Many-Body Perturbation Theory Using the Density-Functional Concept: Beyond the *GW* Approximation

Fabien Bruneval,[1] Francesco Sottile,[1] Valerio Olevano,[1] Rodolfo Del Sole,[2] and Lucia Reining[1]

[1]*Laboratoire des Solides Irradiés, UMR 7642 CNRS/CEA, École Polytechnique, 91128 Palaiseau, France*

[2]*Istituto Nazionale per la Fisica della Materia e Dipartimento di Fisica dell'Università di Roma "Tor Vergata," Via della Ricerca Scientifica, 00133 Roma, Italy*

(Received 4 October 2004; published 9 May 2005)

We propose an alternative formulation of many-body perturbation theory that uses the density-functional concept. Instead of the usual four-point integral equation for the polarizability, we obtain a two-point one, which leads to excellent optical absorption and energy-loss spectra. The corresponding three-point vertex function and self-energy are then simply calculated via an integration, for any level of approximation. Moreover, we show the direct impact of this formulation on the time-dependent density-functional theory. Numerical results for the band gap of bulk silicon and solid argon illustrate corrections beyond the *GW* approximation for the self-energy.

DOI: 10.1103/PhysRevLett.94.186402 PACS numbers: 71.10.−w, 71.15.−m, 71.35.−y, 78.20.Bh

The electronic structure of materials and its response to an external perturbation are key quantities for the interpretation of many experimental results or for the design of technological devices. In this context, *ab initio* electronic structure calculations have become a tool of choice. One can already obtain useful information about the band structure because of the widely used Kohn-Sham (KS) framework of the density-functional theory (DFT) [1]; for the response function, one can use an independent particle Fermi's golden rule [which is equivalent to the random phase approximation (RPA)]. Beyond this, state-of-the-art calculations for solids are based on the many-body perturbation theory. In that case, quasiparticle (QP) band structure energies are obtained from the solution of an equation similar to the KS one, but with the KS exchange-correlation (xc) potential $v_{\rm xc}$ replaced by the electron self-energy Σ, most often using Hedin's *GW* approximation [2]. In this approximation, Σ is the product of the one-particle Green's function G and the screened Coulomb interaction W calculated in the RPA. The resulting band structures, and, in particular, the band gap, are generally much closer to the measured ones than the KS results [3]. In order to get improved response functions, the electron-hole interaction can then be included by adding the so-called "vertex corrections" beyond the RPA, which is done in practice by solving the four-point Bethe-Salpeter equation (BSE) for the polarizability P; this leads in general to excellent absorption and electron energy-loss spectra [4]. In particular, one correctly describes the important excitonic effects. When needed [5], one could then use the vertex and the improved response function to construct a new self-energy beyond the RPA *GW* approximation.

Unfortunately, calculations of vertex corrections are cumbersome essentially because of the four-point (electron-hole scattering) nature of the BSE (see, e.g., [4] for the BSE, or [6,7] for the self-energy beyond *GW*). Calculations of the response function have been limited to relatively simple systems. Cancellation effects on quasiparticle energies between the vertices in P and Σ have been discussed only for selected cases, especially for the homogeneous electron gas [8], or for bulk silicon but using a vertex derived from a *local* approximation to the self-energy, namely, the KS potential in the local density approximation (LDA) [9]. It is therefore an important challenge to find an efficient way to evaluate the effects of vertices arising from more realistic KS potentials or, better, from nonlocal self-energies.

Alternatively, it is known that *in principle* one could obtain the polarizability directly from a two-point equation: this is the case when one works in the framework of time-dependent DFT (TDDFT) [10], since one propagates the density instead of the Green's function. TDDFT could therefore clearly be a prominent alternative to the BSE for the calculation of P. Recently, a reliable (long-range) approximation for the two-point xc kernel $f_{\rm xc}(\mathbf{r}, \mathbf{r}', t, t')$ of the TDDFT linear response equation for P has been derived from the BSE; this combines the precision of the latter with the computational advantages of TDDFT [11]. TDDFT is, of course, not designed to access one-QP properties, like the band structure; yet, one may try to use the progress made concerning P in order to find approximations for Σ beyond the *GW* approximation, and a first attempt in this sense has already been made concerning QP lifetimes [6].

The ultimate goal would be, of course, to combine the density functional and the QP concepts in such a way that systematic and efficient improvements to the spectroscopic quantities of interest could be obtained. In this Letter, we show how this goal can be reached.

We start from the Dyson equation

$$G^{-1}(12) = G_0^{-1}(12) - V(1)\delta(12) - \Sigma(12), \qquad (1)$$

where (12) stands for two space, time, and spin coordinates $(\mathbf{r}_1 t_1 \sigma_1, \mathbf{r}_2 t_2 \sigma_2)$, G_0 is the free-particle Green's function, and $V(1) = U(1) + V_{\rm H}(1)$, the total classical potential,

 © 2005 The American Physical Society

where U is a time-dependent external potential that goes to the static physical potential at times $\pm\infty$ and whose fictitious time-dependent part is made vanishing at the end of the derivation, and $V_{\rm H}$ is the Hartree potential. One can express the self-energy in terms of variations of the Green's function with respect to the external potential, $\Sigma = -i\upsilon G\delta G^{-1}/\delta U$ [2], or

$$\Sigma(12) = iG(14)\Gamma(42;5)\frac{\delta V(5)}{\delta U(3)}\upsilon(31^+), \tag{2}$$

with the irreducible vertex function

$$\Gamma(12;3) = -\frac{\delta G^{-1}(12)}{\delta V(3)} = \delta(13)\delta(23) + \frac{\delta\Sigma(12)}{\delta V(3)} \tag{3}$$

and υ the bare Coulomb interaction (integration over indices not present on the left is implicit throughout this Letter). Disregarding Σ on the right-hand side of Eq. (3) yields the GW approximation.

The derivative $\delta\Sigma/\delta V$ is usually replaced by the chain rule $(\delta\Sigma/\delta G)(\delta G/\delta V)$, which, using the relation between the derivatives of G and of G^{-1}, leads to the term $(\delta\Sigma/\delta G)GG\Gamma$ that transforms Eq. (3) into an integral equation for Γ [2]. This equation, or an equivalent form, with its four-point kernel dominated by $\delta\Sigma(12)/\delta G(34)$, has to be solved in order to get the irreducible polarizability $P = -iGG\Gamma$ and an improved self-energy from Eq. (2). This is the main obstacle on the way to a calculation of polarizabilities or self-energies beyond the RPA.

The fundamental idea of the present Letter is to benefit from the Runge-Gross theorem of TDDFT [10] *in order to rewrite $\delta\Sigma/\delta V$ in Eq.* (3). The one-to-one relation between time-dependent densities and external potentials, or consequently between the densities and the classical potentials V, allows one to use an alternative chain rule to express $\delta\Sigma/\delta V$, namely, $(\delta\Sigma/\delta\rho)(\delta\rho/\delta V)$ [12]. This transformation is hence exact whenever the linear response version of TDDFT, which is the only ingredient needed here, is exact [13]. It leaves observables like QP energies of the original equations accessible and suggests straightforward approximations. Equation (3) hence becomes

$$\Gamma(12;3) = \delta(13)\delta(23) + \frac{\delta\Sigma(12)}{\delta\rho(4)}P(43), \tag{4}$$

where $P = \delta\rho/\delta V$ is the irreducible polarizability that, as explained above, is usually calculated by solving the vertex equation. However, by integrating Eq. (4) with two Green's functions G, one directly obtains

$$P(12) = P_0(12) + P_0(13)f_{\rm xc}^{\rm eff}(34)P(42), \tag{5}$$

with $P_0(12) = -iG(12)G(21)$ and the two-point kernel

$$f_{\rm xc}^{\rm eff}(34) = -iP_0^{-1}(36)G(65)G(5'6)\frac{\delta\Sigma(55')}{\delta\rho(4)}. \tag{6}$$

In other words, *one can now first determine the two-point irreducible polarizability P from the integral Eq.* (5)*, and subsequently the three-point vertex Γ via the integration of Eq.* (4). From P, the reducible polarizability $P^{\rm red}$ [15] is obtained via $P^{\rm red} = P + P\upsilon P^{\rm red}$.

Finally, the self-energy becomes

$$\Sigma(12) = iG(12)W^{\text{TC-TC}}(21) + iG(14) \times \frac{\delta\Sigma(42)}{\delta\rho(5)}P^{\rm red}(53)\upsilon(31+). \tag{7}$$

The first term has the GW form, but with the test charge–test charge (TC-TC) screened Coulomb interaction $W^{\text{TC-TC}} = (1 + \upsilon P^{\rm red})\upsilon$, instead of the RPA one. It has been discussed [8,9] that $\Sigma = iGW^{\text{TC-TC}}$ would yield unreliable results, because of the cancellation effects coming from the second term. In fact, the term $\upsilon P^{\rm red}\upsilon$ contributing to $W^{\text{TC-TC}}$ creates the induced Hartree potential felt by a classical charge. The additional term $(\delta\Sigma/\delta\rho)P^{\rm red}$ is responsible for the missing induced xc potentials that act on an electron or hole. It is therefore useful to reformulate Eq. (4) as

$$\Gamma(12;3) = \delta(13)\delta(23) + \delta(12)f_{\rm xc}^{\rm eff}(14)P(43) + \Delta\Gamma(12;3), \tag{8}$$

where

$$\Delta\Gamma(12;3) = \left(\frac{\delta\Sigma(12)}{\delta\rho(4)} - \delta(12)f_{\rm xc}^{\rm eff}(14)\right)P(43). \tag{9}$$

The most important effects are, in fact, contained in the first two (one- and two-point) contributions to Γ (called $\Gamma^{(2)}$ in the following), whereas the three-point remainder $\Delta\Gamma$ can be interpreted as a subsequent "nonlocality" correction. $\Delta\Gamma$ has no effect on P, as one can see by integrating Eq. (9) with two Green's functions and using Eq. (6). In the self-energy of Eq. (7), the inclusion of $\Gamma^{(2)}$ leads to $\Sigma = iG\tilde{W}$ with a modified screened Coulomb interaction $\tilde{W} = [1 + (\upsilon + f_{\rm xc}^{\rm eff})P^{\rm red}]\upsilon$. This is a test charge–test electron (TC-TE) screened Coulomb interaction instead of $W^{\text{TC-TC}}$; this expresses the fact that an additional electron or hole in the system cannot be described as a classical charge. $\Delta\Gamma$ yields then in Σ a correction term to this physically intuitive contribution; it is entirely due to the nonlocality of the self-energy in Eq. (9).

Equation (5) is a two-point, but QP-derived, equation for the polarizability. The link with TDDFT can be made by the fact that the diagonal of G yields the exact time-dependent density $-iG(11^+) = \rho(1)$ [16]. $\delta G/\delta\rho = -G(\delta G^{-1}/\delta\rho)G$ leads to

$$iG(13)G(41^+)\frac{\delta G^{-1}(34)}{\delta\rho(2)} = \delta(12). \tag{10}$$

Since the same exact density, and hence the same Hartree potential, should also be obtained from the Kohn-Sham potential $\upsilon_{\rm KS} = V + \upsilon_{\rm xc}$, we can write

$$G^{-1}(12) = G_0^{-1}(12) - \delta(12)[\upsilon_{\rm KS}(1) - \upsilon_{\rm xc}(1)] - \Sigma(12). \tag{11}$$

As $\delta G_0^{-1}/\delta\rho = 0$, Eq. (10) becomes

$$P_0(13)\chi_0^{-1}(32) - iG(13)G(41^+)\frac{\delta\Sigma(34)}{\delta\rho(2)} - P_0(13)f_{\rm xc}(32) = \delta(12), \quad (12)$$

where $\chi_0(12) = \delta\rho(1)/\delta v_{\rm KS}(2)$ is the KS independent particle polarizability and $f_{\rm xc}(12) = \delta v_{\rm xc}(1)/\delta\rho(2)$ is the xc kernel of TDDFT. This kernel turns out to consist of two terms, namely, $f_{\rm xc}^{(1)}$ and $f_{\rm xc}^{(2)}$, with $f_{\rm xc}^{(2)}$ exactly equal to the $f_{\rm xc}^{\rm eff}$ arising from our previous approach and

$$f_{\rm xc}^{(1)}(12) = \chi_0^{-1}(12) - P_0^{-1}(12). \quad (13)$$

$f_{\rm xc}^{(1)}$ serves to change the KS response function into the independent QP one, in particular, to solve the so-called band gap problem. $f_{\rm xc}^{(2)}$ accounts for the electron-hole interaction. This splitting [17] is physically intuitive. Altogether, TDDFT yields then for the irreducible polarizability P

$$P = \chi_0 + \chi_0(\chi_0^{-1} - P_0^{-1} + f_{\rm xc}^{\rm eff})P. \quad (14)$$

This is equivalent to Eq. (5).

To get an explicit expression for $f_{\rm xc}^{\rm eff}$, we choose a starting approximation for the self-energy, and consistent approximations for the functional derivative of Σ and for G, on the right-hand side of Eqs. (6) and (7). A simple choice could be to take Σ, G, and P_0 as derived from a *local* and *adiabatic* xc potential, e.g., the LDA one. This leads, of course, to the TDLDA and the $GW\Gamma$ approach of Ref. [9]. A better choice is to start from the GW approximation for Σ, taking W as a screened (e.g., static RPA) Coulomb interaction. For the functional derivative, it is then reasonable (i) to neglect the derivative of W as is usually done in the BSE: (ii) to approximate $\delta G/\delta\rho = -G(\delta G^{-1}/\delta\rho)G$ by $GP_0^{-1}G$, truncating the chain of derivatives $\delta\Sigma/\delta\rho$ that would appear if one continued to calculate all terms of $\delta G^{-1}/\delta\rho$ and that would lead to an integral equation similar to Fig. 2(b) of Ref. [17]. (Note that this is equivalent to supposing G is created by a local potential.) We obtain hence from Eq. (6)

$$f_{\rm xc}^{\rm eff}(34) = P_0^{-1}(36)G(65)G(5'6)W(55') \times G(57)G(75')P_0^{-1}(74). \quad (15)$$

Equation (15) is the electron-hole xc kernel of Refs. [11,17]. In those works, $f_{\rm xc}^{\rm eff}$ was derived in two completely different ways (by mapping matrix elements of the BSE electron-hole interaction on those of TDDFT and by a first-order expansion in W, respectively) and extensively tested for the dielectric function, with excellent results. In other words, the results of Refs. [11,17] yield the numerical validation of the present derivation, whereas, in turn, the latter *a posteriori* yields new physical insight about *why* the former had led to such (unexpectedly) good results: *the physics of the variation of the self-energy upon excitation, which gives rise to the electron-hole interaction, can be captured in terms of density variations only.* This is very important, since it encourages the use of the present scheme also for cases where one wishes to go beyond the above approximations.

It is interesting to see what one obtains when using this scheme to go beyond the GW approximation for the calculation of band gaps in semiconductors and insulators. Since systematic $GW\Gamma$ studies in literature are available only for a short-range (LDA) kernel [9], we provide here a discussion on the influence of a *long-range* contribution on QP energies.

For illustration, we present in Tables I and II results on bulk silicon and solid argon (obviously, the effect of a long-range contribution is particularly interesting in a solid, and silicon and argon represent two extreme cases, the first one with strong screening and continuum excitons, the second one with almost no screening and strongly bound electron-hole pairs).

The first series of results, presented in Table I, uses LDA as the starting approximation for the right-hand side of Eq. (7), whereas the second series in Table II uses the static but nonlocal "Coulomb-hole-plus-screened exchange" (COHSEX) approximation to GW [2]. In the latter case, we use the kernel $f_{\rm xc}^{\rm eff}$ given by Eq. (15), which, although approximate, has the correct long-range behavior [11,18]. Furthermore, LDA wave functions are used throughout: we suppose them to be similar to the COHSEX QP ones [19].

The two tables show the band gap at Γ for both materials under study. The first column gives the band gap that is obtained from the respective starting approximation (i.e., LDA or COHSEX). The second column uses this band structure, and provides the subsequent standard non-self-consistent $GW^{\rm RPA}$. Columns 3 and 4 show the band gap for the approximations to the self-energy derived in this work, using either $W^{\rm TC\text{-}TC}$ [first part of Eq. (7)] or $\tilde{W}$ (neglect of only $\Delta\Gamma$). Finally, the experimental value is given in the last column [21]. Both materials show similar tendencies. In particular, there is a significant influence of the single-

TABLE I. Direct gap (in eV) at Γ in bulk silicon and solid argon, calculated using a local approximation (LDA) for the starting self-energy (see text).

	LDA	$GW^{\rm RPA}$	$GW^{\rm TC\text{-}TC}$	$G\tilde{W}$	Expt.
Si	2.53	3.17	3.08	3.18	3.40
Ar	8.18	12.95	12.64	12.75	14.2

TABLE II. Same as Table I, but based on a nonlocal approximation (COHSEX) for the starting self-energy.

	COHSEX	$GW^{\rm RPA}$	$GW^{\rm TC\text{-}TC}$	$G\tilde{W}$	Expt.
Si	3.64	3.30	3.18	3.32	3.40
Ar	14.85	14.00	14.16	14.76	14.2

particle energies on the GW^{RPA} (second columns). The choice of COHSEX energies in W simulates the effect of the contribution $f_{\mathrm{xc}}^{(1)}$ of Eq. (13). In most cases, the electron-hole vertex correction $f_{\mathrm{xc}}^{\mathrm{eff}} = f_{\mathrm{xc}}^{(2)}$ in $W^{\mathrm{TC\text{-}TC}}$ closes the gap (third columns) with respect to RPA. When $f_{\mathrm{xc}}^{\mathrm{eff}}$ is included according to Eq. (13) in order to evaluate the explicit vertex in $\Sigma = iGW^{\mathrm{TC\text{-}TC}}\Gamma^{(2)} = iG\tilde{W}$ (fourth columns), there is a strong opening of the gap. Our most complete result is hence determined by a series of cancellations. The overall behaviors of both kernels under study (arising from LDA or the nonlocal COHSEX scheme) are very similar, even though the LDA kernel does not have the crucial, correct long-range contribution [18]. These results roughly justify calculations using the RPA GW form constructed with QP energies instead of KS ones. The $GW\Gamma$ gap turns out to be slightly bigger than the experimental value. In order to obtain improved agreement, one should, of course, avoid some of the above approximations; in particular, we expect the nonlocality correction to decrease the gap, since the neglected term should reduce the effect of the external vertex. Those and other more sophisticated numerical calculations (including, e.g., self-consistency in the wave functions) are, however, beyond the scope of this illustration.

In conclusion, using the concept of the density as a crucial quantity, we have derived a complete new set of equations for the many-body vertex, polarizability, and self-energy. This approach does not require the solution of integral equations containing a four-point kernel. In particular, the polarizability is directly obtained from a two-point equation, containing a two-point many-body kernel $f_{\mathrm{xc}}^{\mathrm{eff}}$, which completely changes the way, e.g., excitonic effects can be calculated. We have shown that the same expression for the polarizability can also be derived from the relation between the Green's function and the charge density. Moreover, the latter derivation yields the exchange-correlation kernel of TDDFT, which turns out to differ from $f_{\mathrm{xc}}^{\mathrm{eff}}$ by a term that is essentially responsible for the gap correction. Our approach explains the success of previously published approximations for the kernel and allows one to go beyond in a systematic way. On the other hand, it opens the way for better approximations to the self-energy and other many-body quantities. For the gap corrections in bulk silicon and solid argon, we have put into evidence cancellation effects of different contributions to the vertex corrections beyond the GW approximation.

We are grateful for discussions with C.-O. Almbladh, U. von Barth, J. F. Dobson, A. Marini, A. Rubio, G. Stefanucci, R. van Leeuwen, and N. Vast, support from the the EU's 6th Framework Programme through the NANOQUANTA Network of Excellence (NMP4-CT-2004-500198), and computer time from IDRIS (Project No. 544).

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[2] L. Hedin, Phys. Rev. **139**, A796 (1965); G. Strinati, Riv. Nuovo Cimento **11**, 1 (1988).
[3] W. G. Aulbur, L. Jonsson, and J. W. Wilkins, Solid State Phys. **54**, 1 (1999), and references therein.
[4] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002), and references therein.
[5] See, e.g., F. Aryasetiawan, Phys. Rev. B **46**, 13 051 (1992).
[6] A. Marini and A. Rubio, Phys. Rev. B **70**, 081103(R) (2004).
[7] E. L. Shirley and R. M. Martin, Phys. Rev. B **47**, 15 404 (1993); R. T. M. Ummels, P. A. Bobbert, and W. van Haeringen, Phys. Rev. B **57**, 11 962 (1998).
[8] G. D. Mahan and B. E. Sernelius, Phys. Rev. Lett. **62**, 2718 (1989).
[9] R. Del Sole, L. Reining, and R. W. Godby, Phys. Rev. B **49**, 8024 (1994).
[10] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
[11] L. Reining *et al.*, Phys. Rev. Lett. **88**, 066404 (2002); F. Sottile, V. Olevano, and L. Reining, Phys. Rev. Lett. **91**, 056402 (2003); G. Adragna, R. Del Sole, and A. Marini, Phys. Rev. B **68**, 165108 (2003); A. Marini, R. Del Sole, and A. Rubio, Phys. Rev. Lett. **91**, 256402 (2003).
[12] It should be stressed that the derivative of Σ with respect to ρ in the density-functional sense is different from computing the derivative of $\Sigma(12)$ with respect to $G(34)$ and then taking the limit $4 \rightarrow 3^+$. Such a procedure would, indeed, lead to a rough approximation.
[13] An important requirement of TDDFT is a well defined initial state. This is not an issue here as the initial state is, by definition of the equilibrium Green's function, the N particle ground state in the absence of any time-dependent perturbing potential U. Moreover, the linear response TDDFT needs weaker requirements than the general case [14]. Limiting cases as the example of a stationary current in an infinite solid, which one might not be able to handle in TDDFT, might instead require additional terms in the functional derivative; this is, however, beyond the scope of this work.
[14] R. van Leeuwen, Int. J. Mod. Phys. B **15**, 1969 (2001).
[15] The type of time dependence of quantities appearing in the equations has to be specified; a clean way is the formulation on the Keldysh contour [14]. In any case, one has, of course, to use the adequate form when evaluating the formulas in practice.
[16] R. van Leeuwen, Phys. Rev. Lett. **76**, 3610 (1996).
[17] R. Stubner, I. V. Tokatly, and O. Pankratov, Phys. Rev. B **70**, 245119 (2004).
[18] Ph. Ghosez, X. Gonze, and R. W. Godby, Phys. Rev. B **56**, 12 811 (1997).
[19] We do not enter here the debated subject of fully self-consistent QP calculations (see, e.g., [20]).
[20] B. Farid, Philos. Mag. B **76**, 145 (1997); B. Holm and U. von Barth, Phys. Rev. B **57**, 2108 (1998); W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **89**, 126401 (2002).
[21] R. R. L. Zucca and Y. R. Shen, Phys. Rev. B **1**, 2668 (1970); R. Haensel, Phys. Rev. Lett. **23**, 1160 (1969).