | \psi_i \rangle$ and $\sum_i f_i |\psi_i(\mathbf{r})|^2$, respectively, where $0 \leq f_i \leq 1$ is the occupation of the localized atomic Kohn-Sham orbital ψ_i . For each considered set of occupation numbers, self-consistent calculations are made for each atom, and the set that minimizes the sum of the two

^{a)}Electronic mail: csonkagi@gmail.com

atomic energies is chosen. The resulting total energy is¹³ (apart from the inessential spherical averaging) just what would be predicted by a LSDA calculation for a diatomic molecule at infinite bond length. Figure 6 of Ref. 13 predicts the LSDA dissociation limit for NaCl to be $\text{Na}^{+0.4}\dots\text{Cl}^{-0.4}$. At fixed total electron number, the total LSDA energy minimizes with only 0.6 (not 1) electrons in the Na $3s\uparrow$ orbital and with 2.4 (not 2) electrons in the Cl $3p\uparrow$ orbital, making these two orbitals degenerate at the Fermi level. To go beyond semilocal approximations to SIC or Hartree-Fock requires extensions¹⁵ of these approximations to fractional electron number.

Over subsequent years, several searches¹⁶ for this effect were attempted by plotting binding energy curves for the molecule NaCl using standard molecular codes,¹⁷ but the solutions approached $\text{Na}^0\dots\text{Cl}^0$ as the bond length approached infinity. (We speculate that dissociating NaCl may show multiple Aufbau-respecting solutions or near solutions of the Kohn-Sham equations, although only one is of lowest total energy.) It is only in the past year that we found the lowest-energy, fractional-charge solution. (For technical details see Sec. III.) In the past year Dutoi and Head-Gordon¹⁸ also observed the effect for $XY=\text{LiF}$, LiCl , NaF , and NaCl using a GGA and its hybrids with exact exchange, and showed how the binding energy curves approach a limiting energy below that of $X^0\dots Y^0$. Ossowski *et al.*¹⁹ observed the effect in a molecular calculation for LiF and its nonoccurrence for H_2O , in 2003. As we will show here, the effect is widespread and not limited to strongly ionic molecules.

From the simplest perspective, the spurious fractional-charge dissociation arises from the self-interaction error inherent to all the semilocal functionals,^{13–15} i.e., from their failure to be exact for all one-electron densities. As we will also show here, however, for a correct dissociation it is not enough to make the functionals exact for all one-electron densities. More importantly, one must ensure that the energy of an open system with fractional electron number N has a realistic linear variation¹¹ with N between any two integers $M-1$ and M . Equivalently, one must make the highest-occupied and lowest-unoccupied orbital energies for the atoms realistic, in a sense to be discussed in the next section.

A related problem of semilocal density functionals is somewhat better known.^{20–25} Charged radicals such as H_2^+ also show very unrealistic binding energy curves. In the case of the one-electron system H_2^+ ,²⁰ having half an electron on each dissociated proton is not spurious, but the energy associated with a fractional electron number is still seriously too low as an unambiguous consequence of self-interaction error. For compact subsystems such as separated atomic fragments, the semilocal functionals are typically more accurate for an integer than for a fractional electron number;²⁶ for the latter case, they tend to make the energy relatively too low, as a consequence of an exchange-correlation hole sum rule discussed in Secs. 4A and 4C of Ref. 13: Over a separated subsystem, the semilocal hole density always integrates to -1 , while the exact and Perdew-Zunger self-interaction-corrected hole densities integrate to -1 for an integer electron number but to a less-negative value for a fractional number.^{13,15}

II. ROLE OF THE ATOMIC ORBITAL ENERGIES

Let ε_i and f_i be the orbital energy and occupation number of orbital $\psi_i(\mathbf{r})$, respectively. Then the total energy change δE arising from infinitesimal changes δf_i of the occupation numbers is

$$\delta E = \sum_i \varepsilon_i \delta f_i. \quad (2)$$

Equation (2) was derived by Slater¹⁰ within the $X\alpha$ method and by Janak²⁷ within the approximate or exact Kohn-Sham density functional theory. For the PZ self-interaction correction (without off-diagonal Lagrange multipliers), it was derived by Perdew and Zunger.¹⁵

Physically meaningful energies E are typically ground-state energies for which the orbitals obey an “Aufbau principle,” with $f_i=1$ for $\varepsilon_i < \mu$, $0 \leq f_i \leq 1$ for $\varepsilon_i = \mu$, and $f_i=0$ for $\varepsilon_i > \mu$, where μ is the chemical potential. Let q' ($1 > q' \geq 0$) be any positive fractional charge on X (not necessarily the energy-minimizing value q). Then, in a process where $\delta q > 0$ electrons are transferred from a separated atomic fragment $X^{+q'}$ to a separated atomic fragment $Y^{-q'}$, the energy change is

$$\begin{aligned} E(X^{+[q'+\delta q]} \dots Y^{-[q'+\delta q]}) - E(X^{+q'} \dots Y^{-q'}) \\ = \varepsilon^{\text{HO}}(X^{+q'})(-\delta q) + \varepsilon^{\text{LU}}(Y^{-q'}) (\delta q) \\ = [\varepsilon^{\text{LU}}(Y^{-q'}) - \varepsilon^{\text{HO}}(X^{+q'})] \delta q, \end{aligned} \quad (3)$$

by Eq. (2). Here $\varepsilon^{\text{HO}}(X^{+q'})$ is the highest-occupied ($q'=0$) or partially occupied orbital energy of $X^{+q'}$, and $\varepsilon^{\text{LU}}(Y^{-q'})$ is the lowest-unoccupied ($q'=0$) or partly unoccupied orbital energy of $Y^{-q'}$. This process lowers the energy if

$$\varepsilon^{\text{LU}}(Y^{-q'}) < \varepsilon^{\text{HO}}(X^{+q'}), \quad (4)$$

and continues until the two sides of Eq. (4) equalize at $q'=q$.

For neutral atom pairs (X^0 and Y^0 , $q'=0$), condition (4) can be satisfied within the LSDA, leading to a spurious fractional-charge dissociation. Since the orbital energies change little in higher-level semilocal approximations, a LSDA-like fractional charge is also expected within the GGA and meta-GGA. A close look at Figs. 1 and 2 shows that gradient corrections tend to open the highest-occupied–lowest-unoccupied (HO-LU) gap slightly.

The situation is different in the *exact* Kohn-Sham density functional theory,¹¹ in which for $0 \leq q' < 1$

$$\varepsilon^{\text{HO}}(X^{+q'}) = -I(X^0), \quad (5)$$

$$\varepsilon^{\text{LU}}(Y^{-q'}) = -A(Y^0), \quad (6)$$

where $I(X^0)$ is the first ionization energy of neutral atom X , and $A(Y^0)$ is the electron affinity of neutral atom Y . Since the largest electron affinity of the periodic Table [$A(\text{Cl})=3.62$ eV] is smaller than the smallest ionization energy [$I(\text{Cs})=3.89$ eV], Eq. (4) is never satisfied in the exact Kohn-Sham theory, in which neutral molecules always dissociate to neutral atoms.

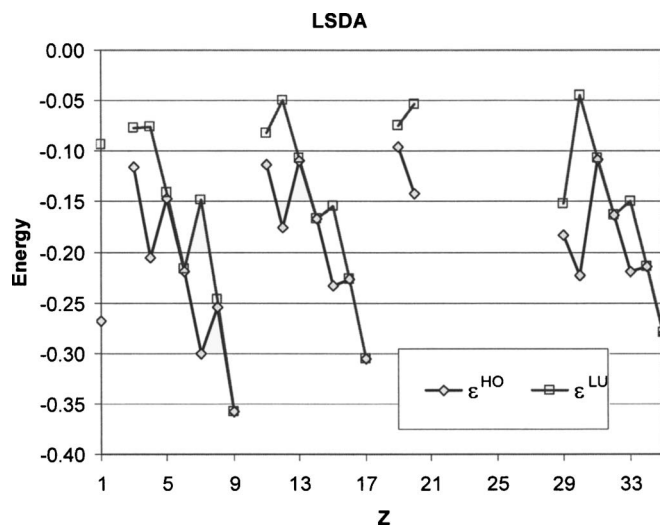


FIG. 1. LSDA/6-311G(*d*) highest-occupied orbital energy ϵ^{HO} (hartree) and lowest-unoccupied orbital energy ϵ^{LU} (hartree) vs Z for the 24 open *sp*-shell atoms with atomic number $Z < 36$. The LSDA ϵ^{HO} is one-half to two-thirds of minus the experimental ionization energy I , and follows the periodic trend of $-I$. Even for Ca ($Z=20$), the LU is not a *d* orbital. The Cu and Zn ($Z=29$ and 30 , respectively) atoms were calculated with the 6-311+G(*d*) basis set. Note that adding diffuse functions to the basis set causes relatively small (less than 0.003 hartree) changes except for Cu and Zn. In Zn, however, it decreases ϵ^{LU} by 0.16 hartree.

Combining Eqs. (5) and (6) with Eq. (3) shows that the exact total energy varies linearly with q' in the range $0 \leq q' < 1$, consistent with an ensemble interpretation of the fractional electron number.¹¹ The derivative of the exact total energy with respect to the electron number, and the exact exchange-correlation potential, can jump discontinuously at integer electron numbers, a prediction of Ref. 11 that has received much recent confirmation.^{28–30} Within semilocal approximations, the exact straight-line segments of the total

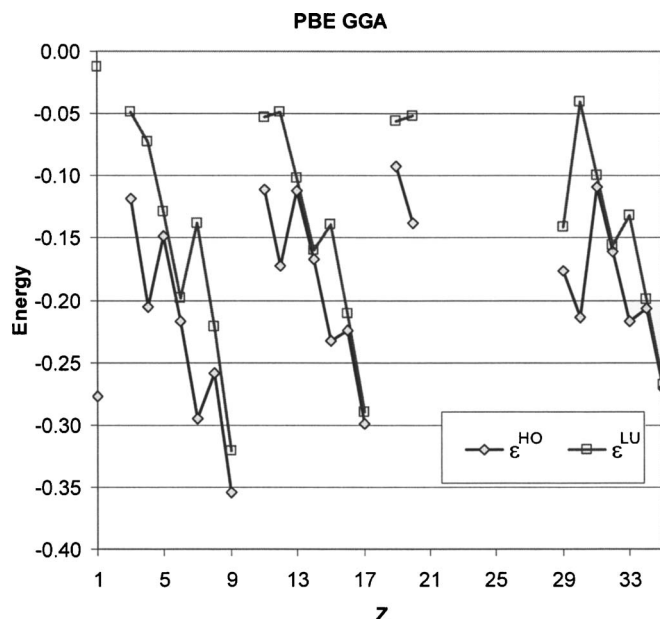


FIG. 2. PBE GGA/6-311G(*d*) highest-occupied orbital energy ϵ^{HO} (hartree) and lowest-unoccupied orbital energy ϵ^{LU} (hartree) vs Z for 24 open *sp*-shell atoms with atomic number $Z < 36$. The Cu and Zn ($Z=29$ and 30 , respectively) atoms were calculated with the 6-311+G(*d*) basis set.

energy versus number between integer numbers are replaced by segments resembling upward-curving parabolas, but a nearly linear behavior is recovered after the Perdew-Zunger¹⁵ self-interaction correction (see Fig. 6 of Ref. 13).

III. COMPUTATIONAL DETAILS

All calculations were carried out with the 03 and developmental versions of the GAUSSIAN molecular program.¹⁷ Except where otherwise noted, the 6-311G(*d*) basis set and the ultrafine grid were used. For 22 of our 24 atoms the ϵ^{HO} and ϵ^{LU} energies are insensitive (energy effect less than 0.003 hartree) to the addition of diffuse functions to the basis set. The two exceptions are Cu and Zn, for which the addition of the diffuse functions led to considerable stabilization in ϵ^{HO} and especially in ϵ^{LU} . A particularly large effect (−0.16 hartree) was found for ϵ^{LU} (Zn).

The densities of *p* and *d* orbitals were nonspherical. The flavor of LSDA was SVWN5.⁵ The Hartree-Fock (HF), LSDA, PBE, TPSS, and SIC (Refs. 15 and 31) calculations were spin-unrestricted. SIC was implemented as in Refs. 32 and 33. Because we avoid approximations made in the original Perdew-Zunger SIC solutions¹⁵ for atoms (neglect of off-diagonal Lagrange multipliers and spherical averaging of orbital densities), and because the original solutions are only for LSDA SIC, conclusions based upon them may be only qualitatively applicable to ours. For the SIC calculations (available only in the developmental version), the large unpruned (99 590) grid and the Guess=Alter keyword were used. The latter is important for avoiding the higher-energy ionic state at the dissociation limit in SIC and HF. For example, near the critical bond length R_c of Sec. V, one might need to switch an electron from an orbital localized on *Y* to a nearly degenerate orbital localized on *X*.

Most importantly, we used quadratic convergence (SCF =qc) and (for the LSDA, PBE, and TPSS calculations) Stable=Opt to help find the solution of the lowest possible energy. Without this, the molecule *XY* might tend to the higher-energy $X^0 \cdots Y^0$ as the bond length $R \rightarrow \infty$.

IV. Pervasiveness of Fractional-Charge Dissociation

To see how widespread the fractional-charge dissociation error is, we considered the first 24 open *sp*-shell atoms of the Periodic Table, i.e., H–Ca and Cu–Br (excluding the rare gas and most of the first transition-metal series elements, for reportorial and computational conveniences). For each neutral atom, we evaluated ϵ^{HO} and ϵ^{LU} within the LSDA. The results are reported in Table I, where they are presented first in order of decreasing $\epsilon^{\text{HO}}(X^0)$ and then in order of increasing $\epsilon^{\text{LU}}(Y^0)$.

From Table I, it is easy to see for which pairs *XY* Eq. (4) is satisfied and spurious fractional-charge dissociation will be found. F, at the top of the *Y* column (the lowest ϵ^{LU} at $Z=9$ in Figs. 1 and 2), can take a fraction of an electron from every other atom in the *X* column but itself. On the other hand, the atoms H, Na, Li, Be, K, Ca, Mg, and Zn, at the bottom of the *Y* column (with $\epsilon^{\text{LU}} > -0.1$ hartree in Figs. 1

TABLE I. The 24 open *sp*-shell atoms with atomic number $Z < 36$, arranged as follows. Column X : in order of decreasing highest-occupied orbital energy ϵ^{HO} (hartrees). Column Y : in order of increasing lowest-unoccupied orbital energy ϵ^{LU} (hartrees). The calculations were performed with LSDA using the 6-311 $G(d)$ basis set. The spin of the HO and LU orbitals is also shown, where α is the majority and β is the minority spin. For a spin-unpolarized atom, we show only α spin.

Z	X	$\epsilon^{\text{HO}} (X^0)$	Spin	Z	Y	$\epsilon^{\text{LU}} (Y^0)$	Spin
19	K	-0.096	α	9	F	-0.357	β
31	Ga	-0.109	α	17	Cl	-0.305	β
13	Al	-0.110	α	35	Br	-0.279	β
11	Na	-0.113	α	8	O	-0.247	β
3	Li	-0.116	α	16	S	-0.227	β
20	Ca	-0.141	α	6	C	-0.216	α
5	B	-0.147	α	34	Se	-0.214	β
32	Ge	-0.163	α	14	Si	-0.167	α
14	Si	-0.167	α	32	Ge	-0.163	α
12	Mg	-0.175	α	15	P	-0.154	β
29	Cu ^a	-0.183	α	29	Cu ^a	-0.152	β
4	Be	-0.205	α	33	As	-0.149	β
34	Se	-0.214	β	7	N	-0.148	β
6	C	-0.219	α	5	B	-0.140	α
33	As	-0.219	α	31	Ga	-0.107	α
30	Zn ^a	-0.223	α	13	Al	-0.107	α
16	S	-0.227	β	1	H	-0.093	β
15	P	-0.234	α	11	Na	-0.082	β
8	O	-0.255	β	3	Li	-0.077	β
1	H	-0.268	α	4	Be	-0.076	α
35	Br	-0.280	β	19	K	-0.075	β
7	N	-0.300	α	20	Ca	-0.053	α
17	Cl	-0.306	β	12	Mg	-0.050	α
9	F	-0.358	β	30	Zn ^a	-0.045	α

^aCalculated with the 6-311+ $G(d)$ basis set.

and 2), cannot take a fraction of an electron from any separated atom in the X column. Figures 1 and 2 show the periodic nature of ϵ^{HO} and ϵ^{LU} .

For Table I, there are $24 \times 23/2 = 276$ distinct pairs $XY = YX$ ($Y \neq X$). Simple counting shows that 174 of these (63%) display some spurious fractional-charge dissociation. The effect is pervasive and is expected to be largest for pairs near the top of the table, such as KF. The effect is, however, absent in the large class of molecules containing only C, H, N, and/or P. The table confirms Perdew's prediction¹³ of the LSDA fractional-charge dissociation in NaCl and not in LiH. It is also consistent with the fractional-charge dissociation of LiF, LiCl, NaF, and NaCl seen by Dutoi and Head-Gordon¹⁸ using a GGA, and with the calculation of significant fractional-charge dissociation in LiF, but none in H₂O by Ossowski *et al.*¹⁹ using LSDA and GGA.

Table II shows a further test of the predictive power of Table I. It shows the LSDA fractional (Mulliken) charge q on X in XY at $R=6$ Å, a finite but highly stretched bond length where the exact q is expected to be very close to zero for the systems studied in Table II. The predictions of Table I are confirmed. We will present detailed results for $XY=\text{NaCl}$ in later sections.

V. DEPENDENCE OF FRACTIONAL CHARGE AND ENERGY ON BOND LENGTH IN NaCl

When the densities of two atoms overlap, there is no unique physical definition for the net charge on one of the

atoms. We have used the Mulliken, the natural population analysis (NPA),³⁴ and the ChelpG³⁵ charges on Na and Cl in NaCl. Mulliken charges are known to be basis-set dependent and tend to give unphysical results.³⁶ We observed that for shorter interatomic distances ($R < 5$ Å) there is a considerable difference between the various charges;³⁷ the wavefunction-based Mulliken charges are always the smallest and the NPA charges are the largest. The ChelpG charges derived from the electrostatic potential are between the two. At $R = 2.5$ Å the HF Mulliken, ChelpG, and NPA charges on Na are 0.77, 0.84, and 0.96, respectively. The LSDA and PBE Na charges are equal to each other and show the same order, but the Mulliken, ChelpG, and NPA charges are somewhat smaller (0.66, 0.77, and 0.91, respectively) than the HF ones.

TABLE II. LSDA orbital energy differences (in hartrees) for the neutral atoms and LSDA fractional (Mulliken) charges q on X in XY at bond length $R=6$ Å.

XY	$\epsilon^{\text{LU}} (Y^0) - \epsilon^{\text{HO}} (X^0)$	q ($R=6$ Å)
CaO	-0.106	0.32
MgS	-0.052	0.22
AlC	-0.106	0.14
AlSi	-0.057	0.13
BeSe	-0.009	0.04
BN	-0.001	0.01
HC	+0.126	0.00

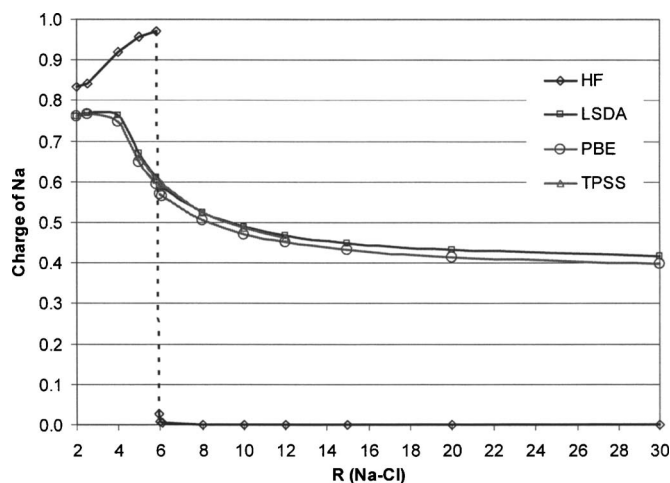


FIG. 3. Fractional (ChelpG) (Ref. 35) charge q on Na in NaCl as a function of bond length R (in Å). The Hartree-Fock “quantum phase transition” is qualitatively right, as discussed in the text.

At larger distances ($R > 6$ Å) the overlap of the orbitals becomes negligible and the Mulliken, ChelpG, and NPA charges agree closely with each other.

Figure 3 shows the fractional (ChelpG) charge q on Na in NaCl as a function of bond length R in the range from equilibrium (around 2.4 Å) to the highly stretched $R = 30$ Å. The predictions of the HF theory are compared to those of three semilocal approximations. Although HF is far from exact, it is self-interaction-free and thus probably qualitatively right.

All methods agree that q is large at equilibrium. In HF, q grows toward 1 until a critical radius $R_c^{\text{HF}} \approx 5.8$ Å is reached, then drops suddenly to near 0 for larger R . The semilocal approximations show the Mulliken q increasing, the NPA q decreasing, and the ChelpG q almost constant up to $R = 4$ Å, then gradually decreasing toward a spurious limit around 0.4 as $R \rightarrow \infty$. Semilocal approximations are “super polarizing”¹⁸ for $R > R_c$, but we observe that they are also “subpolarizing” for $R < R_c$.

The *exact* fractional charge on Na should also show a sharp transition from ~ 1 to ~ 0 around a critical R_c^{HF} (possibly with some rounding). In a molecule XY , when R is big enough to produce negligible overlap of the atomic densities and negligible contribution of $q \geq 2$, the energy difference between the cases $q = 1$ and $q = 0$ should be

$$E(X^{+1} \dots Y^{-1}) - E(X^0 \dots Y^0) = I(X^0) - A(Y^0) - \frac{e^2}{R}. \quad (7)$$

This difference vanishes at the critical radius¹¹

$$R_c = \frac{e^2}{[I(X^0) - A(Y^0)]}. \quad (8)$$

For $XY = \text{NaCl}$, using experimental $I(\text{Na}^0)$ and $A(\text{Cl}^0)$, Eq. (8) predicts $R_c(\text{NaCl}) = (27.21 \text{ eV/a.u.}) / (5.14 \text{ eV} - 3.62 \text{ eV}) = 17.9 \text{ bohrs} = 9.5$ Å. Using unrestricted Hartree-Fock/6-311G(*d*) values, $R_c^{\text{HF}}(\text{NaCl}) = 27.21 / (4.94 - 2.48) = 11 \text{ bohrs} = 5.8$ Å.

Figure 4 plots the binding energy curve $E(XY) - E(X^0) - E(Y^0)$ versus bond length R for $XY = \text{NaCl}$ in the HF,

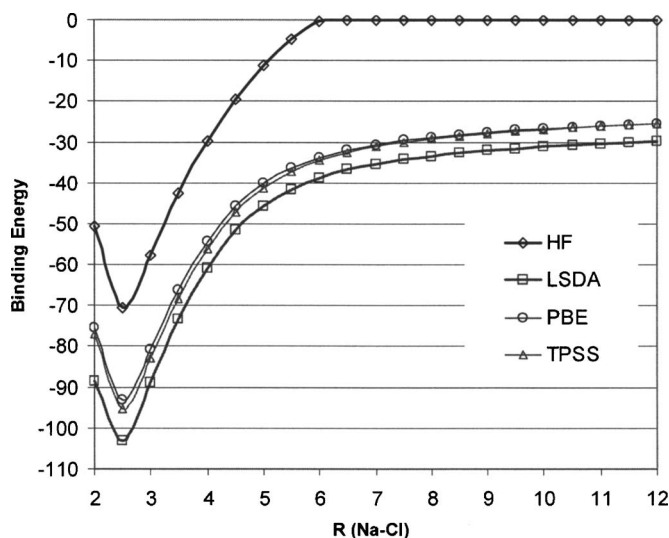


FIG. 4. Binding energy (kcal/mol) curves for NaCl $E(\text{NaCl}) - E(\text{Na}^0) - E(\text{Cl}^0)$ vs $R(\text{Na-Cl})$. The sum of the energies of the neutral atoms is taken as a reference, and the zero-point vibrational energy of NaCl is omitted. The distance R is in angstroms. The experimental atomization energy at 0 K is 97.4 ± 0.5 kcal/mol at an equilibrium $R = 2.361$ Å (Ref. 38) (1 hartree = 627.5 kcal/mol and 1 bohr = 0.5292 Å).

LSDA, PBE GGA, and TPSS meta-GGA approximations. The experimental atomization energy is 97.4 ± 0.5 kcal/mol at an equilibrium $R = 2.361$ Å.³⁸ All approximations predict a reasonable equilibrium bond length. At equilibrium, as usual LSDA overbinds (but by less than usual) and HF underbinds, with improvements from LSDA to PBE to TPSS. But only the HF curve properly tends to zero at $R \geq 9.5$ Å, where the semilocal functionals are 25–30 kcal/mol too low.

The HF behavior and the presumed exact behavior are those of a quantum phase transition in stretched NaCl. The bond length R is a parameter in the Hamiltonian; the character of the ground state can change suddenly as R passes through a critical value R_c where the energies of two different stationary states cross. At R_c , q changes rapidly (Fig. 3) and so can the derivative of the energy (Fig. 4). We note that the latter behavior is also observed for the “long-range corrected PBE functional,”³⁹ and at the predicted critical radius $R_c = 9.5$ Å.

VI. SELF-INTERACTION CORRECTION

Given any density functional approximation $E_{\text{xc}}^{\text{DFA}}[n_{\uparrow}, n_{\downarrow}]$, we can construct a self-interaction-corrected version

$$E_{\text{xc}}^{\text{DFA-SIC}} = E_{\text{xc}}^{\text{DFA}}[n_{\uparrow}, n_{\downarrow}] - \sum_i \chi_i \{E_{\text{xc}}[n_i, 0] + U[n_i]\} \quad (9)$$

that is exact for any fully spin-polarized one-electron density. Here $n(\mathbf{r}) = \sum_i n_i(\mathbf{r})$ is a sum of occupied orbital densities $[n_i(\mathbf{r}) = f_i |\psi_i(\mathbf{r})|^2]$,

$$U[n] \equiv \frac{e^2}{2} \int d\mathbf{r} \int d\mathbf{r}' \frac{n(\mathbf{r}) \cdot n(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}, \quad (10)$$

and χ_i reduces to 1 for any one-electron density. Equation (9) gives no correction to the exact $E_{\text{xc}}[n_{\uparrow}, n_{\downarrow}]$. The energy-

minimizing SIC orbitals $\psi_i(\mathbf{r})$ can be localized even when the canonical Kohn-Sham orbitals are not.

Perdew and Zunger¹⁵ made the simple choice $\chi_i=1$, as motivated by analogy with the Hartree-Fock theory, and found remarkable improvements over LSDA for atoms. In particular, they found that the energy-minimizing orbitals $\psi_i(\mathbf{r})$ (without off-diagonal Lagrange multipliers to ensure strict orthogonality) had orbital energies that satisfy Eq. (2) and nearly satisfy Eqs. (5) and (6). Historically, the PZ SIC preceded and helped motivate the exact density functional theory for fractional particle number.¹¹

Early SIC calculations for molecules included N₂ (Ref. 40) and Li₂.⁴¹ Considerable time passed before the PZ SIC could be implemented for many molecules. In molecules, the off-diagonal Lagrange multipliers are unavoidable;^{40,41} moreover the method can be quite slow in comparison with the implementation of the underlying E_{xc}^{DFA} . Eventually, it became clear^{32,33,42–45} that the PZ SIC degrades the equilibrium properties of molecules, such as bond lengths, although it improves energy barriers.^{31,43,44}

Good equilibrium properties can be restored by scaling down^{31,46} the PZ SIC. In particular, the choice

$$\chi_i = \int d\mathbf{r} |\psi_i(\mathbf{r})|^2 \cdot \left(\frac{\tau_{\sigma}^W}{\tau_{\sigma}} \right)^k \leq 1 \quad (11)$$

is exact for all one-electron densities and has certain other advantages.³¹ In Eq. (11), τ_{σ} and $\tau_{\sigma}^W = |\nabla n_{\sigma}|^2 / 8n_{\sigma} \leq \tau_{\sigma}$ are the exact and von Weizsäcker positive kinetic energy densities for electrons of the same spin σ as orbital i , and $k \geq 0$ is an exponent, typically 1 or 2. $k=0$ recovers the PZ SIC ($\chi_i=1$), which is also recovered for one-electron ground states where $\tau_{\sigma}^W/\tau_{\sigma}=1$. For densities with more than one electron of each spin, $k \rightarrow \infty$ recovers the uncorrected E_{xc}^{DFA} ($\chi_i=0$).

If the PZ SIC (choice $k=0$) approximately satisfies Eqs. (5) and (6) for all atoms, then it is clear from Eq. (11) that the scaled-down SIC will not, when k is sufficiently greater than 0, by the argument at the end of the preceding paragraph. The same can be said for the fractional-electron-number version of the exchange-correlation hole sum rule.^{13,15}

VII. RESILIENCE OF SPURIOUS FRACTIONAL-CHARGE DISSOCIATION: RESULTS FOR NaCl

Table III shows the fractional (Mulliken) charge q on Na in NaCl for the stretched bond lengths $R=6$ Å and $R=12$ Å, which bracket the critical value $R_c=9.5$ Å discussed in Sec. V. Results are presented for LSDA and PBE GGA [or, equivalently for this many-electron system, Eq. (11) with $k \rightarrow \infty$] and for various scaled-down ($k=2$ and 1) or full ($k=0$) self-interaction corrections. As k decreases toward 0, the results improve, but only the full PZ SIC ($k=0$) is qualitatively correct (as expected from the argument at the end of the preceding section). Figure 5 shows that for the highly stretched bond lengths, the binding energy curve shows a similar improvement as $k \rightarrow 0$.

The problem is that a good description of the equilib-

TABLE III. Fractional (Mulliken) charge q on Na in NaCl at bond lengths $R=6$ and 12 Å, evaluated in the HF, LSDA ($k=\infty$), and PBE GGA ($k=\infty$), with scaled-down self-interaction corrections ($k=2$ and 1), and with full Perdew-Zunger (PZ) self-interaction correction ($k=0$). The PZ results are qualitatively correct.

Method	k	q ($R=6$ Å)	q ($R=12$ Å)
HF	...	0.01	0.00
LSDA	∞	0.59	0.47
LSD-SIC	2	0.73	0.44
LSD-SIC	1	0.83	0.33
LSD-SIC-PZ	0	0.97	0.10
PBE	∞	0.57	0.45
PBE-SIC	2	0.71	0.39
PBE-SIC	1	0.81	0.22
PBE-SIC-PZ	0	0.92	0.05

rium properties of molecules requires $k=2$ or 1, not 0. When we improve the equilibrium properties in SIC, the resilient fractional-charge error springs back, in part.

VIII. CONCLUSIONS

In the density functional theory, a molecule XY ($Y \neq X$) will dissociate to spurious fractional-charge fragments $X^{+q} \cdots Y^{-q}$, with an energy below $X^0 \cdots Y^0$, if the lowest-unoccupied orbital energy of Y lies below the highest-occupied orbital energy of X [Eq. (4) of Sec. II, as $q' \rightarrow 0+$]. Then there will be an error in the large-separation region of the binding energy curve or energy surface, although seeing these errors with standard molecular codes may require special care (Sec. III). This error is not uncommon for the approximate semilocal functionals LSDA, GGA, and meta-

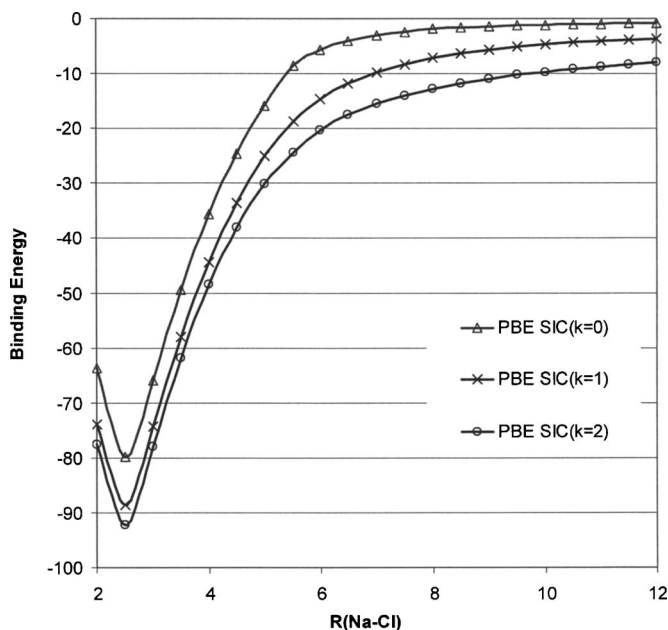


FIG. 5. Binding energy (kcal/mol) curves for NaCl, $E(\text{NaCl}) - E(\text{Na}^0) - E(\text{Cl}^0)$ vs $R(\text{Na-Cl})$, calculated by self-interaction-corrected PBE GGA. $k=0$ is the original SIC introduced by Perdew and Zunger, and k is the scaling exponent of Eq. (11). The sum of the energies of the neutral atoms is taken as a reference.

GGA (Sec. IV) and is typically similar at all three levels (even though the TPSS meta-GGA correlation energy is exact for any one-electron density).

In the exact density functional theory, Eqs. (5) and (6) are satisfied. For NaCl and similar molecules, there is then a critical bond length R_c , Eq. (8), beyond which the exact q drops to zero. The semilocal approximations make q too large for $R > R_c$ and too small for $R < R_c$ (Sec. V).

The self-interaction correction of Perdew and Zunger,¹⁵ which approximately satisfies Eqs. (5) and (6) for atoms (when the orbital energies are calculated as in Ref. 15), eliminates the spurious fractional-charge dissociation of NaCl and tends to improve energy barriers, but it worsens the equilibrium properties of molecules. When we scale down the self-interaction correction via Eq. (11),³¹ we improve the equilibrium properties and keep the functional exact for any one-electron density, but the resilient fractional-charge error returns in part because Eqs. (5) and (6) are now satisfied less well.

We can say that a functional is “one-electron self-interaction-free” if it is exact for any one-electron density, and that it is nearly “many-electron self-interaction-free” if its total energy has a realistic linear variation with electron number N between all pairs of integers $M-1$ and M . Both the original Perdew-Zunger self-interaction correction and our scaled-down version of it are one-electron self-interaction-free, but only the former is nearly many-electron self-interaction-free for atoms (even more so than is the Hartree-Fock theory) (see Fig. 6 of Ref. 13).

More precisely, we can say that a functional is nearly “ M -electron self-interaction-free” (M =positive integer) if its total energy for a system of fixed external potential with electron number N in the range $M-1 < N \leq M$ has a realistic linear variation with N , or equivalently if its highest-occupied orbital energy for $M-1 < N \leq M$ is nearly equal to minus the electron-removal energy from the M -electron system. (This removal energy is zero if the system cannot bind M electrons, and in that case the unbound electrons go to the bottom of the free-electron continuum.) Our scaled-down SIC reduces to PZ SIC for any subsystem ground state with two or fewer electrons, so it is exactly one- and nearly two-electron self-interaction-free, which improves its description of dissociation not only for H_2^+ but also for He_2^+ (in which each dissociation fragment $\text{He}^{+0.5}$ has $1.5 < 2$ electrons).⁴⁷

The development of accurate approximations that are nearly M -electron self-interaction-free for all M is one of the most pressing problems in the density functional theory. The solution of this problem would be relevant not only to the calculation of energy surfaces but also to the problems of charge transfer,⁴⁸ charge transport,⁴⁹ polarizability,⁵⁰ and time-dependent density functional theory.⁵¹ It remains to be seen to what extent hyper-GGA's (Refs. 52–55) that use full exact exchange and nonlocal correlation can be many-electron self-interaction-free.

ACKNOWLEDGMENTS

Two of the authors (A.R. and G.I.C.) dedicate this article to the memory of their teacher and friend József Réffy. One of the authors (A.R.) acknowledges the support of an OTKA

postdoctoral fellowship PD-050014. NSF support is acknowledged by one of the authors (J.P.P.) for DMR-0501588 and by two of the authors (G.E.S. and O.A.V.) for CHE-0457030. OTKA-NSF support is acknowledged by one of the authors (G.I.C.).

- ¹W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
- ²*A Primer in Density Functional Theory*, edited by C. Fiolhais, F. Nogueira, and M. Marques (Springer, Berlin, 2003).
- ³V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, J. Chem. Phys. **119**, 12129 (2003); **121**, 11507(E) (2004).
- ⁴J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
- ⁵S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
- ⁶J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); **78**, 1396(E) (1997).
- ⁷J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).
- ⁸F. Furche and J. P. Perdew, J. Chem. Phys. **124**, 044103 (2006).
- ⁹A. Ruzsinszky, J. P. Perdew, and G. I. Csonka, J. Phys. Chem. A **109**, 11006 (2005); **109**, 11015 (2005).
- ¹⁰J. C. Slater, *The Self-Consistent Field for Molecules and Solids* (McGraw-Hill, New York, 1974).
- ¹¹J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).
- ¹²J. P. Perdew and J. R. Smith, Surf. Sci. **141**, L295 (1984).
- ¹³J. P. Perdew, in *Density Functional Method in Physics*, edited by R. M. Dreizler and J. da Providencia (Plenum, New York, 1985).
- ¹⁴J. P. Perdew, Adv. Quantum Chem. **21**, 113 (1990).
- ¹⁵J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
- ¹⁶M. Ernzerhof and J. P. Perdew (to be published); A. Ruzsinszky, J. P. Perdew, and G. I. Csonka (to be published).
- ¹⁷M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN Development Version, Revision D.02, Gaussian Inc., Wallingford, CT, 2004.
- ¹⁸A. D. Dutoi and M. Head-Gordon, Chem. Phys. Lett. **422**, 230 (2006).
- ¹⁹M. M. Ossowski, L. L. Boyer, M. J. Mehl, and M. R. Pederson, Phys. Rev. B **68**, 245107 (2003).
- ²⁰J. P. Perdew and M. Ernzerhof, in *Electron Density Functional Theory: Recent Progress and New Directions*, edited by J. F. Dobson, G. Vignale, and M. P. Das (Plenum, New York, 1998).
- ²¹T. Bally and G. N. Sastry, J. Phys. Chem. A **101**, 7923 (1997).
- ²²B. Braïda, P. C. Hiberty, and A. Savin, J. Phys. Chem. A **102**, 7872 (1998).
- ²³M. Grüning, O. V. Gritsenko, S. J. A. van Gisbergen, and E. J. Baerends, J. Phys. Chem. A **105**, 9211 (2001).
- ²⁴D. J. Tozer, N. C. Handy, and A. J. Cohen, Chem. Phys. Lett. **382**, 203 (2003).
- ²⁵M. Lundberg and P. E. M. Siegbahn, J. Chem. Phys. **122**, 224103 (2005).
- ²⁶Y. Zhang and W. Yang, J. Chem. Phys. **109**, 2604 (1998). Note that these authors found spurious fractional-charge dissociation of CO^- , but were apparently unaware of previous work on fractional-charge dissociation of neutral molecules.
- ²⁷J. F. Janak, Phys. Rev. B **18**, 7165 (1978).
- ²⁸A. Savin, F. Colonna, and R. Pollet, Int. J. Quantum Chem. **93**, 166 (2003); see Fig. 12.
- ²⁹C. M. Horowitz, C. R. Proetto, and S. Rigamonti, Phys. Rev. Lett. **97**, 026802 (2006).
- ³⁰M. Grüning, A. Marini, and A. Rubio, J. Chem. Phys. **124**, 154108 (2006).
- ³¹O. A. Vydrov, G. E. Scuseria, J. P. Perdew, A. Ruzsinszky, and G. I. Csonka, J. Chem. Phys. **124**, 094188 (2006).
- ³²O. A. Vydrov and G. E. Scuseria, J. Chem. Phys. **121**, 8187 (2004).
- ³³O. A. Vydrov and G. E. Scuseria, J. Chem. Phys. **122**, 184107 (2005).
- ³⁴A. E. Reed, R. B. Weinstock, and F. Weinhold, J. Chem. Phys. **83**, 735 (1985), and references cited therein.
- ³⁵C. M. Breneman and K. B. Wiberg, J. Comput. Chem. **11**, 361 (1990).
- ³⁶S. Kristyán and G. I. Csonka, Chem. Phys. Lett. **307**, 469 (1999).
- ³⁷A. Ruzsinszky, C. Van Alsenoy, and G. I. Csonka, J. Phys. Chem. A **106**, 12139 (2002).
- ³⁸NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database No. 101, Release 12, edited by R. D. Johnson III, 2005, <http://srdata.nist.gov/cccbdb>
- ³⁹O. A. Vydrov and G. E. Scuseria (submitted).
- ⁴⁰M. R. Pederson, R. A. Heaton, and C. C. Lin, J. Chem. Phys. **80**, 1972

- (1984).
- ⁴¹M. R. Pederson, R. A. Heaton, and C. C. Lin, J. Chem. Phys. **82**, 2688 (1985).
- ⁴²S. Goedecker and C. J. Umrigar, Phys. Rev. A **55**, 1765 (1997).
- ⁴³G. I. Csonka and B. G. Johnson, Theor. Chem. Acc. **99**, 158 (1998).
- ⁴⁴S. Patchkovskii and T. Ziegler, J. Chem. Phys. **116**, 7806 (2002).
- ⁴⁵J. Gräfenstein, E. Kraka, and D. Cremer, J. Chem. Phys. **120**, 524 (2004).
- ⁴⁶K. Tsemekhman, E. J. Bylaska, E. C. Brown, and H. Jonsson (to be published); E. J. Bylaska, K. Tsemekhman, and F. Gao, Physica Scripta T **124**, 86 (2006).
- ⁴⁷A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, J. Chem. Phys. (submitted).
- ⁴⁸D. J. Tozer, J. Chem. Phys. **119**, 12697 (2003).
- ⁴⁹C. Toher, A. Filipetti, S. Sanvito, and K. Burke, Phys. Rev. Lett. **95**, 146402 (2005).
- ⁵⁰S. Kümmel, L. Kronik, and J. P. Perdew, Phys. Rev. Lett. **93**, 213002 (2004).
- ⁵¹M. Mundt and S. Kümmel, Phys. Rev. Lett. **95**, 203004 (2005).
- ⁵²J. P. Perdew and K. Schmidt, in *Density Functional Theory and Its Applications to Materials*, edited by V. E. Van Doren, C. Van Alsenov, and P. Geerlings (American Institute of Physics, Melville, NY, 2001).
- ⁵³A. D. Becke, J. Chem. Phys. **119**, 2972 (2003).
- ⁵⁴P. Mori-Sanchez, A. Cohen, and W. Yang, J. Chem. Phys. **124**, 091102 (2006).
- ⁵⁵J. P. Perdew, V. N. Staroverov, G. E. Scuseria, and J. Tao (to be published).