

## GW Approximation of the Many-Body Problem and Changes in the Particle Number

Fabien Bruneval

CEA, DEN, Service de Recherches de Métallurgie Physique, F-91191 Gif-sur-Yvette, France

(Received 25 August 2009; published 23 October 2009)

A stringent test for an exchange-correlation approximation in electronic structure calculations is the equality between the ionization energy of the neutral system and the affinity of the singly positively charged system. All of the commonly used approximations (local, semilocal, hybrid) for the exchange correlation within density functional theory fail badly with this test. They consequently present a localization or delocalization error, resulting in a highest occupied molecular orbital or lowest unoccupied molecular orbital gap over- or underestimation. The *GW* approximation appears as the best available framework to describe particle number changes. The small remaining error can be further reduced by devising a  $\Delta$ SCF-like method within the *GW* approximation. The proposed approach is necessary as soon as localized states are involved, e.g., in finite systems or defect states in crystals.

DOI: 10.1103/PhysRevLett.103.176403

PACS numbers: 71.15.-m, 32.10.Hq, 71.10.-w, 71.55.-i

For many years, density functional theory (DFT) has been seeking for the most correct approximation for the exchange-correlation functional [1]. The exact exchange-correlation term is of course unknown, since it has to account for all the quantum effects contained in the Schrödinger equation for interacting electrons. Several shortcomings of its existing approximations have been identified. A crucial issue for the available exchange-correlation functionals is the behavior of the energy along with the change of the number of electrons. Practical difficulties are tightly related to that problem: for instance, the infamous band gap problem arises from the poor description of electron addition and removal for most of the existing functionals.

Generalizing the idea of particle number changes to fractional number of electrons, it has been shown [2,3] from ensemble arguments that the total energy should be linear in between integral numbers of electrons. Unfortunately, all the usual approximations fail with this crucial property: local density approximation (LDA) and generalized gradient approximation (GGA) are convex, whereas Hartree-Fock (HF) is concave [4]. The straight line behavior arising from the exact exchange correlation is not just a playground for theoreticians. Indeed, the first ionization energy  $I$  can be obtained as the derivative on the left-hand side of the total energy with respect to the number of particles and the electron affinity  $A$  can be reached by the derivative on the right-hand side. If the slope of the total energy between  $N - 1$  and  $N$  electrons is not a constant, the electron affinity of the  $N - 1$  electron system is not to be equal to the ionization energy of the  $N$  electron system, even though they should represent the same total energy difference:

$$A(N - 1) = - \left. \frac{\partial E_0}{\partial n} \right|_{n=(N-1)^+} = E_0(N - 1) - E_0(N) \quad (1)$$

$$I(N) = - \left. \frac{\partial E_0}{\partial n} \right|_{n=N^-} = E_0(N - 1) - E_0(N), \quad (2)$$

where  $E_0(N)$  is the ground-state energy of the system with  $N$  electrons. The  $+$  and  $-$  signs indicate the side of the derivative. In most cases, the derivative reduces to the eigenvalue thanks to the Janak's theorem [5].

In a recent work, Cohen, Mori-Sánchez, and Yang [4,6] have clarified the relation between convexity or concavity, localization error and band gap error. The convex approximations, like LDA or GGA, lower the energy in spreading electrons as much as possible so that a fractional number of electrons is preferred. The concave approximations, like HF, instead find energetically favorable to localize electrons so that they integrate to an integer. HF is indeed a quasiparticle theory, which relies on integral number of particles. The true exchange correlation should be insensitive to these two situations. Only in this case, the total energy difference, called  $\Delta$ SCF method, is to match the eigenvalue estimate of the ionization or affinity.

Originating from an other framework, many-body perturbation theory, the *GW* approximation to the exchange correlation [7] has been extremely powerful in describing the band structure of solids [8,9]. The *GW* approximation is an improvement over the HF approximation. It is based on the concept of screened Coulomb interaction. In practice, the *GW* approximation is usually evaluated as a first-order perturbation with respect to LDA, in the so-called  $G_0W_0$  approach. This assumes, in particular, that the LDA and the  $G_0W_0$  wave functions are identical. The currently best implementation of the *GW* approximation, the quasi particle self-consistent *GW* (sc*GW*) [10–12], proposes a static approximation to the complete *GW* case, which allows one to recalculate self-consistently the *GW* wave functions and eigenvalues. This approach is compulsory in the case of atoms and molecules, as shown in the following.

In the present Letter, we evaluate the quality of local, hybrid, and  $GW$  approximations to the exchange correlation in terms of localization error, band gap error, and ionization or affinity consistency. By testing them on small sodium clusters, we show that the  $GW$  approximation prevails over all the usual exchange-correlation approximations. The small remaining error in the ionization or affinity determination can be integrated in the framework in devising an extension of the  $\Delta$ SCF procedure to the  $GW$  approach. We finally demonstrate the effectiveness of the procedure for localized defect states in a crystal.

It is unfortunately difficult in practice to obtain the total energy within the  $GW$  framework. Furthermore, the generalization of the  $GW$  equations to fractional numbers of electrons would require some care. As a consequence, we propose two alternative routes to evaluate the behavior of the  $GW$  approximation as a function of a fractional number of electrons. (i) Analyze the  $GW$  wave functions of an extra electron (or hole) in a doubled system: consider the system of two distant molecules with one additional electron. Will the extra electron be spread onto the two molecules or will it localize on one of them? (ii) Compare the ionization energy of a neutral system with the affinity of the positively charged one. If  $-A(+)<-I(0)$  [ $-A(+)>-I(0)$ ], the approximation is to be convex (concave). Furthermore, the difference between  $I(0)$  and  $A(+)$  evaluates the magnitude of the localization error.

Following our route (i), we considered a system consisting of two distant molecules of  $\text{Na}_2$  in a supercell with one extra electron. The technical details are provided below. If the chosen scheme were exact, the extra electron spread onto the two molecules or the electron localized on one single molecule should be two stationary points of the functional giving the same total energy. LDA is subjected to delocalization error so that the extra electron has an equal weight on the two molecules. The panel (a) of Fig. 1 represents the LDA highest occupied molecular orbital (HOMO) that places half an electron on the two molecules. Note that the lowest unoccupied molecular orbital (LUMO) is degenerate with the HOMO within LDA. The HF framework localizes easily the extra electron on one of the two molecules and breaks the HOMO/LUMO degeneracy. We performed then scGW calculations starting either from HF and from LDA. Initiating the scGW evaluation from the HF wave functions leads to a rapidly converging result, which maintains the extra electron on one  $\text{Na}_2$  molecule [panel (b) of Fig. 1]. Starting the scGW calculation from the LDA wave functions leads to a slowly converging result: after a dozen iterations with the extra electron spread over the two molecules, the HOMO wave function finally turns into the HF one. Whatever the starting point, the  $GW$  calculation ends up in the same stationary point, localizing the electron on one single  $\text{Na}_2$  molecule. As a consequence, the  $GW$  approach yields a concave total energy and suffers from a localization error.

In the following, we quantify the nonlinearity of the scGW approximation with calculations for small sodium

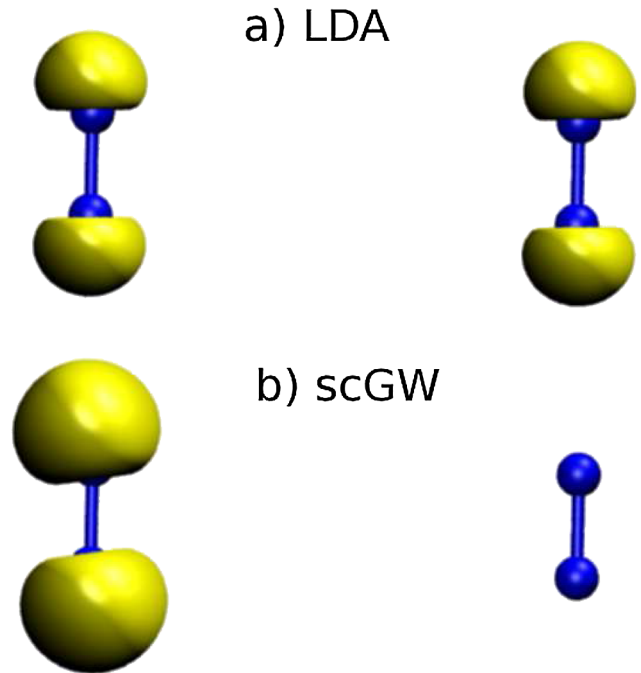


FIG. 1 (color online). Isosurface of an extra electron added in the system of two distant  $\text{Na}_2$  molecules, or in other words, isosurface of the highest occupied molecular orbital of the  $(2\text{Na}_2)^-$  system. Panel (a) represents the LDA results, whereas panel (b) provides the scGW result.

clusters according to our route (ii): we compare the consistency of ionizations and affinities. The sodium clusters are a system of choice for such a study, since accurate experiments [13] and configuration-interaction calculations [14] are available. Furthermore, these systems are practical enough so that we still can use a plane-wave code, which is customary in the  $GW$  framework. We performed  $\Gamma$ -point calculations in large face-centered cubic supercells with a 60 Bohr cubic edge length. We use a plane-wave cutoff of 14 Ha for wave functions and of 2 Ha for dielectric matrices. We use a norm-conserving pseudo-potential, where the semicore states ( $2s^22p^6$ ) of sodium are treated as valence. These states are indeed very important for the exchange operator and are noticeably polarizable. We employ a plasmon-pole approximation, and we have verified that this is not an issue. The number of states is 512, which is reasonably low thanks to the acceleration scheme of Ref. [15]. The Coulomb interaction has been cut off in order to remove spurious image interactions [16,17]. As the geometries are not the issue here, we performed all the calculations in the neutral configurations obtained from a reference quantum chemistry work [14].

Some care is required due to the periodic approach. Periodic charged systems in presence a neutralizing background slowly converge [18]. However, the eigenvalues of neutral molecule experience also slow convergence in the supercell approach. The eigenvalues are shifted with respect to an isolated system calculation since the potential does not vanish at infinity in the periodic calculation. The

difference between a finite system calculation within Gaussian formalism using a 6-311++G\*\* basis and our periodic supercell approach appears to be a mere shift of the eigenvalues, whatever the charge, we simulate. Both neutral and charged systems can be corrected by shifting the supercell eigenvalues onto the isolated ones within LDA, for instance. Using the same shift, we were able to superimpose the HF eigenvalues from the periodic calculations onto the isolated results with a 0.1 eV accuracy. In the following, this shifting procedure is systematically applied. Note that we do not present results for the affinity of the clusters, as they show strong dependence with respect to the supercell size.

Table I compares the ionization energy of small sodium clusters, as obtained from removing an electron from  $\text{Na}_n$  or from adding an electron to  $\text{Na}_n^+$ , within LDA, HF, B3LYP,  $G_0W_0$ , and scGW. For all the approximations considered here, there is no discontinuity in the exchange-correlation functional so that the ionizations and affinities reduce to the HOMO and LUMO eigenvalues. Within LDA, the LUMO energy for the positively charged clusters is much lower than the HOMO energy of the neutral ones. LDA (and GGA, not presented here) is a convex approximation, which is consistent with the band gap underestimation problem [4]. HF gives the exact answer for a one electron system, since it is devoid of any self-interaction. The sodium atom, which has a single 3s valence electron, is well described within HF. The agreement between ionization and affinities then deteriorates up to  $\sim 0.8$  eV for the largest clusters. The HF approximation is clearly concave, which is consistent with the observed band gap overestimation. The hybrid functional family that mixes LDA, GGA, and exact-exchange could be a potential answer. The B3LYP functional [19] that includes 20% of exact-exchange is still convex: B3LYP predicts systematically the LUMO energies of the  $\text{Na}_n^+$  clusters much lower than the HOMO of the  $\text{Na}_n$ .

Turning to *GW* calculations, we first provide for completeness the standard  $G_0W_0$  results. Our results agree well with the published data for neutral species from Ref. [20]. Though reasonable compared to experiment, the  $G_0W_0$  data are difficult to interpret and do not show clear trends. This is mainly due to the inadequacy of the perturbative approach in the case of the unoccupied states in a finite

system. The scGW approach, which recalculates the wave functions, provides the most sensible results. The LUMO of  $\text{Na}_n^+$  is systematically slightly higher than the HOMO of  $\text{Na}_n$ , but the difference is always lower than 0.45 eV. This shows a small, but noticeable, localization error in agreement with the result from route (i).

Because of the inconsistency between the eigenvalues, the ionizations and affinities are generally obtained from either total energy differences, the  $\Delta$ SCF method, or from Slater's transition state theory [21]. Both approaches generally agree very well. The  $\Delta$ SCF results provided in Table I for LDA, HF, and B3LYP supersede the eigenvalue estimate within the corresponding approximations. Following the argument of Slater, if the total energy within our approximation is not linear for fractional number of electrons as it should be, we may expand it as a second order polynomial. Under this assumption, the ionization energy can be approximated by the eigenvalue at the half charge  $N - 1/2$ . The Slater's transition state approach gives a very good estimate for the total energy difference. Following the same arguments, we observe that the total energy difference can be also evaluated as the mean value,

$$I(N) \approx -\frac{1}{2} \left[ \left. \frac{\partial E_0}{\partial n} \right|_{n=(N-1)^+} + \left. \frac{\partial E_0}{\partial n} \right|_{n=N^-} \right]. \quad (3)$$

The evaluation of Eq. (3) does not require total energy nor fractional charge calculations, but only the derivatives with respect to the particle number, which reduce in most cases to the eigenvalues of the neutral and the charged system. It can be readily evaluated from the data provided in Table I. The final result, labeled  $\Delta$ SCF within scGW approximation, gives the best estimate of all approximations for the ionization energy of the sodium clusters. Furthermore, this  $\Delta$ SCF procedure allows for a reconciliation between total energy and eigenvalue approaches. A direct evaluation (beyond reach by now) of the scGW total energy differences would be consistent with the proposed procedure.

As a final illustration of the inconsistency between eigenvalues between charged systems, we consider the localized state in the band gap of crystal created by a point defect. We exemplify with the carbon split  $\langle 100 \rangle$  interstitial in cubic silicon carbide [22]. The calculations have been performed in a 65 atom cubic supercell with a  $2 \times 2 \times 2$   $k$  point sampling. The structure of the defect (inset of Fig. 2)

TABLE I. Ionization energy in eV of the small sodium clusters evaluated from the HOMO of the neutral species  $-I(0)$ , from the LUMO of the cationic species  $-A(+)$ , or from the difference in total energies ( $\Delta$ SCF), within different approximations to the exchange correlation: LDA, HF, B3LYP,  $G_0W_0$ , and scGW.

	LDA			HF			B3LYP			$G_0W_0$		scGW			Expt. [13]
	$-A(+)$	$-I(0)$	$\Delta$ SCF	$-A(+)$	$-I(0)$	$\Delta$ SCF	$-A(+)$	$-I(0)$	$\Delta$ SCF	$-A(+)$	$-I(0)$	$-A(+)$	$-I(0)$	$\Delta$ SCF	
$\text{Na}_1$	-6.96	-3.08	-5.36	-4.94	-4.95	-4.94	-7.10	-3.48	-5.42	-4.88	-5.40	-5.05	-5.49	-5.27	-5.139
$\text{Na}_2$	-7.12	-3.20	-5.25	-3.90	-4.48	-4.08	-6.78	-3.52	-5.19	-5.10	-5.05	-4.66	-5.10	-4.88	-4.934
$\text{Na}_3$	-5.99	-2.62	-4.30	-3.47	-4.17	-3.80	-5.59	-2.91	-4.24	-4.42	-4.24	-4.15	-4.42	-4.29	-3.97
$\text{Na}_4$	-6.01	-2.77	-4.38	-3.00	-3.78	-3.34	-5.55	-2.98	-4.25	-4.71	-4.29	-4.18	-4.38	-4.28	-4.27
$\text{Na}_5$	-5.77	-2.78	-4.27	-3.34	-4.43	-3.94	-5.32	-3.03	-4.17	-4.54	-4.17	-4.18	-4.39	-4.28	-4.05

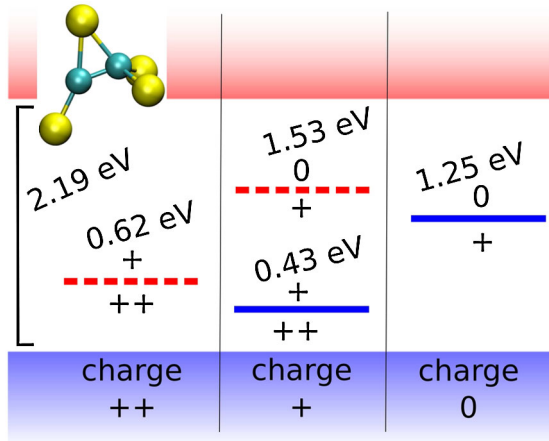


FIG. 2 (color online). Vertical charge transition levels of the carbon split  $\langle 100 \rangle$  interstitial of 3C-SiC, evaluated within  $G_0W_0$  using different charges for the 65 atom supercell. The occupied levels are plain lines, whereas the empty ones are dashed. The structure of the defect has been kept frozen in the neutral geometry, as shown in the inset in the upper left corner. The energy of the defect levels is referred to the top valence band.

has been frozen in the neutral optimized configuration in order to isolate the purely electronic behavior we want to address. We performed LDA and  $G_0W_0$  evaluation of the position of the defect level inside the band gap for different charged supercells (charges 0, + and ++). In this case, scGW is not required since we have verified that the LDA wave functions constitute a good approximation for the scGW wave functions. Comparing the density of states, we carefully checked that the very tiny band shifts were not an issue. The trends are consistent with the sodium clusters. The eigenvalues of the levels within LDA deviate strongly when changing their occupation, showing a strong convex behavior (the HOMO of the neutral system is higher than the LUMO of the positive system). As seen from Fig. 2, the discrepancy between ionizations and affinities is small within  $G_0W_0$  ( $\sim 0.2$ – $0.3$  eV) and confirms the slight localization error. For the defect calculations, we again propose the  $\Delta$ SCF procedure within  $GW$ , which simulates total energy differences without the need to perform such calculations. The final  $\Delta$ SCF value for the charge transitions are  $\epsilon_{GW}(+/0) = 1.39$  eV and  $\epsilon_{GW}(++/+) = 0.53$  eV.

In conclusion, we proposed to judge the quality of the exchange-correlation approximations on the discrepancy between ionization of the neutral system and affinity of the positively charged one. The exact exchange-correlation functional should not have any. From all the approximations tested here (LDA, HF, B3LYP,  $GW$ ), the  $GW$  approach offers the lowest discrepancy. The small remaining error within  $GW$  is consistent with a systematic localization error and the slight band gap overestimation observed in practice [11,23]. In order to provide the most meaningful

results, we support the use of a  $\Delta$ SCF-like procedure to conciliate total energy and quasiparticle energy evaluations of the ionization and affinity energies. Finally, the ionization or affinity consistency can give insights concerning the properties of the vertex function that should fix the  $GW$  errors.

The present calculations were performed using the ABINIT code [24] and GAUSSIAN03 [25]. We are grateful to Silvana Botti for her helpful comments. This work was performed using HPC resources from GENCI-CINES (Grant No. 2009-GEN6018).

- [1] W. Kohn, Rev. Mod. Phys. **71**, 1253 (1999).
- [2] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Jr., Phys. Rev. Lett. **49**, 1691 (1982).
- [3] W. T. Yang, Y. Zhang, and P. W. Ayers, Phys. Rev. Lett. **84**, 5172 (2000).
- [4] P. Mori-Sánchez, A. J. Cohen, and W. T. Yang, Phys. Rev. Lett. **100**, 146401 (2008).
- [5] J. F. Janak, Phys. Rev. B **18**, 7165 (1978).
- [6] A. J. Cohen, P. Mori-Sánchez, and W. T. Yang, Science **321**, 792 (2008).
- [7] L. Hedin, Phys. Rev. **139**, A796 (1965).
- [8] F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998).
- [9] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985).
- [10] S. V. Faleev, M. van Schilfhaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004).
- [11] M. van Schilfhaarde, T. Kotani, and S. Faleev, Phys. Rev. Lett. **96**, 226402 (2006).
- [12] F. Bruneval, N. Vast, and L. Reining, Phys. Rev. B **74**, 045102 (2006).
- [13] A. Herrmann, E. Schumacher, and L. Wöste, J. Chem. Phys. **68**, 2327 (1978).
- [14] V. Bonačić-Koutecký, P. Fantucci, and J. Koutecký, J. Chem. Phys. **93**, 3802 (1990).
- [15] F. Bruneval and X. Gonze, Phys. Rev. B **78**, 085125 (2008).
- [16] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni, Phys. Rev. Lett. **75**, 818 (1995).
- [17] J. Spencer and A. Alavi, Phys. Rev. B **77**, 193110 (2008).
- [18] G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).
- [19] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
- [20] Y. Noguchi, S. Ishii, K. Ohno, and T. Sasaki, J. Chem. Phys. **129**, 104104 (2008).
- [21] J. C. Slater, *The Self-Consistent Field for Molecules and Solids* (McGraw-Hill, New York, 1974), Vol. 4.
- [22] M. Bockstedte, A. Mattausch, and O. Pankratov, Phys. Rev. B **68**, 205201 (2003).
- [23] M. Shishkin, M. Marsman, and G. Kresse, Phys. Rev. Lett. **99**, 246403 (2007).
- [24] X. Gonze *et al.*, Z. Kristallogr. **220**, 558 (2005).
- [25] M. J. Frisch *et al.*, *Gaussian 03, Revision C.02* (Gaussian, Inc., Wallingford, CT, 2004).