Total Energies from *GW* Calculations

Bengt Holm

Department of Theoretical Physics, University of Lund, S-223 62 Lund, Sweden
and Condensed Matter Theory Group, Department of Physics, Uppsala University, Box 530, S-751 21 Uppsala, Sweden
(Received 21 December 1998)

The possibilities of obtaining accurate total energies from *GW* calculations at various levels of self-consistency are investigated. In the conserving approximation known as the fully self-consistent *GW* approach there is no ambiguity in calculating the total energy, and the results for the electron gas compare well with those of accurate Monte Carlo calculations. We demonstrate how to obtain chemical potentials and total energies of similar accuracy from partially self-consistent (GW_0) calculations.

The *GW* approximation has for some time now been a major method to go beyond mean-field theories in order to describe many-body effects in solids. Thus, e.g., the band structure of different types of materials, like semiconductors [1], alkali [2] and transition metals [3], as well as insulators [4], has been obtained using this method. However, as originally proposed, the *GW* approximation implies a scheme of self-consistency [5], which usually has not been invoked in the applications. Instead, the self-energy, which is the quantity that describes the many-body effects, is obtained from a convolution over a Green's function G and a screened interaction W, both most often obtained from some first-principle mean-field calculation. This way of treating the *GW* approximation was applied to the electron gas already in the 1960s [6], a study which has served as a valuable point of reference ever since, also in this work. In spite of the neglect of self-consistency, the results have generally been satisfactory. However, they depend on the input parameters, and there is no general theoretical justification for the good results that are obtained. For this reason, recently, the effects of self-consistency have been studied for some different systems, such as the electron gas [7–9], various model systems [10,11], and even a real system, bulk silicon [12]. The significance of self-consistency has also been investigated in the case of semicore states [13]. It is by now well known that self-consistency actually worsens some key results of a *GW* calculation, e.g., the occupied bandwidth is increased, quasiparticle weight is transferred towards the edge of the spectra, plasmon structure is not improved, rather the opposite tends to be true. All this is contrary to what is expected from exact results.

However, it has been hinted that a self-consistent treatment of the *GW* approximation could produce good total energies, at least for the electron gas [8]. In this Letter it is shown that this indeed is the case, and some generally useful methods and observations of total energy calculations are presented. The results of two studies are used, Refs. [7] and [8]. These were mainly concerned with the quality and the accuracy of the description of the quasiparticles and the satellite structure in the one-electron Green's function, in order to direct our attention towards the accuracy of the total ground-state energies.

In Ref. [8], an approximation known as the fully self-consistent *GW* method was examined. In this method, the self-energy is constructed in a Hartree-Fock like fashion but by replacing the bare Green's function with the resulting interacting Green's function. In addition, the bare Coulomb interaction is replaced by a dynamically screened interaction W. The latter is constructed as within the normal RPA but with the constituent Green's function equal to that which one obtains from Dyson's equation involving the resulting self-energy. This approximation is conserving in the spirit of Kadanoff and Baym [14,15]. Although the method has several valuable physical properties, especially when one studies the response to external perturbations, it also has major shortcomings as discussed above. These also persist when a real system is studied [12].

In Ref. [7], a partially self-consistent scheme was investigated, which, in the present work, will be referred to as the GW_0 method. This method is similar to the *GW* method defined above, but the screened interaction W is no longer involved in the self-consistency procedure. Instead, W is kept fixed ($=W_0$) at the RPA level, i.e., W_0 is simply obtained from the Lindhard function for the polarizability of the electron gas. The resulting approximation is not conserving but the description of quasiparticles and satellite structures is reasonably good. Moreover, there is some evidence that self-consistency with regard to the Green's function is an important ingredient in the understanding of systems with localized and highly correlated electrons [4], as well as for semicore states [13]. Also, the application of the GW_0 method to real system constitutes a drastic simplification to the *GW* method. Thus, it would be highly desirable to be able to obtain accurate total energies from the GW_0 method.

As mentioned above, we use the results of the scheme which have been referred to as the G_0W_0, or the first-iteration method as a reference [6]. This scheme corresponds to a straightforward calculation of the electronic self-energy through the noninteracting Green's function G_0 and the same fixed W_0 as discussed above. The

0031-9007/99/83(4)/788(4)$15.00

resulting Green's function is obtained from Dyson's equation with this self-energy. This method corresponds to the way in which GW calculations are routinely performed in real systems, i.e., self-consistency is not at all attempted.

Most textbooks on many-body perturbation theory (MBPT) explain how one, in principle, can obtain the total ground-state energy from a knowledge of only the one-electron Green's function, assuming the validity of certain exact relations between the one- and two-particle Green's functions. They also show how to obtain the total energy directly from the two-particle Green's functions. In approximate theories, exact relations might be only poorly obeyed or not at all. As a consequence, different ways of calculating the total energy might give very different results, and one is faced with the task of deciding which result is to be preferred. One of the virtues of a conserving approximation like, e.g., the self-consistent GW method discussed above, is that this ambiguity in the total energy does not arise. All practically conceivable ways of obtaining the total energy yield the same result. On the other hand, the perhaps more useful GW_0 method is not conserving and has the mentioned problem. As we shall see, however, there are well-defined ways of obtaining the total energy also within this method.

The total energy of a many-electron system described by the one-particle Green's function G is given by

$$E = \sum_{\mathbf{k}} \int_{-\infty}^{\mu} (\epsilon_{\mathbf{k}} + \omega) A(\mathbf{k}, \omega) \, d\omega, \tag{1}$$

with the shorthand notation $\sum_{\mathbf{k}} \equiv \Omega \int d^3k (2\pi)^{-3}$. The usual Lehmann representation of the Green's function G in terms of the spectral function $A(\mathbf{k}, \omega)$ has been introduced, $G(\mathbf{k}, \omega) = \int \frac{A(\mathbf{k},\omega') d\omega'}{\omega - \omega' + i\delta \, \mathrm{sgn}(\omega' - \mu)}$.

The quantity μ is the chemical potential given by the solution to the quasiparticle equation at the Fermi surface, $k = |\mathbf{k}| = k_F$:

$$\mu = \epsilon_F + \Sigma(k_F, \mu). \tag{2}$$

Here, $\epsilon_F = k_F^2/2$. Equation (1) is often referred to as the Galitskii-Migdal expression [16], and it is simply the sum of the kinetic and the potential energies both calculated from the one-electron Green's function. For the benefit of the later discussion we can also rewrite these two contributions, T and U, in terms of the spectral function $A(\mathbf{k}, \omega)$:

$$T = 2 \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} n_{\mathbf{k}}, \tag{3}$$

$$U = \sum_{\mathbf{k}} \int_{-\infty}^{\mu} (\omega - \epsilon_{\mathbf{k}}) A(\mathbf{k}, \omega) \, d\omega, \tag{4}$$

where the momentum distribution function $n_{\mathbf{k}}$ is given by

$$n_{\mathbf{k}} = \int_{-\infty}^{\mu} A(\mathbf{k}, \omega) \, d\omega. \tag{5}$$

It can be shown that knowing precisely two of the quantities T, U, or μ, the total energy can be obtained from a simple linear combination,

$$E = T + U = \frac{1}{5}[3\mu N + U] = \frac{1}{4}[3\mu N - T]. \tag{6}$$

At this stage, note that each of the equations are also valid for each component of the different energies, i.e., the noninteracting part (0), the exchange part (x), and the correlation part (c).

The total energy as a function of the usual density parameter r_s has been calculated from the Galitskii-Migdal expression [Eq. (1)] using the Green's functions of the three computational schemes defined in the introduction. The computational details are explained in [17].

In Fig. 1 and Table I we display the correlation energy per electron (ϵ_c). The result is compared to Monte Carlo data for the electron gas according to Ceperley and Alder [18], which are generally believed to be rather accurate.

This figure demonstrates that the total energies come out very well from the fully self-consistent GW scheme, in fact only mismatching somewhat in the third digit. The two other schemes are off by some 5% to 10%. Note that the method based on a partial self-consistency already constitutes an improvement on the G_0W_0 method.

In order to understand the shortcomings of the two nonconserving methods, we display the kinetic and the potential contributions to the total correlation energy per electron in separate figures (Fig. 2). These quantities are calculated from Eqs. (3) and (4). Notice that the curves in Figs. 1 and 2 are not in the same scale.

From the conserving properties of the fully self-consistent GW scheme it follows that an accurate total energy also implies accuracy in the different pieces (t_c, u_c). Accordingly, we see that this method again performs well. The mismatch in the third digit is now slightly larger, but there is a cancellation of errors when
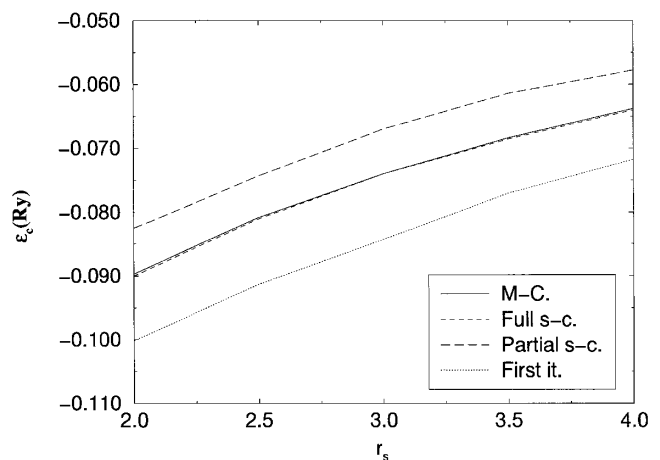

FIG. 1. The correlation part of the total energy per electron obtained through different levels of self-consistency in the GW approximation compared to Monte Carlo data. Notice the overlap of the curves corresponding to Monte Carlo data and to the fully self-consistent method.

TABLE I. The correlation energy per electron, ϵ_c, for the different GW schemes, compared to quantum Monte Carlo data.

r_s	QMC	GW	GW_0	G_0W_0
2.0	−0.0897	−0.0901	−0.0825	−0.101
4.0	−0.0638	−0.0640	−0.0578	−0.0718

the two contributions are added. Although the GW_0 scheme is particle conserving [17], it fails to reproduce the correct t_c. The G_0W_0 scheme does even worse. In order to understand these shortcomings, we study the form of the momentum distributions $n_{\mathbf{k}}$ [Eq. (5)] of the various approximations. The quantity $n_{\mathbf{k}}$ enters t_c through the factor $\sum_{\mathbf{k}} \epsilon_{\mathbf{k}}(n_{\mathbf{k}} - n_{\mathbf{k}0})$ ($n_{\mathbf{k}0}$ is the momentum distribution of the noninteracting electron gas). The quasiparticle renormalization factor, Z_F, which describes the discontinuity of $n_{\mathbf{k}}$ at k_F, is rather large in the GW method, as discussed in Ref. [8]. Therefore, as the GW method produces the correct t_c, and since Z_F decreases as we go to the GW_0 scheme, and decreases even further as we go to the G_0W_0 scheme, we can expect t_c to be increasingly too large in these schemes. This is because more and more states are transferred from below to above the Fermi surface $|\mathbf{k}| = k_F$.

In the lower part of Fig. 2 we display the correlation part of the interaction energy per electron. The result is very interesting. The schemes of full and partial self-consistency virtually yield the same result. However, most textbooks on MBPT will also show how E can be obtained as an integration over the interaction strength. By rescaling all the position variables $\mathbf{r}$ to $r_s\mathbf{r}$ in the Hamiltonian of the interacting gas, one realizes that the
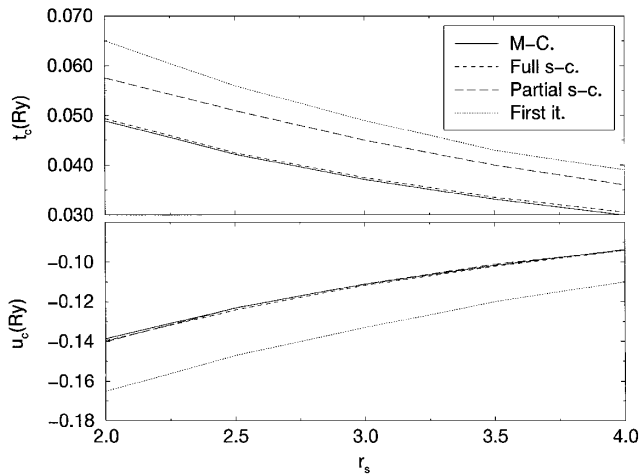

FIG. 2. The correlation part of the kinetic (upper) and interaction (lower) energy per electron obtained through different levels of self-consistency in the GW approximation compared to Monte Carlo data. Notice the overlap of the curves corresponding to Monte Carlo data and to the fully self-consistent method *and*, for the interaction energy, also for the partially self-consistent method.

total energy of the gas has the form $H(\lambda r_s, N)/r_s^2$, where λ is the strength of the Coulomb interaction. It is also immediately clear that, in the limit of a very large volume Ω, boundary effects are negligible and we must be able to write the total energy as $N_g(r_s, \lambda)$. Combining these two expressions, we realize that there exists a function $f(x)$ such that total energy E can be written $E = (N/r_s^2) \times f(\lambda r_s)$. By using the Hellmann-Feynman trick with respect to the strength λ of the Coulomb interaction we see that $U = \lambda(\partial E/\partial \lambda) = (N/r_s)f'(r_s)$, where we have set λ to one after taking the derivative after the last equal sign. Using the fact that $E = T + U$ we can write $T = (N/r_s^2)[f(r_s) - r_s f'(r_s)]$. This renders the following, for our purposes, useful formula:

$$\epsilon_c = (1/r_s^2) \int_0^{r_s} u_c(r_s') r_s' \, dr_s' . \qquad (7)$$

Thus, we can equally well obtain total energies from the computationally much simpler and physically more realistic partially self-consistent GW_0 scheme. As an example, ϵ_c has been calculated from u_c obtained from different methods of interest. The data points were fitted to a rather simple polynomial, consisting of a constant, a logarithmic, and a linear term. As a test of the accuracy, the same coarseness of the mesh was used in the exact case as in the others. Equation (7) now becomes an analytic expression, and the results are displayed in Table II. Remark that the fit was made over metallic densities only, and naturally, high densities (low r_s values) are more sensitive to the accuracy of the fit and its corresponding high density behavior.

It can be speculated whether the procedure of using the GW_0 method to calculate total energies could be extended to the case of valence electrons, by simulating different densities through a variation of the lattice parameter. That would imply that the method could be generalized to the case of real crystals where pseudopotential calculations are useful. Whether or not the above statement proves true, we cannot make any conclusion about the case of core electrons in this issue from the present work, as we have studied the gas, which is no proper model to describe them.

No similar way was found to obtain good total energies from the G_0W_0 scheme. To understand these results, we study the integrand in Eq. (4). It consists of two terms, ωA and $-\epsilon_{\mathbf{k}} A$. From the study of t_c, we know that

TABLE II. The correlation energy per electron, ϵ_c, as obtained from exact QMC data and different GW schemes through Eq. (7), compared to quantum Monte Carlo data.

r_s	QMC[a]	QMC[b]	GW	GW_0
2.0	−0.0897	−0.0891	−0.0889	−0.0919
4.0	−0.0638	−0.0637	−0.0638	−0.0643

[a]Exact.
[b]From exact u_c data through Eq. (7).

the contribution from the second term when using the GW_0 method is lower than when using the GW. This is thus compensated by the first term, which can be well understood, as the spectral function A is much more extended in the energy regime in the GW case [8], and thus the contribution from ωA in the range of integration is lower. In the case of the G_0W_0 method, the contribution from the second term is even more off the correct value. However, in this case there is no similar compensation in the first term, because now Z_F is smaller, and thus, more spectral weight transferred to the side of the main peak, as compared to the GW_0 case, and we get a more negative contribution in the integration.

We also investigate the chemical potential μ obtained through Eq. (2), and from the relation

$$\mu_c = \epsilon_c - \frac{r_s}{3} \frac{\partial \epsilon_c}{\partial r_s}. \qquad (8)$$

In Fig. 3 we display the correlation part of the chemical potential obtained directly through Eq. (2) and as obtained through Eq. (8) from the energies displayed in Fig. 1. Again, we compare to Monte Carlo data. We see that, within the GW scheme, μ_c equals the Monte Carlo result satisfactorily. We also see that μ_c from Eq. (8) is very close. However, these curves cannot claim the same accuracy as the previous energy curves, because they result either from an implicit equation or polynomial fit. To improve, coarser meshes in energy and density parameter would be required, a task which would severely increase computing time and effort.
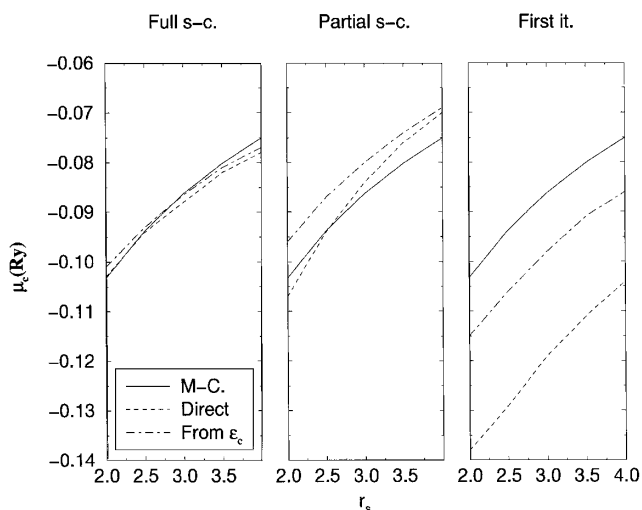

FIG. 3. The correlation part of the chemical potential obtained from a fully self-consistent GW calculation (left), partially self-consistent GW_0 calculation (middle), and from a G_0W_0 calculation, "first iteration" (right) compared to Monte Carlo data. Notice the similarity between the curve obtained directly [Eq. (2)] and the one obtained from ϵ_c [Eq. (8)] in the fully self-consistent case.

As we go on to study the corresponding curves resulting from the GW_0 and the G_0W_0 schemes, we see that the values for μ are gradually getting worse, and that the relation in Eq. (8) is more and more violated.

In summary, we have calculated total energies for the electron gas as given by the Galitskii-Migdal formula [Eq. (1)], using different levels of self-consistency in the GW approximation. We have found numerically, that for the fully self-consistent GW scheme, the total energies come out very well, and the examined consistency relations are fulfilled. As for the GW_0 method, we find that different consistency relations are much better satisfied as compared to the G_0W_0 scheme. We further present a method which uses the results of the GW_0 scheme to calculate very accurate total energies. The straightforward G_0W_0 scheme fails to produce satisfactory results. In particular, the relation between μ and ϵ [Eq. (8)] is strongly violated. We have found no simple means of obtaining correct total energies.

[1] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985); Phys. Rev. B **32**, 7005 (1985); **34**, 5390 (1986).
[2] J. E. Northrup, M. S. Hybertsen, and S. G. Louie, Phys. Rev. Lett. **59**, 819 (1987).
[3] F. Aryasetiawan, Phys. Rev. B **46**, 13 051 (1992).
[4] F. Aryasetiawan and O. Gunnarsson, Phys. Rev. Lett. **74**, 3221 (1995).
[5] L. Hedin, Phys. Rev. **139**, A796 (1965).
[6] B. I. Lundqvist, Phys. Kondens. Mater. **6**, 193 (1967); **6**, 206 (1967); **7**, 117 (1968); B. I. Lundqvist and V. Samathiyakanit, Phys. Kondens. Mater. **9**, 231 (1969).
[7] U. von Barth and B. Holm, Phys. Rev. B **54**, 8411 (1996); **55**, 10 120(E) (1997).
[8] B. Holm and U. von Barth, Phys. Rev. B **57**, 2108 (1998).
[9] E. L. Shirley, Phys. Rev. B **54**, 7758 (1996).
[10] H. J. de Groot, P. A. Bobbert, and W. van Haeringen, Phys. Rev. B **52**, 11 000 (1995).
[11] A. Schindlmayr, T. J. Pollehn, and R. W. Godby, Phys. Rev. B **58**, 12 684 (1998).
[12] W.-D. Schöne and A. G. Eguiluz, Phys. Rev. Lett. **81**, 1662 (1998).
[13] M. Rohlfing, P. Krüger, and J. Pollman, Phys. Rev. B **56**, 7065(R) (1997).
[14] G. Baym and L. P. Kadanoff, Phys. Rev. **124**, 287 (1961).
[15] G. Baym, Phys. Rev. **127**, 1391 (1962).
[16] V. Galitskii and A. Migdal, Sov. Phys. JETP **7**, 96 (1958).
[17] B. Holm and U. von Barth, in B. Holm, Ph.D. thesis, Department of Theoretical Physics, University of Lund, 1997, paper III. A copy can be requested from the author at holm@teorfys.lu.se
[18] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).