Fully self-consistent *GW* self-energy of the electron gas

B. Holm and U. von Barth
Department of Theoretical Physics, University of Lund, S-22362 Lund, Sweden
(Received 28 January 1997)

We present fully self-consistent results for the self-energy of the electron gas within the *GW* approximation. This means that the self-consistent Green's function G, as obtained from Dyson's equation, is used not only for obtaining the self-energy but also for constructing the screened interaction W within the random-phase approximation. Such a theory is particle and energy conserving in the sense of Kadanoff and Baym. We find an increase in the weight of the quasiparticle as compared to ordinary non-self-consistent calculations but also to calculations with partial self-consistency using a fixed W. The quasiparticle bandwidth is larger than that of free electrons and the satellite structure is broad and featureless; both results clearly contradict the experimental evidence. The total energy, though, is as accurate as that from quantum Monte Carlo calculations, and its derivative with respect to particle number agrees with the Fermi energy as obtained directly from the pole of the Green's function at the Fermi level. Our results indicate that, unless vertex corrections are included, non-self-consistent results are to be preferred for most properties except for the total energy.
[S0163-1829(97)04148-9]

I. INTRODUCTION

The present paper represents the second in a series of papers describing our investigations of the effects of self-consistency within the *GW* approximation for the electronic self-energy. In our previous paper,[1] we stressed that the efficiency of our present-day computers and our thorough knowledge of the one-electron structure of real solids have now enabled us to apply various approximations within many-body perturbation theory (MBPT) to real solids. As a result, the last decade has seen a wealth of calculations of the physical properties of real solids using MBPT. Most calculations employ the so-called *GW* approximation[2] (GWA) meaning that the electronic self-energy is obtained from the one-electron Green's function in a Hartree-Fock-like fashion but with a screened interaction W in place of the bare Coulomb interaction v.

$$\Sigma = iGW. \tag{1}$$

Such calculations can be carried out in a number of different ways. The one-electron Green's function is usually, and often as a matter of convenience, taken to be that (G^{LD}) obtained from a self-consistent density-functional (DF) calculation[3] based on the local-density approximation (LDA).[4] Another choice would be the Green's function obtained from Dyson's equation:

$$G = G^{\mathrm{LD}} + G^{\mathrm{LD}}\left(\Sigma - v_{\mathrm{xc}}^{\mathrm{LD}}\right)G, \tag{2}$$

where Σ is the *GW* self-energy and $v_{\mathrm{xc}}^{\mathrm{LD}}$ is the local energy-independent exchange-correlation potential of the LDA. Sometimes a hybrid scheme is used in which the Green's function is almost that of the LDA. The constituent LDA eigenenergies are, however, replaced by quantities closer to true one-electron excitation energies. Such a scheme has proven to be valuable and necessary, e.g., in NiO.[5] For the screened interaction W, on the other hand, most calculations employ the random-phase approximation (RPA) using orbitals and eigenenergies from the LDA. Thus,

$$W = v + vP_0W, \tag{3}$$

where the irreducible polarizability P_0, within the RPA, is approximated by the non-interacting density-density response function χ_0 given by

$$iP_0 = i\chi_0 = 2G^{\mathrm{LD}}G^{\mathrm{LD}}. \tag{4}$$

Sometimes, further approximations are introduced such as, e.g., describing the energy dependence of W as a sum of plasmonlike poles. Other calculations have included particle-hole interactions in the polarizability P_0. Most often such vertex corrections have been energy independent and they are usually constructed from the LDA. Since the early days of MBPT, it has, however, been known to be inappropriate to include even static vertex corrections without simultaneously dressing up the Green's function. There are strong cancellations between the self-energy effects on the Green's function and the effects of adding particle-hole interactions.

The point we wish to make here is that we have so far no *a priori* reason to prefer one or the other of all these different computational procedures. We simply try them out and judge their merits by comparing to experiment. The problem is fundamental in nature and can only be rectified by finding a systematic way of going beyond the GW A. We then remind the reader that we are dealing with a divergent or conditionally convergent perturbation expansion and an infinity of terms must always be summed in order to obtain reasonable results. Thus, physical reason and intuition must be made to bear when constructing approximations beyond the GWA. Until we have found a systematic way of proceeding, we believe it to be worthwhile to investigate the consequences of the different computational procedures proposed so far.

In the early 1960s, bearing in mind transport properties, Kadanoff and Baym[6,7] proposed to judge the merits of different approximations by their ability to conserve quantities

such as particle number, energy, and momentum under the influence of external perturbations. Several such schemes are known, such as the Hartree-Fock approximation, the self-consistent *T*-matrix approach, and the self-consistent GWA.

The essence of the latter method is that the Green's function should be obtained from Dyson's equation [Eq. (2)] and the screened interaction *W* should be obtained from the RPA but again using the same interacting Green's function. Thus,

$$iP_0 = 2GG. \tag{5}$$

Until now, this approximation has not been applied to any system with realistic screening properties. An interesting first attempt was made by de Groot *et al.*,[8] who applied the method to a semiconducting wire. We speculate that the lack of an efficient computational procedure has discouraged people from undertaking this interesting task. Their hesitation could, however, also have a deeper reason. The fully self-consistent procedure clearly violates the rule stated above. Within this scheme, the polarizability P_0 is constructed from dressed-up Green's functions, i.e., Green's functions with self-energy insertions, without including the corresponding vertex corrections. As a consequence, the resulting screened interaction *W* is unphysical and, e.g., it does not obey the *f*-sum rule like a normal response function. Within the self-consistent GWA, *W* is thus merely an auxiliary quantity for obtaining a, hopefully, better self-energy Σ.

Because of the doubts raised above about the fully self-consistent procedure, our first paper dealt with a partial self-consistency in which the Green's function entering the calculation of the self-energy was that obtained from Dyson's equation. The screened interaction W_0 was, however, kept fixed at the level of the RPA with noninteracting Green's functions. Thus, W_0 is a plausible screened interaction that, e.g., obeys the *f*-sum rule. The resulting self-consistent Green's function gave a smaller reduction of the strength of the quasiparticles as compared to non-self-consistent calculations and the bandwidth was 5–10% larger than that of free electrons. Another undesirable feature of that calculation was that it resulted in a total energy whose derivative with respect to particle number did not correspond to the calculated Fermi energies.

The last deficiency mentioned in the previous paragraph is remedied in our present fully self-consistent calculations. Furthermore, we obtain very accurate total energies, i.e., the energies of these calculations are very close to those of sophisticated quantum Monte Carlo calculations.[9] Otherwise, we obtain a slightly inferior description of the quasiparticles and the satellite part of the one-electron spectrum becomes much worse. Thus, the new results do not justify the very large increase in the computational effort caused by going to full self-consistency.

The outline of the paper is the following: In Sec. II we go through the theory and mention a few computational details and in Sec. III we present the results. Then, in Sec. IV we compute the total energy from the Galitskii-Migdal[10] expression and make a few remarks of theoretical nature and in Sec. V we give our conclusions and speculate on future developments.

II. THEORY AND BASIC PROCEDURES

In the present section, we will summarize the basic theoretical expressions needed in this work. We will also give a few details concerning the computational methods. Most of the formulas have appeared in previous work[1] with comments and, in some cases, derivations. We remind the reader that the two features that change this project from a formidable challenge into a reasonable computational task are as follows: (1) Working with the electron gas, the necessary **k**-space integrations become much simpler integrations over one polar angle and the length of the **k** vector, and (2) our physical way of representing the Green's function in terms of its real non-negative spectral function which, in turn, is represented as a sum of Gaussian functions. The latter representation has many advantages.

(a) A true spectral function of the electron gas has a main quasiparticle peak of weight less than unity and a series of successively smaller satellite peaks corresponding to the shakeup mainly of plasmons and, to a smaller extent, of particle-hole pairs. Such a structure is easily modeled by a series of Gaussians.

(b) Any real non-negative function can be represented by a sum of Gaussians, albeit, sometimes with a very large number of terms.

(c) The necessary frequency integrals involving products of spectral functions at different momenta can be carried out analytically regardless of the sharpness of the structures.

(d) Most importantly, the sensitivity of the spectral function that one obtains from the calculation is relatively insensitive to details of the spectral function that is used as input to the calculation. This was the case in our previous work on partial self-consistency, a circumstance that carries over to the present case and that greatly facilitates the numerical procedures.

As an assistance to the reader we list the relevant formulas below.

The Green's function $G(\mathbf{k},\omega)$ is obtained from the spectral function $A(\mathbf{k},\omega)$ according to the usual Lehman representation,

$$G(\mathbf{k},\omega) = \int_C \frac{A(\mathbf{k},\omega')}{\omega-\omega'}\, d\omega', \tag{6}$$

where *C* is the contour in the complex frequency plane defined as a straight line from $-\infty$ to 0 just above the real axis and another straight line from 0 to ∞ just below the real axis. At this point we mention that here and throughout this paper, the zero of energy is always chosen at the Fermi energy. Thus, all propagators, fermion or boson, interacting or noninteracting, always change their analytic structure at $\omega=0$.

It follows directly from the approximation defined by Eq. (5) that the corresponding irreducible polarization propagator $P_0(\mathbf{q},\omega)$ has a spectral representation of the usual form

$$P_0(\mathbf{q},\omega) = \int_C \frac{S_0(\mathbf{q},\omega')}{\omega-\omega'}\, d\omega' \tag{7}$$

in terms of a spectral function $S_0(\mathbf{q},\omega)$ given by

$$S_0(\mathbf{q},\omega) = 2\sum_{\mathbf{k}} \int_0^{\omega} A(\mathbf{k},\omega'-\omega)A(\mathbf{k}+\mathbf{q},\omega')\, d\omega'. \tag{8}$$

Note that $\Sigma_{\mathbf{k}}$ is short for $\int d^3k(2\pi)^{-3}$ and the factor of two is associated with the spin degeneracy.

Also the screened interaction $W(\mathbf{q},\omega)$ has a similar spectral representation,

$$W(\mathbf{q},\omega)=v(\mathbf{q})+\int_C \frac{B(\mathbf{q},\omega')}{\omega-\omega'}d\omega'. \tag{9}$$

It follows from Eq. (3) that the spectral function $B(\mathbf{q},\omega)$, in this case, is given by

$$B(\mathbf{q},\omega)=\frac{v^2(\mathbf{q})S_0(\mathbf{q},\omega)}{|1-v(\mathbf{q})P_0(\mathbf{q},\omega)|^2}, \tag{10}$$

where $v(\mathbf{q})=4\pi/q^2$ is the Fourier transform of the Coulomb interaction.

Finally, the spectral decomposition of the self-energy $\Sigma(\mathbf{k},\omega)$ follows directly from the GWA, i.e., from Eq. (1),

$$\Sigma(\mathbf{k},\omega)=\Sigma_{\mathrm{HF}}(\mathbf{k})+\int_C \frac{\Gamma(\mathbf{k},\omega')}{\omega-\omega'}\,d\omega' \tag{11}$$

in terms of the spectral function $\Gamma(\mathbf{k},\omega)$, which, in the fully self-consistent case has the form

$$\Gamma(\mathbf{k},\omega)=\sum_{\mathbf{q}}\int_0^{\omega} A(\mathbf{k}+\mathbf{q},\omega-\omega')B(\mathbf{q},\omega')d\omega'. \tag{12}$$

As in previous work[1], the Hartree-Fock self-energy $\Sigma_{\mathrm{HF}}(\mathbf{k})$ is given by

$$\Sigma_{\mathrm{HF}}(\mathbf{k})=-\sum_{\mathbf{q}} v(\mathbf{q})n_{\mathbf{k}+\mathbf{q}} \tag{13}$$

in terms of the momentum distribution function $n_{\mathbf{k}}$,

$$n_{\mathbf{k}}=\int_{-\infty}^{0} A(\mathbf{k},\omega)d\omega. \tag{14}$$

In our previous work[1] we derived a few sum rules of pertinence also to the present investigation. For instance, it follows from the fact that the spectral function $A(\mathbf{k},\omega)$ is normalized to unity and from Eq. (12) that

$$\int_{-\infty}^{\infty}\Gamma(\mathbf{k},\omega)d\omega=\sum_{\mathbf{q}}\int_0^{\infty} B_0(\mathbf{q},\omega)d\omega. \tag{15}$$

From Eq. (8) we similarly obtain

$$\int_0^{\infty} S_0(\mathbf{q},\omega)\omega d\omega=2\sum_{\mathbf{k}} n_{\mathbf{k}}(1-n_{\mathbf{k}+\mathbf{q}})[E^p_{\mathbf{k}+\mathbf{q}}-E^h_{\mathbf{k}}], \tag{16}$$

where we have defined the average particle ($E^p_{\mathbf{k}}$) and hole ($E^h_{\mathbf{k}}$) energies through the integrals

$$\int_0^{\infty} A(\mathbf{k},\omega)\omega d\omega=(1-n_{\mathbf{k}})E^p_{\mathbf{k}},$$
$$\int_{-\infty}^{0} A(\mathbf{k},\omega)\omega d\omega=n_{\mathbf{k}}E^h_{\mathbf{k}}. \tag{17}$$

It is clear that when $\mathbf{k}$ is below the Fermi surface ($k<k_F$), there is very little spectral weight at positive energies, i.e., above the Fermi energy. This fact is accounted for by the factor $1-n_{\mathbf{k}}$ in the first equation just above. Notice, however, that due to the interactions, this factor never vanishes below k_F but it is close to zero at the bottom of the Fermi sea ($k=0$). Similarly, there is very little spectral weight below the Fermi level at momenta above the Fermi surface. This is the motivation for the factor $n_{\mathbf{k}}$ in the second equation above. The momentum distribution function $n_{\mathbf{k}}$ approaches zero rather slowly at large momenta $\mathbf{k}$ but it never vanishes (see Fig. 11). In the free-electron case, both energies E^p and E^h are simply $k^2/2-k_F^2/2$, but in the present case, they are both closer to the Hartree-Fock energy $E^{\mathrm{HF}}_{\mathbf{k}}=\epsilon_{\mathbf{k}}+\Sigma_{\mathrm{HF}}(\mathbf{k})$.

Bearing these facts in mind we see from Eq. (16) that its right-hand side is not proportional to q^2 and it does not even tend to zero with q. This demonstrates the breakdown of the f-sum rule. We stress, however, that this breakdown does not preclude a reasonable description of the quasiparticle properties as we shall soon see.

Incidently, it follows from the similar analytic structure and large frequency behavior of $P_0(\mathbf{q},\omega)$ and $W(\mathbf{q},\omega)$ [Eq. (3)] that

$$\int_0^{\infty} B(\mathbf{q},\omega)\omega d\omega=v^2(\mathbf{q})\int_0^{\infty} S_0(\mathbf{q},\omega)\omega d\omega. \tag{18}$$

We have found this sum rule and the one given in Eq. (15) to be quite convenient as a check on our numerical accuracy.

For the purpose of closing our self-consistency loop, we use Dyson's equation

$$G(\mathbf{k},\omega)=\frac{1}{\omega-\epsilon_{\mathbf{k}}-\Sigma(\mathbf{k},\omega)}, \tag{19}$$

where the free-electron energy $\epsilon_{\mathbf{k}}$ has been shifted by the chemical potential μ in order to stick to our convention of having the pole of the Fermi-surface Green's function at $\omega=0$. Thus,

$$\epsilon_{\mathbf{k}}=\tfrac{1}{2}k^2-\mu, \tag{20}$$

with

$$\mu=\tfrac{1}{2}k_F^2+\Sigma(k_F,0). \tag{21}$$

Using the spectral representations given by the Eqs. (6) and (11), Eq. (19) can more conveniently be written in terms of the spectral functions $A(\mathbf{k},\omega)$ and $\Gamma(\mathbf{k},\omega)$,

$$A(\mathbf{k},\omega)=\frac{\Gamma(\mathbf{k},\omega)}{|\omega-\epsilon_{\mathbf{k}}-\Sigma(\mathbf{k},\omega)|^2}. \tag{22}$$

Our computational procedure should now be evident. We start from some reasonable spectral function $A(\mathbf{k},\omega)$, e.g., the one obtained from our older procedure with partial self-consistency.[1] This $A(\mathbf{k},\omega)$ is thus given as a sum of Gaussians,

$$A(\mathbf{k},\omega)=\sum_{\nu}\frac{W_\nu(\mathbf{k})}{\sqrt{2\pi}\Gamma_\nu(\mathbf{k})}\exp\left[-\frac{[\omega-E_\nu(\mathbf{k})]^2}{2\Gamma^2_\nu(\mathbf{k})}\right]. \tag{23}$$

We then obtain $S_0(\mathbf{q},\omega)$ from Eq. (8), which amounts to a two-dimensional integral. Due to the Gaussians, the frequency integral can be done analytically and the integral over azimuthal angle simply gives a factor of 2π. The real part of the ''polarizability'' $P_0(\mathbf{q},\omega)$ is then taken from the Hilbert transform defined by Eq. (7), and the spectral function $B(\mathbf{q},\omega)$ is obtained from Eq. (10). The most time consuming task of the project is the calculation of the spectral function $\Gamma(\mathbf{k},\omega)$ from Eq. (12). This gives rise to a three-dimensional integral for each $\mathbf{k}$ and ω. Finally, the real part of the self-energy $\Sigma(\mathbf{k},\omega)$ is obtained from the Hilbert transform of Γ [Eq. (11)] where after a new spectral function $A(\omega)$ is given by Eq. (22). (If not nessecary for the context, we will sometimes supress the momentum dependence of spectral functions.) This new $A(\omega)$ is then refitted to a sum of Gaussians [Eq. (23)] and the whole procedure is repeated until no further changes occur in the output $A(\omega)$. We here stress that the final, converged $A(\omega)$ is independent of the starting $A(\omega)$ and convergence is usually reached after relatively few iterations. This is due to the fact that the output $A(\omega)$ is rather insensitive to the input $A(\omega)$ as long as the latter has the correct physical behavior. It should, e.g., have a main quasiparticle peak of approximately the correct weight and position and some kind of satellite or incoherent region carrying the remainder of the total weight $(=1)$.

At this stage one could almost anticipate the outcome of a full calculation. Suppose that we start from a spectral function $A(\omega)$ from our partial self-consistency. Such an $A(\omega)$ had a main quasiparticle peak at approximately the free-electron energy with a weight of ~ 0.7 $(r_s=4)$. In addition it had a single plasmon satellite starting at the plasma frequency ω_p below the quasiparticle energy and having its maximum approximately half a plasma frequency further down and a weight of ~ 0.3. (According to experiment, the satellite should have been smaller, more narrow, and positioned closer in energy to the quasiparticle. Also, there should have been additional smaller plasmon satellites further down in energy.) The occurrence of the plasmon satellite in our previous results is simply understood in terms of the structure of the bare or unscreened polarizability P_0 of free electrons, a quantity used in all our older calculations.[1] For free electrons, the spectral function $S_0(\omega)$ vanishes around the plasma frequency where the denominator of Eq. (10) has a root. Consequently, at the plasma frequency, there is a large δ-function contribution to the spectral function $B(\mathbf{q},\omega)$. This contribution dominates $B(\omega)$, especially at small $\mathbf{q}$, and is strongly reflected in $\Gamma(\omega)$ at ω_p below the quasiparticle [Eq. (12)]. It is this structure in $\Gamma(\omega)$ that gives the plasmon satellite in $A(\omega)$ [Eq. (22)]. Now, using instead the partially self-consistent result for $A(\omega)$ in computing $S_0(\omega)$ from the convolution integral in Eq. (12) we obtain three structures in $S_0(\omega)$. One strong feature near zero frequency corresponding to the overlap of the quasiparticle and the quasihole, one weaker feature at the plasma frequency corresponding to the overlap of a quasiparticle or hole with a plasmon, and at twice the plasma frequency there is the very weak feature corresponding to the overlap of two plasmons. Thus, even though the real part of the dielectric function $\epsilon(\mathbf{q},\omega)=1-v(\mathbf{q})P_0(\mathbf{q},\omega)$ still might have a zero near ω_p, there is now a peak in the imaginary part $S_0(\mathbf{q},\omega)$ and, therefore, no strong structure in $B(\mathbf{q},\omega)$ [Eq. (10)]. Neither is there a structure in $\Gamma(\omega)$ at ω_p below the quasiparticle and, consequently, no plasmon satellite in the spectral function $A(\mathbf{q},\omega)$. In the next iteration there will be no peak in $S_0(\omega)$ at ω_p, but still a finite value giving a larger but broad structure in B at ω_p and in $\Gamma(\omega)$ at ω_p below the quasiparticle. Thus, the incoherent structure in $A(\omega)$ will increase but there will be no sharp structure. In the next iteration again, $S_0(\omega)$ will thus increase somewhat at the plasma frequency thereby broadening the incoherent structure in $A(\omega)$. After a few further iterations things settle down to a situation with broad, featureless structures in both $S_0(\mathbf{q},\omega)$ and $A(\mathbf{k},\omega)$ as can be seen from Figs. 2, 8(a), and 8(b).

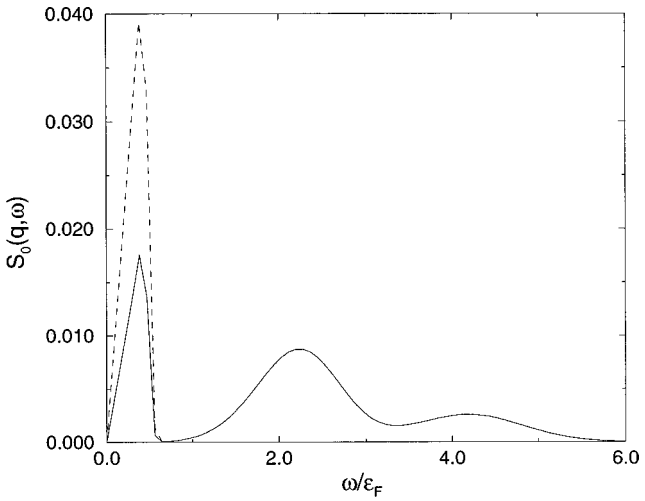

FIG. 1. The spectral function $S_0(\mathbf{q},\omega)$ (solid) of the irreducible ''polarizability,'' here obtained from a Green's function from our previous partially self-consistent calculations (Ref. 1), is compared to the corresponding quantity for noninteracting electrons (dashed). Momenta and energies are measured in units of the Fermi momentum (k_F) and the Fermi energy $(k_F^2/2)$, respectively. In these reduced units the noninteracting result is independent of r_s. Notice the peaks at the plasma frequency ω_p and at $2\omega_p$. The first peak is responsible for killing the plasmon satellite in the subsequent iteration. Here, $r_s=4.$ and $|\mathbf{q}|=0.25$.

III. RESULTS

The emphasis in the present work is on the qualitative aspects of different ways of doing *GW* calculations. We are thus mainly interested in trends and not in actual numbers. Therefore, we believe it to be instructive to present our results in the form pictures. Most of the effects in which we are interested increase with the effective strength of the Coulomb interaction, i.e., with r_s giving the average distance between electrons $[n=3/(4\pi r_s^3)]$. Therefore, most figures picture the situation at $r_s=4$ corresponding to the low electron density of sodium.

In Fig. 1 we present the spectral function $S_0(\mathbf{q},\omega)$ of the irreducible polarization propagator $P_0(\mathbf{q},\omega)$ obtained from a spectral function $A(\mathbf{k},\omega)$ with a proper plasmon structure. The quantity $S_0(\omega)$ is compared with the corresponding quantity for free electrons, which produces the well-known Lindhard function.[11] The peaked structures at the plasma frequency ω_p and at $2\omega_p$ are clearly discerned. It is the one at ω_p that is responsible for killing the plasmon structure in the

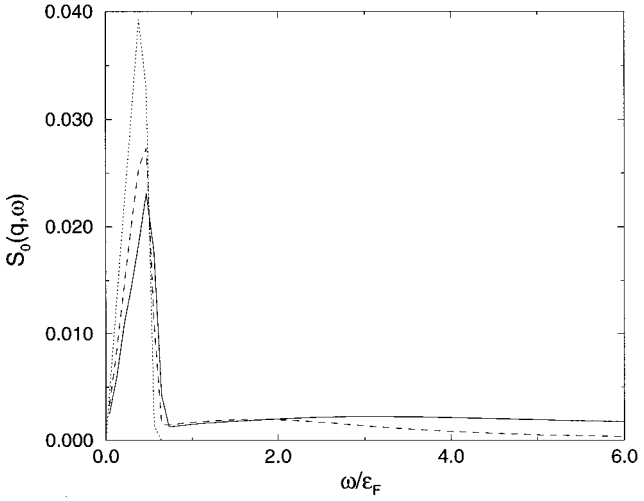

FIG. 2. The spectral function $S_0(\mathbf{q},\omega)$ of the irreducible ''polarizability'' at full self-consistency [$r_s=2$ (dashed) and $r_s=4$ (solid)] is compared to the corresponding quantity for noninteracting electrons (dotted). The latter is independent of r_s in the reduced units defined under Fig. 1 and used throughout. Notice the much more extended tail in the more strongly interacting case ($r_s=4$). Here, $|\mathbf{q}|=0.25$.

subsequent $A(\omega)$ as described in Sec. II. Thus, the fully self-consistent $S_0(\omega)$ has a normal particle-hole structure at small frequencies supplemented by a very broad structure extending to very large energies (see Fig. 2).

In Fig. 2 we also see that the extent of this tail increases with the effective strength of the Coulomb interaction, i.e., with r_s. The tail has the effect of wiping out structure in the screened version of $S_0(\omega)$ [essentially $B(\mathbf{q},\omega)$ without the $v^2(\mathbf{q})$] as seen in Fig. 4. The long tail also causes some problems concerning numerical convergence.

The unphysical nature of $S_0(\omega)$ also causes the real part of the dielectric function $\epsilon(\mathbf{q},\omega)=1-v(\mathbf{q})P_0(\mathbf{q},\omega)$ to have

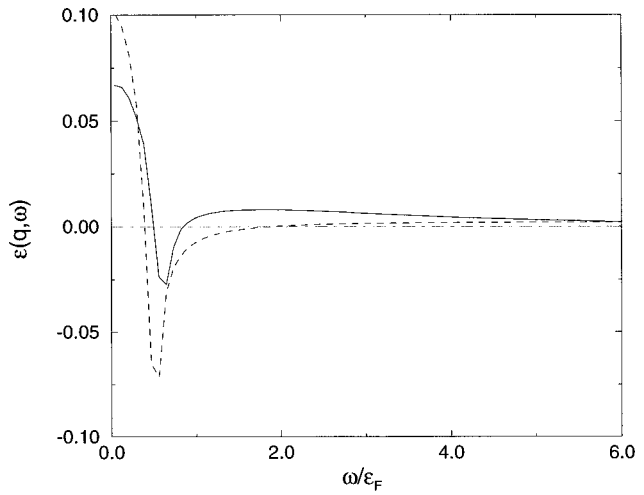

FIG. 3. The real part of the dielectric function $\epsilon=1-vP_0$ at full self-consistency (solid) is compared to the dielectric function of Lindhard (RPA) (dashed). The latter ϵ has a zero at the plasma frequency ω_p where the noninteracting $S_0(\omega)$ vanishes, giving rise to the well-known plasmon pole. At the zeros of the self-consistent ϵ the self-consistent $S_0(\omega)$ has a reasonable magnitude (Fig. 2) and the plasmon is killed. Here, $r_s=4$ and $|\mathbf{q}|=0.25$.

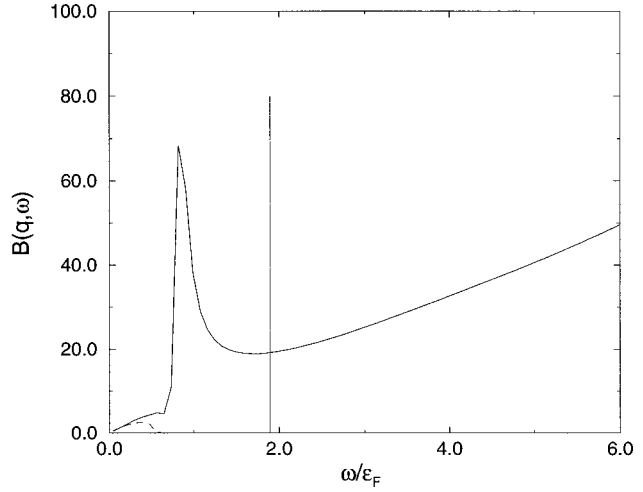

FIG. 4. The spectral function $B(\mathbf{q},\omega)$ of the screened interaction $W(\mathbf{q},\omega)$ (solid) is compared to the corresponding quantity within normal RPA (dashed). Notice the plasmon pole at ω_p in the case of the RPA. The residue of this pole is essentially the same as the weight under the entire $B(\omega)$ curve. Here, $r_s=4$, and $|\mathbf{q}|=0.25$.

several zeros, none of which occur at ω_p (see Fig. 3). Fortunately, these zeros do not have a large effect on the resulting spectral function $B(\mathbf{q},\omega)$, Fig. 4. The reason is that $S_0(\mathbf{q},\omega)$ is always finite at the zeros of $\epsilon(\mathbf{q},\omega)$. It should be noted that these anomalies occur mainly at small momentum transfers (small $\mathbf{q}$) and their possible adverse effects are washed out by the ensuing momentum integration involved in computing $\Gamma(\omega)$ from Eq. (12).

In Figs. 5(a) and 5(b), we compare the self-consistent spectral function Γ for the self-energy with the corresponding quantity obtained previously in partially self-consistent calculations. Around the Fermi level the two results for Γ are similar; i.e., both results show the well-known, correct quadratic energy dependence. At a distance ω_p and further away from the Fermi level, however, the differences are striking. The sharp structures produced by the Lindhard screening function in the partially self-consistent case have almost vanished in the fully self-consistent result. Instead, there is broad and rather structureless spectral weight of considerable magnitude at large energies—positive and negative. Such weight is of little consequence to the spectral function A as seen in Eq. (22). Needless to say, the loss of the sharp structure at the plasma frequency in the self-consistent calculation is the reason for the disappearance of the plasmon satellite in that calculation.

In Fig. 6 we do the same comparison for the real part of the correlation contribution to the self-energy, i.e., the last term in Eq. (11). Again, we notice the loss of sharp structure and the larger magnitudes far away from the Fermi level in the fully self-consistent calculation. At the Fermi level, however, the two curves have almost the same value, indicating that the correlation contribution to the chemical potential is similar in the two cases. As a matter of fact, the same is approximately true also for the Hartree-Fock part of the self-energies and the chemical potentials of the GW and the GW_0 calculations are thus rather similar.

We also notice the smaller slope of the GW result at the

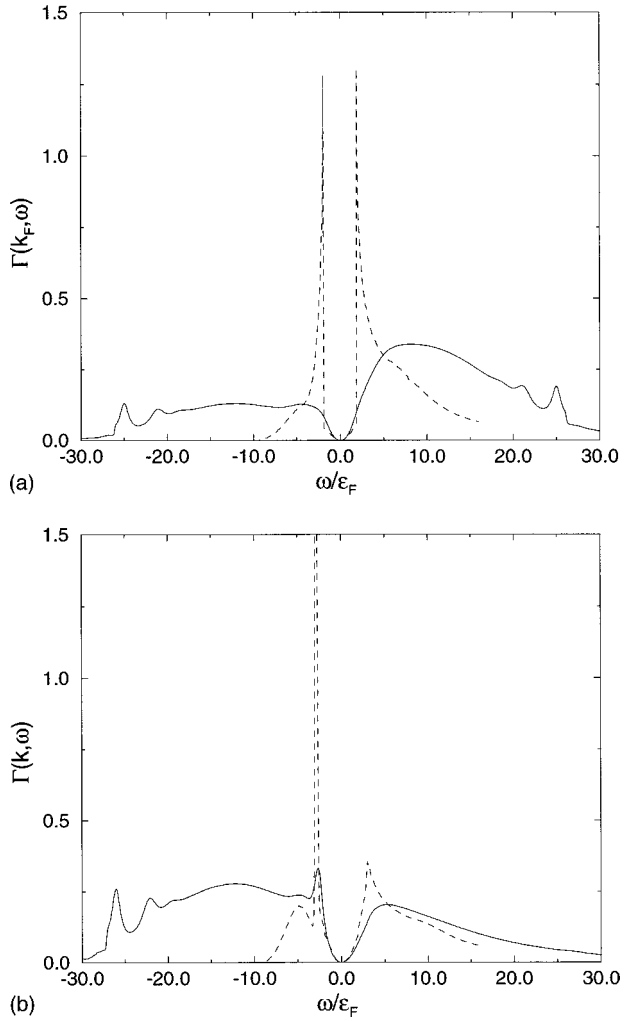

FIG. 5. The spectral function $\Gamma(\mathbf{k},\omega)$ of the self-energy at full self-consistency (solid) is compared to corresponding quantity (dashed) from a partially self-consistent GW_0 calculation (Ref. 1) at the Fermi surface [$k=k_F$ (a)] and at the bottom of the band [$k=0$ (b)]. Notice the same parabolic shapes around the Fermi level but also the transformation of the sharp structures in the GW_0 case to the very wide structures obtained in the full GW calculation. Here, $r_s=4$.

Fermi level. This slope determines the so-called quasiparticle renormalization factor Z_k according to

$$Z_k=\left[1-\frac{\partial\ \mathrm{Re}\Sigma(\mathbf{k},\omega)}{\partial\omega}\right]^{-1}_{\omega=E_{\mathbf{k}}}, \tag{24}$$

where the quasiparticle energy $E_{\mathbf{k}}$, as usual, is defined to be the solution to the Dyson-like equation that determines the real part of the poles of the Green's function [Eq. (19)]. Thus,

$$E_{\mathbf{k}}=\epsilon_{\mathbf{k}}+\Sigma(\mathbf{k},E_{\mathbf{k}})\,. \tag{25}$$

The imaginary part,

$$\Gamma_{\mathbf{k}}=\frac{1}{\pi}\,|\mathrm{Im}\Sigma(\mathbf{k},E_{\mathbf{k}})|=\Gamma(\mathbf{k},E_{\mathbf{k}}) \tag{26}$$

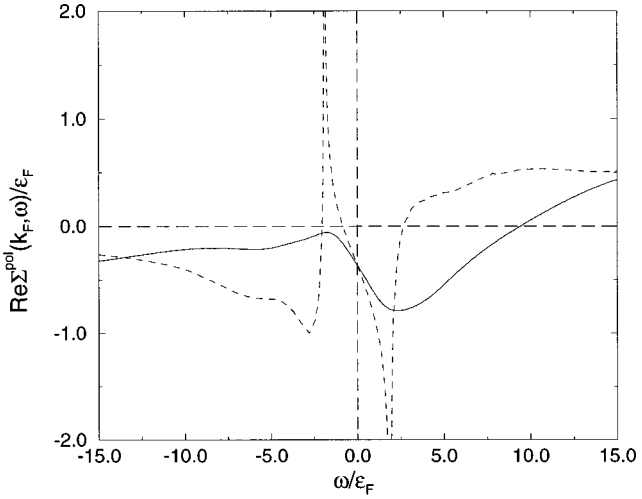

FIG. 6. The real part of the correlation contribution, $\mathrm{Re}\Sigma(k_F,\omega)-\Sigma_{\mathrm{HF}}(k_F)$, to the self-energy at full self-consistency (solid) is compared to the corresponding quantity (dashed) from a GW_0 calculation (Ref. 1). Notice again the loss of sharp structure in the full GW result and the smaller slope at the Fermi level. Here, $r_s=4$.

of the self-energy at $E_{\mathbf{k}}$ is a measure of the width of the quasiparticle peak. As we shall see below, the quasiparticle peak below the Fermi level is so sharp that $E_{\mathbf{k}}$ really corresponds to the maximum of the quasiparticle peak in the spectral function $A(\omega)$.

In Fig. 7 we display the quantity Z_k, which thus gives the weight of the quasiparticle relative to all satellite structures. We see that the weight of the quasiparticle has increased with respect not only to a normal G_0W_0 calculation but also compared to a partially self-consistent GW_0 calculation. This is also evident from Table I, giving the renormalization factor $Z_F=Z_{k_F}$ at the Fermi surface at two different densities. In Fig. 7, we also see the loss of structure in the self-consistent

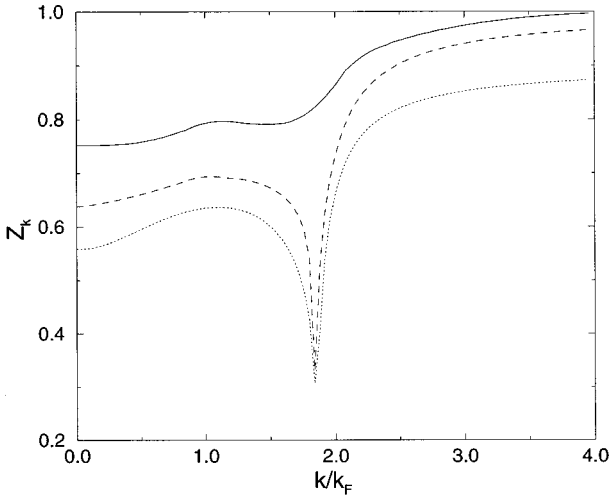

FIG. 7. The quasiparticle renormalization factor Z_k [Eq. (24)] of the full GW calculation (solid) is compared to the corresponding quantity from a GW_0 calculation (dashed) and from a G_0W_0 calculation (dotted). Notice the increase of quasiparticle weight with a larger degree of self-consistency. Notice also the loss of structure due to plasmon decay in the full GW result. Here, $r_s=4$.

TABLE I. The quasiparticle renormalization factor Z_F from three different calculations: G_0W_0, GW_0, and GW. Results are given for the densities corresponding to $r_s=2$ and 4.

	Z_F		
r_s	G_0W_0	GW_0	GW
2	0.764	0.804	0.846
4	0.645	0.702	0.793

result. This is, of course, again associated with the replacement of the plasmon satellite with a broad and featureless structure.

The resulting self-consistent one-electron spectral functions $A(\mathbf{k},\omega)$ are displayed in the Figs. 8(a) and 8(b) and compared with those of our previous partially self-consistent GW_0 results. The most notable differences are the disappearance of the plasmon satellite and the larger bandwidth seen in the fully self-consistent results.

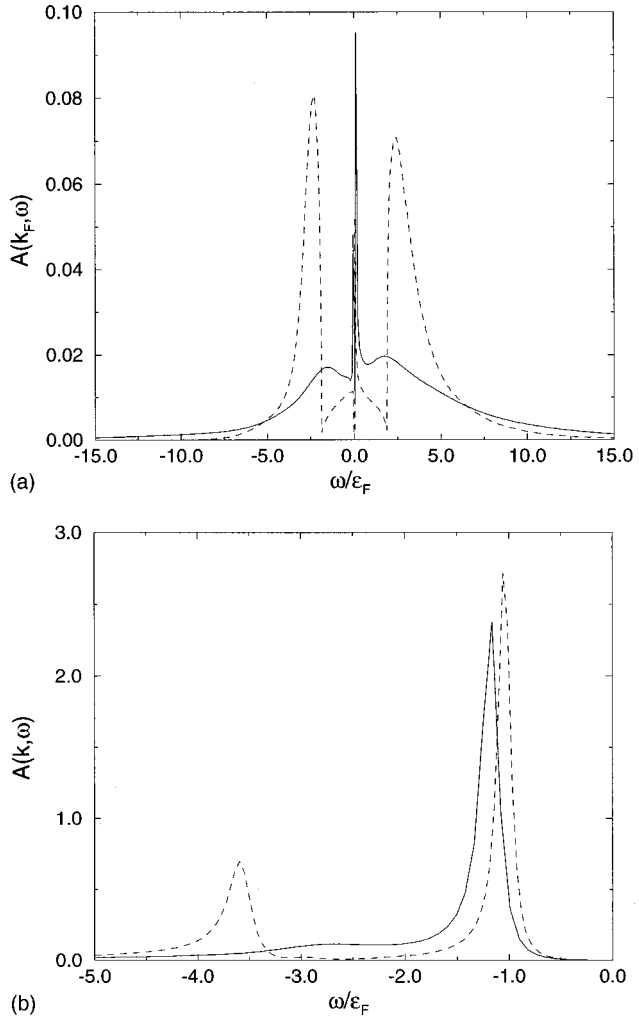

FIG. 8. The resulting one-electron spectral function $A(\mathbf{k},\omega)$ from the full GW calculation (solid) is compared to the corresponding quantity from the partially self-consistent GW_0 calculation (Ref. 1) (dashed). In (a), the comparison at made at the Fermi surface, $|\mathbf{k}|=k_F$, and (b) shows the same comparison at the bottom of the band, $|\mathbf{k}|=0$. In (a) the height of the quasiparticle peak cannot be accommodated within the figure. Notice in both figures the loss of plasmon satellites in the full GW calculation. Notice also in (b) the markedly larger bandwidth in the GW case. As before, $r_s=4$.

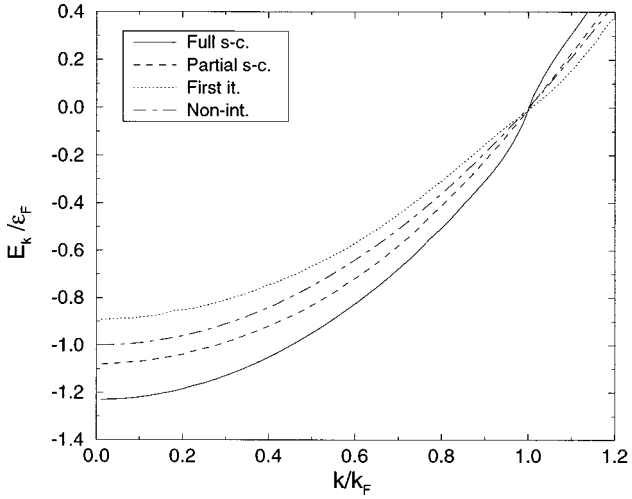

FIG. 9. The dispersion $E_{\mathbf{k}}$ of the quasiparticles is compared to a free-electron parabola (small dots) for the cases G_0W_0 (dotted), GW_0 (dash-dotted), and GW (solid). Only the simplest G_0W_0 shows the desired band narrowing and the GW result is the worst. Here, $r_s=4$.

Experiments show that the bandwidths of simple metals are of the order of 10% more narrow as compared to the results of band-structure calculations employing a local and energy-independent potential, i.e., like that of the LDA. Thus, there was originally some hope that a GW calculation would cure this problem—and a non-self-consistent G_0W_0 calculation for the electron gas does indeed produce a more narrow bandwidth although not narrow enough. One of the conceptual deficiencies of a non-self-consistent procedure is, of course, that the final results depend on the rather arbitrary choice of zeroth-order Green's function (G_0). Therefore, it was disappointing to find that a partially self-consistent GW_0 calculation gives a bandwidth larger than that of noninteracting electrons.[1] This unfortunate circumstance led to the hope that a fully self-consistent calculation (GW) might rectify the situation. It is therefore perhaps even more disappointing to see that the fully self-consistent procedure actually makes the problem even worse.

As seen in Fig. 9, at a density corresponding to that of sodium ($r_s=4$), the bandwidth is more than 20% larger than that of noninteracting electrons and some 15% larger than that predicted by a partially self-consistent calculation. This result is somewhat difficult to understand. In a G_0W_0 calculation, the Hartree-Fock part of the self-energy is that of noninteracting electrons, causing almost a doubling of the bandwidth. This effect is more than compensated for by the dynamic part of the self-energy, which is calculated from G_0 and W_0-v. In a GW_0 calculation, the dominating quasiparticle contribution to the dynamic part of the self-energy is scaled down by the renormalization factor Z_k and that part of Σ is no longer able to cancel the large Hartree-Fock contribution. The fact that the latter is reduced somewhat by the use of interacting occupation numbers in Eq. (13) improves the situation but does not cure the problem. In the fully self-consistent case, the magnitude of the quasiparticle contribution to the Green's function and, therefore, to the dynamical

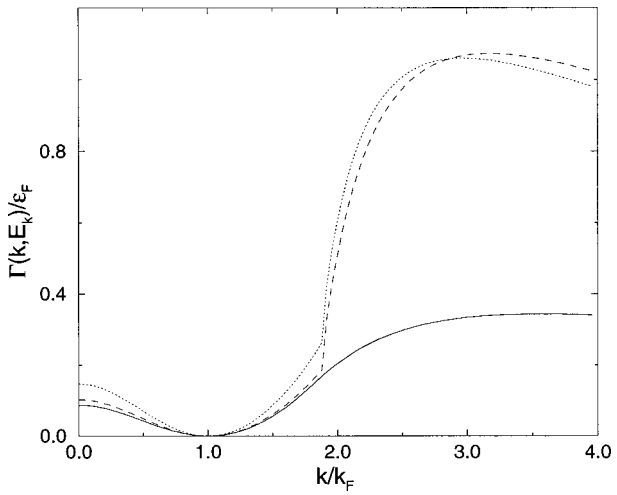

FIG. 10. The inverse lifetime of the quasiparticles Γ_k [Eq. (26)] at full self-consistency (*GW*, solid) is compared to the corresponding quantities from GW_0 (dashed) and G_0W_0 (dotted) calculations. The two latter results show a region with much shorter lifetimes caused by inelastic collision with plasmons. This physically correct feature is absent from the full *GW* result. As before, $r_s=4$.

part of the self-energy is larger and one would expect a better cancellation of the Hartree-Fock enlargement of the bandwidth. Unfortunately, in this case, the effect on the occupation numbers and, therefore, on the Hartree-Fock contribution appears to be even larger, resulting in an increased bandwidth. Consequently, a smaller bandwidth in better harmony with the experimental facts can only be achieved trough vertex corrections. That vertex corrections indeed are capable of giving the desired band narrowing was recently demonstrated by Mahan and Sernelius[12] and by Shirley.[13]

Finally, we say a few words about the momentum distribution function $n_{\mathbf{k}}$ [Eq. (14)] and the lifetime $1/\Gamma_{\mathbf{k}}$ of the quasiparticles. The latter is displayed in Fig. 10 and the former in Fig. 11. With no well-defined plasmons, as in the *GW* case, there is no possibility of energy loss through inelastic collisions with plasmons. Consequently, in the *GW* case, the quasiparticles peaks are rather sharp at all momenta. The G_0W_0 and the GW_0 results clearly show a more physical behavior with a large decrease of the lifetime at the onset of decay through plasmon production.

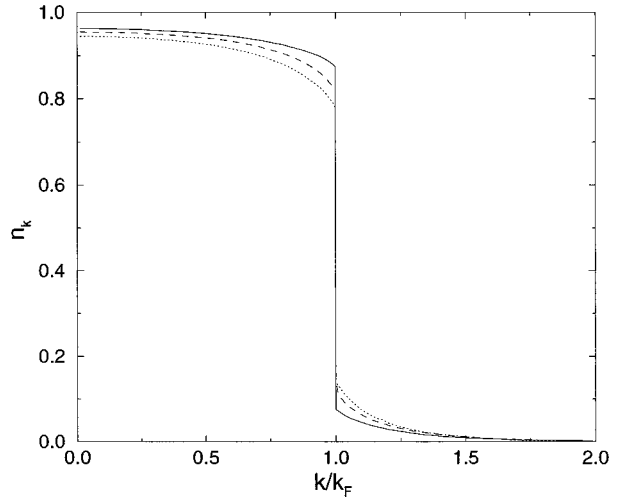

FIG. 11. The momentum distribution function $n_{\mathbf{k}}$ [Eq. (14)] of the full *GW* calculation is compared to $n_{\mathbf{k}}$ from GW_0 (dashed) and G_0W_0 (dotted) calculations. Again, we notice the increase in quasiparticle weight, represented by the size of the discontinuity, as we increase the degree of self-consistency. As before, $r_s=4$.

The discontinuities in the momentum distribution functions in Fig. 11 are equal to the weight of the quasiparticles at the Fermi surface. Consequently, these distribution curves again show an increase in quasiparticle weight with a larger degree of self-consistency. Note that the area under each curve gives the electron density of the system. As will be discussed in Sec. IV, the *GW* calculation reproduces the free-electron result and this can also be shown to be true in the case of the GW_0 calculation.[14] As a matter of principle, the G_0W_0 procedure is known to violate[15] this exact relation but, as was discussed in Ref. 1, it is numerically true beyond the accuracy of our numerical procedures.

IV. THE TOTAL ENERGY

Perhaps one of the most interesting features of the present fully self-consistent way of doing *GW* calculations is the fact that it represents a conserving approximation in the sense defined by Kadanoff and Baym.[6,7] Such approximations preserve particle number, energy, momentum, and angular momentum under external perturbations. They also obey a number of additional consistency requirements that one might desire from a reasonable theory.[7] As an example we might consider for a moment the total energy, which is a well-defined quantity in an exact theory. Within MBPT, however, this uniqueness is easily lost. We often try to find approximations to the one-electron Green's function *G* from which the total energy can be calculated. The road towards the approximate *G* most often passes the two-particle Green's function from which the total energy also can be calculated. In this way, we can obtain two total energies, which, in most approximations, do not agree. We are thus faced with the problem of deciding which is the most accurate total energy. One of the virtues of conserving approximations is that different ways of calculating the total energy give identical results. For instance, having some approximation $G(k)$ to the Green's function [k here represents the four-vector $(\mathbf{k},\omega)$], we easily obtain the expectation value *T* of the operator representing the kinetic energy,

$$T=-2i\Omega\sum_{\mathbf{k}}\int\frac{d\omega}{2\pi}\,\epsilon_{\mathbf{k}}G(k). \tag{27}$$

Here, Ω is the total volume of the gas. Having also some approximation to the self-energy $\Sigma(k)$, we can calculate the expectation value *U* of the Coulomb interaction,

$$U=-i\Omega\sum_{\mathbf{k}}\int\frac{d\omega}{2\pi}\,G(k)\Sigma(k). \tag{28}$$

Adding the pieces *T* and *U* to form the total energy ($E=T+U$), we obtain the Galitskii-Migdal[10] formula after a few manipulations that require that *G* and Σ are coupled through Dyson's equation, Eq. (19):

TABLE II. The total energies from the full GW calculation at different densities (r_s) are compared to the results from the Monte Carlo calculations by Ceperley and Alder (Ref. 9).

	E/N (Ry)	
r_s	GW	MC
2	0.005	0.004
4	−0.156	−0.155

$$E=\Omega\sum_{\mathbf{k}}\int(\omega+\epsilon_{\mathbf{k}})A(\mathbf{k},\omega)d\omega. \tag{29}$$

The total energy can also be obtained through the usual Hellman-Feynman trick, i.e., by integrating the interaction energy $U(\lambda)$ regarded as a function of the strength λ of the Coulomb interaction with respect to that same λ,

$$E=E_0+\int_0^1\frac{d\lambda}{\lambda}U(\lambda). \tag{30}$$

Here, E_0 is the total energy of the same number of noninteracting electrons. The beauty of a conserving approximation is that these two ways of obtaining the total energy [Eqs. (29) and (30)] produce the same result—something that we have verified to within our numerical accuracy (approximately two significant figures).

A third way of obtaining the total energy is through the chemical potential μ defined by Eq. (21). An alternative and perhaps more fundamental definition of μ is the exact relation

$$\mu=\frac{\partial E}{\partial N}, \tag{31}$$

where N is the total number of electrons in the system. Also this relation is obeyed to within two significant figures in our numerical calculations, thus adding to the credibility of the present work.

In Table II we compare our total energies to those obtained from the Green's-function Monte Carlo calculations by Ceperley and Alder.[9] The latter energies are often considered to be very close to the exact ones. It is striking to see how close our total energies are to the Monte-Carlo results.

We did not anticipate such a close agreement. We find it hard to believe that any theory, like the present one, which does not include some reasonable approximation to the interparticle interaction at close range, could ever yield an accurate energy. Such a theory must certainly involve higher orders of the screened interaction W. As far as the total energy is concerned, it appears as if errors due to an improper treatment of the Coulomb interaction at short distances are compensated by errors from a less than adequate treatment of the screening effects at long distances. As will be described in a future publication,[14] the total energy of the self-consistent GWA is a variational quantity that might explain the accurate energies obtained in the present work.

We end this section with a remark about the particle number. We note that the only input parameter of the present theory is the Fermi momentum k_F in Eq. (21). There is thus no immediate reason to believe that the particle density n calculated from the final self-consistent Green's function through the formula

$$n=2\sum_{\mathbf{k}}\int_{-\infty}^{0}A(\mathbf{k},\omega)d\omega \tag{32}$$

should yield $n=k_F^3/(3\pi^2)$ as for free electrons. As proven by Baym,[7] however, this is indeed true in conserving approximations and we have here verified the same thing numerically to a high degree of accuracy.

V. CONCLUSIONS

The purpose of the present work and our previous work on self-consistent GW calculations is to elucidate the consequences of different ways of performing GW calculations in general. We were, however, also motivated by the desire to see a truly conserving approximation at work in a real many-body system. In this section we will try to summarize our most important findings.

During our investigations, we have found that allowing for a broadening of quasiparticles usually has a negligible effect on the final one-electron Green's function. We could, however, imagine cases when the quasiparticles are less well defined broad structures. In such cases, the effect of quasiparticle decay could become important.

In the present investigations on the translationally invariant electron gas, there is no substantial shift of quasiparticle energies. Thus, in simple metals we would anticipate only small changes in the calculated one-electron spectrum from using a starting Green's function with reasonable shifts (10%) in the one-electron energies.

In going from a G_0W_0 to a GW_0 calculation and further to a full GW calculation in the notation of this work, we find a steady deterioration in the description of the bandwidth. A fully self-consistent GW calculation in sodium could easily yield a bandwidth one-third larger than the experimental width. This result clearly indicates the necessity of vertex corrections.

The description of satellite structure is not satisfactory in any kind of GW approach. The description is reasonable at the levels of G_0W_0 and GW_0 but it breaks down at the level of full self-consistency (GW).

In going from a G_0W_0 to a GW_0 calculation and further to a full GW calculation, the total energy as obtained from the Galitskii-Migdal expression becomes systematically better and the energy is very accurate in the fully self-consistent theory. It should be noted, however, that there are variational and other ways of obtaining total energies from less sophisticated GW calculations. This will be discussed in future publications

Calculations for real solids based on MBPT at the level of G_0W_0 is quickly becoming routine. Partially self-consistent calculations for real solids are within reach, particularly if there are not too many atoms per unit cell. We still have to find a few proper shortcuts before fully self-consistent calculations become feasible. The present investigation, however, clearly demonstrates the futility of such huge efforts. In fully self-consistent calculations a lot of computing power will be spent on producing quite inferior results. Instead, the present

calculations demonstrate that we should now direct our efforts towards constructing vertex corrections that properly account for the cancellation between self-energy insertions and particle-hole and hole-hole interactions.

ACKNOWLEDGMENTS

We would like to thank Dr. C.-O. Almbladh and Mikael Hindgren for useful discussions during the course of the present work.

[1] U. von Barth and B. Holm, Phys. Rev. B **54**, 8411 (1996); Phys. Rev. B **55**, 10 120(E) (1997).
[2] L. Hedin, Phys. Rev. **139**, A796 (1965).
[3] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[4] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[5] F. Aryasetiawan and O. Gunnarsson, Phys. Rev. Lett. **74**, 3221 (1995).
[6] G. Baym and L. P. Kadanoff, Phys. Rev. **124**, 287 (1961).
[7] G. Baym, Phys. Rev. **127**, 1391 (1962).
[8] H. J. de Groot, P. A. Bobbert, and W. van Haeringen, Phys. Rev. B **52**, 11 000 (1995).
[9] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
[10] V. Galitskii and A Migdal, Zh. Eksp. Teor. Fiz. **34**, 139 (1958) [Sov. Phys. JETP **7**, 96 (1958)].
[11] J. Lindhard, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **28**, 8 (1954).
[12] G. D. Mahan and B. E. Sernelius, Phys. Rev. Lett. **62**, 2718 (1989).
[13] E. L. Shirley, Phys. Rev. B **54**, 7758 (1996).
[14] B. Holm and U. von Barth (unpublished).
[15] A. Schindlmayr, Phys. Rev. B **56**, 3528 (1997).