Correlation effects in solids from first principles

Ferdi Aryasetiawan
Joint Research Center for Atom Technology-Angstrom Technology Partnership, 1-1-4 Higashi, Tsukuba, Ibaraki 305-0046, Japan

Abstract

First principles calculations of bandstructures of crystals are usually based on one-particle theories where the electrons are assumed to move in some effective potential. The most commonly used method is based on density functional theory within the local density approximation (LDA). There is, however, no clear justification for interpreting the one-particle eigenvalues as the bandstructure. Indeed, the LDA failure to reproduce the experimental bandstructure is not uncommon. The most famous example is the bandgap problem in semiconductors and insulators where the LDA generally underestimates the gaps.

A rigorous approach for calculating bandstructures or quasiparticle energies is provided by the Green function method. The main ingredient is the self-energy operator which acts like an effective potential but unlike in the LDA, it is nonlocal and energy dependent. The self-energy contains the effects of exchange and correlations. An approximation to the self-energy which has proven fruitful in a wide range of materials is the so-called GW approximation (GWA). This approximation has successfully cured the LDA problems and has produced bandstructures with a rather high accuracy. For example, bandgaps in s-p semiconductors and insulators can be obtained typically to within 0.1-0.2 eV of the experimental values.

Despite its success, the GWA has some problems. One of the most serious problems is its inadequacy to describe satellite structures in photoemission spectra. For example, multiple plasmon satellites observed in alkalis cannot be obtained by the GWA. Recently, a theory based on the cumulant expansion was proposed and shown to remedy this problem. Apart from plasmon satellites which are due to long-range correlations, there are also satellite structures arising from short-range correlations. This type of satellite cannot be described by the cumulant expansion. A t-matrix approach was proposed to account for this.

Although traditionally the Green function method is used to calculate excitation spectra, groundstate energies can also be obtained from the Green function. Recent works on the electron gas have shown promising results and some approaches for calculating total energies will be discussed.

1 Introduction

The electronic structures of molecules and solids are now routinely calculated using the Kohn-Sham (KS) density functional theory (DFT) [52, 61] within the local density approximation (LDA). In this theory, one solves a single-particle Hamiltonian with an effective local potential containing the Hartree potential V_H and an exchange-correlation potential V_{xc}:

$$\left[-\frac{1}{2}\nabla^2 + V_H + V_{xc}\right]\psi_i = \varepsilon_i\psi_i \tag{1}$$

The resulting one-particle eigenvalues ε_i are often interpreted as one-particle excitation energies or in the case of a crystal, as quasiparticle energies (bandstructures) measured in photoemission experiments. There is no clear theoretical justification for this, except for the highest occupied states. For *sp* systems, it is found that the bandstructure is often in rather good agreement with experiment. It is remarkable that the LDA, being a simple approximation, can go a long way in accounting for the electronic structures of many materials. However, there are serious discrepancies. The most well-known of these is the band-gap problem in semiconductors (Si, GaAs, Ge, etc.) where the LDA bandgpas are systematically underestimated. This is illustrated in figure 1. The discrepancies become worse in strongly correlated $3d$ and $4f$ systems. In the so called Mott-Hubbard insulators of transition metal oxides, the LDA band gap is much too small. In some cases, the LDA even gives qualitatively wrong results. For example, the Mott-Hubbard insulator CoO and the undoped parent compound of the high T_c material La_2CuO_4 are predicted to be metals whereas experimentally they are insulators. Apart from these problems, LDA sometimes overestimates the valence bandwidth, for example in Na and Ni.

A proper way of calculating single-particle excitation energies or quasiparticle energies [64, 65] is provided by the Green function theory [38, 39]. It can be shown that the quasiparticle energies E_i can be obtained from the quasiparticle equation:

$$[-\frac{1}{2}\nabla^2(\mathbf{r}) + V^H(\mathbf{r})]\Psi_i(\mathbf{r}) + \int d^3r'\ \Sigma(\mathbf{r},\mathbf{r}';E_i)\Psi_i(\mathbf{r}') = E_i\Psi_i(\mathbf{r}) \tag{2}$$

The non-local and energy dependent potential Σ, or the self-energy, contains the effects of exchange and correlations. It is in general complex with the imaginary part describing the damping of the quasiparticle. It is then clear that the different single-particle theories amount to approximating the self-energy operator Σ. Thus we can think of the V_{xc} in DFT as a local and energy independent approximation to the self-energy that gives the correct ground state density. Approximating Σ by the exchange operator

$$\Sigma(\mathbf{r},\mathbf{r}') = -v(\mathbf{r}-\mathbf{r}')\sum_i^{\text{occ}}\psi_i^*(\mathbf{r})\psi_i(\mathbf{r}') \tag{3}$$

results in the Hartree-Fock approximation.

Improving the LDA means finding corrections to the self-energy beyond V_{xc}^{LDA}. There are a number of attempts, amomg these are Generalized Gradient Approximations (GGA) [67, 16, 17, 18, 101, 95, 88], Self-Interaction Correction (SIC) [26, 71, 108, 87, 100, 102, 7], LDA+U [4, 5, 68, 93, 94, 6], Optimized Effective Potential (OEP) Method (Exact Exchange + Correlations) [103, 62, 63, 22, 23], and more recently Dynamical Mean-Field theory (DMFT) (for a review and references, see [40]). This note describes the *GW* approximation (GWA) [46] which has been found to be successful in accounting for quasiparticle energies for a wide range of systems from atoms to solids. The GWA may be thought of as a generalization of the Hartree-Fock approximation (HFA) but with a dynamically screeneed interaction. Recent development in going beyond the GWA is also presented. These include the cumulant expansion method and

the t-matrix approach. In addition, applications of the Green function theory to total energy calculations and the issue of self-consistency will be discussed.

This note also describes practical implementations of the GWA. Conventional methods for calculating the self-energy use plane waves as basis functions which are suitable for sp systems. However, applications to 3d and 4f systems using plane-wave basis are unsuitable due to a large number of plane waves needed to describe the localized 3d or 4f states. A method for treating systems with localized states is described. The method is based on the linear muffin-tin orbital (LMTO) approach [3] and basis functions needed to describe the screened Coulomb interaction and the self-energy are products of the LMTO's.

2 Theory

2.1 Photoemission experiments measure quasiparticle energies

In a photoemission experiment, photons are used to excite electrons out of a crystal leaving holes behind and giving information about the occupied states. In an inverse photoemission experiment (BIS), electrons are sent into the crystal to probe unoccupied states. Photoemission processes are usually interpreted in terms of the three-step model [34, 81, 89, 20] consisting of

- optical excitation of an electron;
- its transport through the solid with the possibility of inelastic scattering with other electrons;
- the escape through the surface into the vacuum.

To illustrate the basic principle, here we consider a simple case of normal emission where photoelectrons are emitted normal to the sample surface. The more general case is considerably more complicated and for a more detailed discussion, we refer to Ref. [92].

Using conservation of energy

$$E_0(N) + \omega = E_f(N-1, \mathbf{k}) + k^2/2 \tag{4}$$

where $E_0(N)$ is the groundstate energy of the solid, ω is the photon energy, $E_f(N-1, \mathbf{k})$ is an excited state of the (N-1) electrons, and $k^2/2$ is the kinetic energy of the outgoing photoelectron. For a given crystal wave vector $\mathbf{k}$, the photoelectron spectrum usually shows peaks at well defined energies $\omega - k^2/2 = E_f(N-1, \mathbf{k}) - E_0(N)$ which may be interpreted as the quasiparticle excitations and they can be shown to correspond to peaks in the imaginary part of the Green function. When the positions of the peaks are plotted as a function of $\mathbf{k}$, we obtain what is called the bandstructure.

¿From the Fermi Golden Rule, the probability per unit time of emitting an electron of momentum $\mathbf{k}$ is

$$I(\mathbf{k}, \omega) = 2\pi \sum_f |\langle N-1, f, \mathbf{k}|\mathbf{A}\cdot\mathbf{p}|N\rangle|^2 \delta(\omega - k^2/2 - \varepsilon_f) \tag{5}$$

where $\mathbf{A}$ is the electromagnetic vector field, $\varepsilon_f = E_f(N-1) - E_0(N)$. Since the photoelectron usually has a high energy we can approximate ("sudden approximation")

$$|N-1, f, \mathbf{k}\rangle \approx c^\dagger_{\mathbf{k}}|N-1, f\rangle, \quad c_{\mathbf{k}}|N\rangle \approx 0 \tag{6}$$

In second quantized form $\mathbf{A}\cdot\mathbf{p} = \sum_{lm} \Delta_{lm} c^\dagger_l c_m$. Using $c_{\mathbf{k}} c^\dagger_l + c^\dagger_l c_{\mathbf{k}} = \delta_{\mathbf{k}l}$ we then have

$$I(\mathbf{k}, \omega) = 2\pi \sum_f |\langle N-1, f| \sum_m \Delta_{\mathbf{k}m} c_m |N\rangle|^2 \delta(\omega - k^2/2 - \varepsilon_f) \tag{7}$$

which can be rewritten

$$I(\mathbf{k},\omega) = 2\pi \sum_{mm'} \Delta_{\mathbf{k}m} A_{mm'}(\omega - k^2/2)\Delta_{\mathbf{k}m'} \tag{8}$$

where

$$A_{mm'}(\omega) = \sum_f \langle N-1, f|c_m|N\rangle\langle N|c^\dagger_{m'}|N-1,f\rangle\delta(\omega-\epsilon_f) \tag{9}$$

The quantity A will be shown later to be proportional to the imaginary part of the Green function, assuming the matrix elements $\Delta_{\mathbf{k}m}$ are constant.

2.2 The Green function and the self-energy

To describe the photoelectron spectra measured in photemission and inverse photoemission experiments we need a quantity that the propagation of a hole or an added electron. A suitable quantity is the one-particle Green function Refs. [36, 78, 59] which is defined as:

$$\begin{aligned} iG(x,x') &= \langle N|T[\hat{\psi}(x)\,\hat{\psi}^\dagger(x')]|N\rangle \\ &= \begin{cases} \langle N|\hat{\psi}(x)\,\hat{\psi}^\dagger(x')|N\rangle & for \quad t > t' \quad \text{(electron)} \\ -\langle N|\hat{\psi}^\dagger(x')\,\hat{\psi}(x)|N\rangle & for \quad t < t' \quad \text{(hole)} \end{cases} \end{aligned} \tag{10}$$

We have used a notation $x \equiv (\mathbf{r}, t)$. T is the time ordering operator, and $|N\rangle$ is the groundstate of N electrons. Thus, for $t > t'$, the Green function is the probability amplitude that an electron added at x' will propagate to x, and for $t' > t$, the probability amplitude that a hole created at x will propagate to x'. A possible spin flip can be incorporated in the definition of G.

¿From the Green function we can obtain:

- The expectation value of any single-particle operator in the ground state.
- The ground state energy
- The excitation spectrum

The field operators are defined in the Heisenberg representation and they satisfy the equation of motion

$$i\frac{\partial}{\partial t}\hat{\psi}(x) = [\hat{\psi}(x), \hat{H}] \tag{11}$$

where the Hamiltonian is given by

$$\begin{aligned} \hat{H} &= \int d^3r\ \hat{\psi}^\dagger(x) h_0(\mathbf{r})\hat{\psi}(x) \\ &+ \frac{1}{2}\int d^3r d^3r'\ \hat{\psi}^\dagger(x)\hat{\psi}^\dagger(x') v(\mathbf{r}-\mathbf{r}')\hat{\psi}(x')\hat{\psi}(x) \end{aligned} \tag{12}$$

h_0 is the kinetic energy operator plus any single-particle operator such as an external field. Evaluating the above commutator, the equation of motion for the Green function is then

$$\begin{aligned} &\left[i\frac{\partial}{\partial t} - h_0(x)\right] G(x,x') \\ &+ i\int d^3r_1\ v(\mathbf{r}-\mathbf{r}_1)\ \langle N|T[\hat{\psi}^\dagger(\mathbf{r}_1,t)\hat{\psi}(\mathbf{r}_1,t)\hat{\psi}(\mathbf{r},t)\hat{\psi}^\dagger(\mathbf{r}',t')]|N\rangle \\ &= \delta(x-x') \end{aligned} \tag{13}$$

The quantity

$$\langle N|T[\hat{\psi}^\dagger(\mathbf{r}_1,t)\hat{\psi}(\mathbf{r}_1,t)\hat{\psi}(\mathbf{r},t)\hat{\psi}^\dagger(\mathbf{r}',t')]|N\rangle$$

is a special case of the two-particle Green function defined by

$$G_2(1,2,3,4) = (i)^2\langle N|T[\hat{\psi}(1)\hat{\psi}(3)\hat{\psi}^\dagger(4)\hat{\psi}^\dagger(2)|N\rangle \tag{14}$$

describing the propagation of two particles from $2,4$ to $1,3$ where we use a short-hand notation $1 \equiv x_1$. The two-particle Green function, in turn, satisfies an equation of motion involving a three-particle Green function, forming a hierarchy of equations. To break the hierarchy, the two-particle Green function is expressed in terms of the one-particle Green functions. Formally, we introduce the mass operator $M(x,x')$ such that

$$\int dx_1 M(x,x_1)G(x_1,x') = \\ -i\int d^3r_1\; v(\mathbf{r}-\mathbf{r}_1)\;\langle N|T[\hat{\psi}^\dagger(\mathbf{r}_1,t)\hat{\psi}(\mathbf{r}_1,t)\hat{\psi}(\mathbf{r},t)\hat{\psi}^\dagger(\mathbf{r}',t')]|N\rangle \tag{15}$$

so that

$$\left[i\frac{\partial}{\partial t} - h_0(x)\right]G(x,x') - \int dx_1 M(x,x_1)G(x_1,x') = \delta(x-x') \tag{16}$$

As will be shown later, the mass operator contains the Hartree potential and the self-energy is defined to be the mass operator without the Hartree potential, thus containing the effects of exchange and correlation only:

$$\Sigma = M - V_H \tag{17}$$

Fourier transforming Eq. (16) and using the definition of the self-energy we get

$$[\omega - h_0(\mathbf{r}) - V_H(\mathbf{r})]\,G(\mathbf{r},\mathbf{r}',\omega) - \int d\mathbf{r}_1\Sigma(\mathbf{r},\mathbf{r}_1,\omega)G(\mathbf{r}_1,\mathbf{r}',\omega) = \delta(\mathbf{r}-\mathbf{r}') \tag{18}$$

We define a noninteracting Green function G_0 as a solution to Eq. (18) with $\Sigma = 0$. Noting that $\omega - h_0 - V_H = G_0^{-1}$, the exact Green function G satisfies the Dyson equation

$$\begin{aligned} G &= G_0 + G_0\Sigma G \\ &= \frac{1}{G_0^{-1} - \Sigma} \end{aligned} \tag{19}$$

3 Spectral representation of the Green function

The spectral representation of the Green fucntion is obtained from the definition in Eq. (10) by noting that

$$\hat{\psi}(\mathbf{r},t) = \exp(i\hat{H}t)\;\hat{\psi}(\mathbf{r},0)\;\exp(-i\hat{H}t) \tag{20}$$

Inserting a complete set of $N\pm1$-particle eigenstates of $\hat{H}$ and performing the Fourier transform give

$$G(\mathbf{r},\mathbf{r}',\omega) = \sum_i \frac{h_i(\mathbf{r})h_i^*(\mathbf{r}')}{\omega + E_i(N-1) - E_0(N) - i\delta} + \sum_j \frac{p_j(\mathbf{r})p_j^*(\mathbf{r}')}{\omega - E_j(N+1) + E_0(N) + i\delta} \tag{21}$$

where

$$h_i(\mathbf{r}) = \langle N-1,i|\hat{\psi}(\mathbf{r},0)|N\rangle \tag{22}$$

$$p_j(\mathbf{r}) = \langle N|\hat{\psi}(\mathbf{r},0)|N+1,j\rangle \tag{23}$$

$|N\pm1,i\rangle$ is the ith eigenstate of the $N\pm1$ electrons with eigen energy $E_i(N\pm1)$ and $E_0(N\pm1)$ is the groundstate energy of the $N\pm1$ electrons.

We can write the Green function in a more compact form

$$G(\mathbf{r},\mathbf{r}',\omega)=\int_{-\infty}^{\mu}d\omega'\,\frac{A(\mathbf{r},\mathbf{r}',\omega')}{\omega-\omega'-i\delta}+\int_{\mu}^{\infty}d\omega'\,\frac{A(\mathbf{r},\mathbf{r}',\omega')}{\omega-\omega'+i\delta} \tag{24}$$

The spectral function A (density of states) is proportional to the imaginary part of G

$$A(\mathbf{r},\mathbf{r}',\omega)=-\frac{1}{\pi}\mathrm{Im}\,G(\mathbf{r},\mathbf{r}',\omega)\mathrm{sgn}(\omega-\mu) \tag{25}$$

and it is given by

$$\begin{aligned}A(\mathbf{r},\mathbf{r}',\omega) &= \sum_i f_i(\mathbf{r})f_i^*(\mathbf{r}')\delta[\omega-\mu+e(N-1,i)]\\ &+ \sum_i g_i(\mathbf{r})g_i^*(\mathbf{r}')\delta[\omega-\mu-e(N+1,i)]\end{aligned} \tag{26}$$

$e_i(N\pm1)$ is the excitation energy of the $N\pm1$ electrons:

$$e_i(N\pm1)=E_i(N\pm1)-E_0(N\pm1) \tag{27}$$

which is positive and the quantity μ is the chemical potential

$$\mu=E(N+1)-E(N)=E(N)-E(N-1)+O(1/N) \tag{28}$$

The poles of the Green function are therefore the exact excitation energies of the $N\pm1$ electrons.

As an example, for a noninteracting case the many-body states become single-Slater determinants. The quantities h_i and p_i are replaced by the one-particle wavefunctions satisfying

$$H_0\psi_i=\varepsilon_i\psi_i \tag{29}$$

and the excitation energies are replaced by ε_i. The noninteracting Green function is then given by

$$G_0(\mathbf{r},\mathbf{r}',\omega)=\sum_i^{\mathrm{occ}}\frac{\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')}{\omega-\varepsilon_i-i\delta}+\sum_j^{\mathrm{unocc}}\frac{\psi_j(\mathbf{r})\psi_j^*(\mathbf{r}')}{\omega-\varepsilon_j+i\delta} \tag{30}$$

The energy is measured with respect to the chemical potential μ.

3.1 Quasiparticles

For atoms and molecules, the excitation energies are discrete and the poles of the Green function are consequently discrete. In crystals, the excitation energies become continuous and the poles of the Green function form a branch cut. It is then not meaningfull to speak about individual poles. It is here the concept of quasiparticle comes in and it may be understood as follows. ¿From Eq. (19), the spectral function A is schematically given by

$$\begin{aligned}A(\omega) &= \frac{1}{\pi}\sum_i|\mathrm{Im}\,G_i(\omega)|\\ &= \frac{1}{\pi}\sum_i\frac{|\mathrm{Im}\,\Sigma_i(\omega)|}{|\omega-\varepsilon_i-\mathrm{Re}\,\Sigma_i(\omega)|^2+|\mathrm{Im}\,\Sigma_i(\omega)|^2}\end{aligned} \tag{31}$$

where G_i is the matrix element of G in an eigenstate ψ_i of the noninteracting system H_0. In a crystal, the state label i corresponds to the wavevector $\mathbf{k}$ and band index n. A is usually

peaked at each energy $E_i = \varepsilon_i - \mathrm{Re}\ \Sigma_i(E_i)$ provided $\mathrm{Im}\Sigma(E_i)$ is small. This peak is called a quasiparticle peak with a life-time given by $|\mathrm{Im}\ \Sigma_i(E_i)|$. The life-time is a consequence of the fact that quasiparticle excitations are not eigenstates of the interacting Hamiltonian.

The weight of the quasiparticle (strength of the Lorentzian) can be obtained by Taylor expanding $\mathrm{Re}\Sigma$ around the quasiparticle energy:

$$\mathrm{Re}\Sigma(\omega) = \mathrm{Re}\Sigma(E_i) + (\omega - E_i)\frac{\partial \mathrm{Re}\Sigma(E_i)}{\partial \omega} + \cdots \tag{32}$$

We obtain for ω close to E_i

$$\mathrm{Im}G_i(\omega) = Z_i \frac{|\mathrm{Im}\ \Sigma_i(\omega)|}{|\omega - E_i|^2 + |\mathrm{Im}\ \Sigma_i(\omega)|^2} \tag{33}$$

where

$$Z_i = \left[1 - \frac{\partial \mathrm{Re}\ \Sigma_i(E_i)}{\partial \omega}\right]^{-1} < 1 \tag{34}$$

¿From the definition of G we have

$$\int d\omega A(\mathbf{r}, \mathbf{r}'; \omega) = \delta(\mathbf{r} - \mathbf{r}') \tag{35}$$

or

$$\int d\omega A_i(\omega) = 1 \tag{36}$$

Since Z_i is usually less than one, this means the rest of the weight must be distributed at other energies. At some other energies ω_p, the denominator may be small and $A(\omega_p)$ could also show peaks or satellite (incoherent) structure which can be due to plasmon excitations or other collective phenomena. This is in contrast to the nonineracting case where the quasiparticle peak is just a delta function with an infinite life-time and without any incoherent features.

¿From the classical theory of the Green functions a solution to equation (18) can be written as

$$G(\mathbf{r}, \mathbf{r}', \omega) = \sum_i \frac{\Psi_i(\mathbf{r}, \omega)\Psi_i^\dagger(\mathbf{r}', \omega)}{\omega - E_i(\omega)} \tag{37}$$

where Ψ_i are solutions to

$$H_0\,(\mathbf{r})\,\Psi_i\,(\mathbf{r}, \omega) + \int d^3r\ \Sigma(\mathbf{r}, \mathbf{r}', \omega)\Psi_i(\mathbf{r}', \omega) = E_i(\omega)\Psi_i(\mathbf{r}, \omega) \tag{38}$$

We define a quasiparticle wave function Ψ_i with energy E_i as a solution to

$$H_0(\mathbf{r})\Psi_i(\mathbf{r}) + \int d\mathbf{r}'\Sigma(\mathbf{r}, \mathbf{r}', E_i)\Psi_i(\mathbf{r}') = E_i\Psi_i(\mathbf{r}) \tag{39}$$

The eigenvalues E_i are in general complex and the quasiparticle wavefunctions are not in general orthogonal because Σ is not Hermitian but both the real and imaginary part of Σ are symmetric.

3.2 Derivation of the self-energy

There are several ways of evaluating the self-energy, either by using Wick's theorem (see, e.g. Ref. [36]) or by functional derivative. We follow the latter because it is physically appealing and give a summary of the steps. For more details we refer to Refs. [47, 59]. We introduce a time varying field $\phi(x)$ which functions as a mathematical device and will be set to zero in the

end. It is similar to the principle of virtual work in classical mechanics. For a time-varying field it is convenient to work in the interaction (Dirac) representation:

$$|\psi_D(\mathbf{r},t)\rangle = \hat{U}(t,t_0)\ |\psi_D(\mathbf{r},t_0)\rangle \tag{40}$$

The time development operator $\hat{U}$ is given by

$$\hat{U}(t,t_0) = T\ \exp\left[-i\int_{t_0}^{t} d\tau \hat{\phi}(\tau)\right] \tag{41}$$

$$\hat{\phi}(t) = \int d^3r\ \phi(\mathbf{r},t)\ \hat{\psi}_D^\dagger(\mathbf{r},t)\hat{\psi}_D(\mathbf{r},t) \tag{42}$$

The relationship between operators in the Heisenberg and interaction representation is

$$\hat{\psi}_H(x) = \hat{U}^\dagger(t,0)\hat{\psi}_D(x)\hat{U}(t,0) \tag{43}$$

The field operator $\hat{\psi}_D(x)$ satisfies

$$i\frac{\partial}{\partial t}\hat{\psi}_D(x) = [\hat{\psi}(x), \hat{H}(\hat{\phi}=0)] \tag{44}$$

so it is the same as the unperturbed ($\phi = 0$) Heisenberg operator. The Green function can now be written as

$$G(x,x') = -i\frac{\langle N^0|\ T[\hat{U}(\infty,-\infty)\hat{\psi}_D(x)\hat{\psi}_D^\dagger(x')|N^0\rangle}{\langle N^0|\hat{U}(\infty,-\infty)]|N^0\rangle} \tag{45}$$

and $|N^0\rangle$ is the unperturbed groundstate (the *interacting* groundstate before the application of ϕ). By taking functional derivative of G with respect to ϕ it can be shown that

$$\frac{\delta G(1,2)}{\delta\phi(3)} = G(1,2)G(3,3^+) - G_2(1,2,3,3^+) \tag{46}$$

Using this result in Eq. (15) we get

$$\int d3\ M(1,3)G(3,2) = V_H(1)G(1,2) + i\int d3\ v(1,3)\frac{\delta G(1,2)}{\delta\phi(3)} \tag{47}$$

where

$$v(1,3) \equiv v(|\mathbf{r}_1 - \mathbf{r}_3|)\delta(t_1 - t_3) \tag{48}$$

and V_H is the Hartree potential

$$V_H(x) = \int dx' v(x,x')\rho(x') \tag{49}$$

Using the definition of the self-energy $\Sigma = M - V_H$ and using the identity

$$\frac{\delta}{\delta\phi}(G^{-1}G) = 0 \rightarrow \frac{\delta G}{\delta\phi} = G\frac{\delta G^{-1}}{\delta\phi}G \tag{50}$$

we have

$$\Sigma(1,2) \;=\; -i\int d3\ d4\ v(1,4)G(1,3)\frac{\delta G^{-1}(3,2)}{\delta\phi(4)} \tag{51}$$

This expression is exact. The quantity $\delta G^{-1}/\delta\phi$ is related to the dielectric or response function. Multiplying both sides of Eq. (18) by G^{-1} and keeping in mind that h_0 contains the probing field ϕ we find

$$\frac{\delta G^{-1}(3,2)}{\delta\phi(4)} = -\delta(3-2)\underbrace{\left[\delta(3-4)+\frac{\delta V_H(3)}{\delta\phi(4)}\right]}_{\epsilon^{-1}(3,4)} - \frac{\delta\Sigma(3,2)}{\delta\phi(4)} \tag{52}$$

We can see very clearly the stucture of the self-energy: Including the first term of $\delta G^{-1}/\delta\phi$ results in the well-known Hartree-Fock approximation. Including further the second term, i.e., the change in the Hartree potential, in fact leads us to the GWA because, as indicated, $1+\delta V_H/\delta\phi$ is the inverse dielectric function so that

$$W(1,2) = \int d3v(1,3)\epsilon^{-1}(3,2) \tag{53}$$

is the screened Coulomb potential. We therefore have

$$\Sigma_{GW}(1,2) = iG(1,2)W(1,2) \tag{54}$$

The last term in $\delta G^{-1}/\delta\phi$ is the response of the self-energy to the probing field and this constitutes the vertex corrections. Thus, apart from the Green function, response functions are very important quantities in the calculations of the self-energy. Note that we have not invoked any diagrams or many-body perturbation theory in deriving the GWA or the self-energy in general.

3.3 The response function

We have seen that the Schwinger functional derivative technique is a very convenient tool in Green's function theory. Here we demonstrate again its usefulness by deriving the equation satisfied by the charge response function. It is convenient to use a convention where repeated index or variable means a summation or an integral. We first define the polarisation function which may be thought of as a response function but with respect to the total potential $V = \phi + V_H$:

$$\begin{aligned}
P(1,2) &= \frac{\delta\rho(1)}{\delta V(2)} \\
&= -i\frac{\delta G(1,1^+)}{\delta V(2)} \\
&= -iG(1,3)\frac{\delta G^{-1}(3,4)}{\delta V(2)}G(4,1^+) \\
&= iG(1,3)\left\{\delta(3-4)\delta(3-2)+\frac{\delta\Sigma(3,4)}{\delta V(2)}\right\}G(4,1^+) \\
&= iG(1,3)\Lambda(3,4,2)G(4,1^+)
\end{aligned} \tag{55}$$

where Λ is known as the vertex function:

$$\Lambda(1,2,3) = \delta(1-2)\delta(1-3)+\frac{\delta\Sigma(1,2)}{\delta V(3)} \tag{56}$$

We have used the identity in Eq. (50) and Eq. (52) for $\delta G^{-1}/\delta\phi$. From

$$\begin{aligned}
\epsilon^{-1} &= \frac{\delta V}{\delta\phi} \\
&= 1+v\frac{\delta\rho}{\delta\phi}
\end{aligned}$$

$$
\begin{aligned}
&= 1 + v\frac{\delta\rho}{\delta V}\frac{\delta V}{\delta\phi} \\
&= 1 + vP\epsilon^{-1}
\end{aligned}
\tag{57}
$$

we find

$$
\epsilon^{-1} = [1 - vP]^{-1} \quad \text{or} \quad \epsilon = 1 - vP \tag{58}
$$

We now calculate the response function:

$$
\begin{aligned}
R(1,2) &= \frac{\delta\rho(1)}{\delta\phi(2)} \\
&= \frac{\delta\rho(1)}{\delta V(3)}\frac{\delta V(3)}{\delta\phi(2)} \\
&= P(1,3)\epsilon^{-1}(3,2)
\end{aligned}
\tag{59}
$$

Thus schematically

$$
R = P[1 - vP]^{-1} \quad \text{or} \quad R = [1 - Pv]^{-1}P \tag{60}
$$

This equation for R is exact. If we use a noninteracting P^0 we arrive at the well-known time-dependent Hartree or the random phase approximation (RPA) equation. The RPA has a simple physical interpretation: The change in density due to a perturbing external field is given by $\delta\rho = P^0\delta(\delta V_{\text{ext}} + \delta V_H)$, i.e., the system response to the *total* field as if it were not interacting. When $1 - Pv = 0$ we expect a new mode of excitation which can be understood as follows: A density fluctuation $\delta\rho$ giving rise to a potential $\delta v = v\delta\rho$ such that $Pv\delta\rho = P\delta v = \delta\rho$, i.e., the potential generated by the density fluctuation induces the same density fluctuation. This new mode of excitation is usually referred to as a plasmon. Thus a plasmon is a self-sustaining density fluctuation.

By using the spectral representation for G in Eq. (24, the polarization function can also be expressed in terms of its spectral representation:

$$
P(\mathbf{r},\mathbf{r}',\omega) = \int_{-\infty}^{0} d\omega' \, \frac{S(\mathbf{r},\mathbf{r}',\omega')}{\omega - \omega' - i\delta} + \int_{0}^{\infty} d\omega' \, \frac{S(\mathbf{r},\mathbf{r}',\omega')}{\omega - \omega' + i\delta} \tag{61}
$$

where S is proportional to the imaginary part of P and defined to be anti-symmetric in ω:

$$
\begin{aligned}
S(\mathbf{r},\mathbf{r}',\omega) &= -\frac{1}{\pi}\text{Im}\, P(\mathbf{r},\mathbf{r}',\omega)\text{sgn}(\omega) \\
&= \sum_{i,j} h_i(\mathbf{r})h_i^*(\mathbf{r}')p_j(\mathbf{r})p_j^*(\mathbf{r}')\delta[\omega - e(N+1,j) - e(N-1,i)]
\end{aligned}
\tag{62}
$$

For a noninteracting system, $h_i \to \psi_i^{\text{occ}}$, $p_j \to \psi_j^{\text{unocc}}$ and $e(N-1,i) \to \varepsilon_i^{\text{occ}}$, $e(N+1,j) \to \varepsilon_j^{\text{unocc}}$. Similarly, the response function R can be written in the form of Eq. (61).

In terms of the many-body states, $|M\rangle$, the exact time-ordered response function can be expressed as

$$
R(1,2;\omega) = \sum_M \left[\frac{\langle 0|\hat{\rho}(1)|M\rangle\langle M|\hat{\rho}(2)|0\rangle}{\omega - E_M + E_0 + i\delta} - \frac{\langle 0|\hat{\rho}(2)|M\rangle\langle M|\hat{\rho}(1)|0\rangle}{\omega + E_M - E_0 - i\delta} \right] \tag{63}
$$

Thus the poles of the response function yield the exact excitation energies.

4 The GW approximation

The GWA is usually regarded as the first term in the expansion of the self-energy in the screened interaction W. This point ov view is based on many-body perturbation theory. This is, however, not very useful from physical point of view. If we do include higher order diagrams, we do not necessarily get better results (in the sense of closer agreement with experiment). Straightforward higher order diagrams can give a self-energy with wrong analytic properties which in turns results in an unphysically negative spectral function (density of states) for some energies.

Eqs. (51) and (52 on the other hand are exact and the GWA has been obtained by neglecting the vertex $\delta\Sigma/\delta\phi$ in $\delta G^{-1}/\delta\phi$. It has not been obtained by summing diagrams as is normally done in many-body perturbation theory. The GW self-energy has the same form as the exchange operator in the HFA with the bare Coulomb potential substituted by a screened Coulomb interaction. It is more physical and useful to regard the GWA as a Hartree-Fock approximation with a frequency-dependent screening which cures the most serious deficiency of the HFA.

The GWA has been applied with success to many systems ranging from alkali metals [99, 77, 82, 83], semiconductors [55, 41], transition metals [8, 11], metal surfaces [30, 31] to clusters [85]. The list is by no means exhaustive.

4.1 Explicit expression for Σ_{GW}

Fourier transforming Eq. (54) we obtain

$$\begin{aligned}\Sigma(\mathbf{r},\mathbf{r}',\omega) &= \frac{i}{2\pi}\int d\omega'\, G(\mathbf{r},\mathbf{r}',\omega+\omega')W(\mathbf{r},\mathbf{r}',\omega') \\ &= \frac{i}{2\pi}\int d\omega'\, G(\mathbf{r},\mathbf{r}',\omega+\omega')v(\mathbf{r}-\mathbf{r}') \\ &\quad +\frac{i}{2\pi}\int d\omega'\, G(\mathbf{r},\mathbf{r}',\omega+\omega')W_c(\mathbf{r},\mathbf{r}',\omega') \qquad (64)\end{aligned}$$

The first term gives the bare exchange

$$\Sigma^x(\mathbf{r},\mathbf{r}') = -\sum_i f_i(\mathbf{r})f_i^*(\mathbf{r}')v(\mathbf{r}-\mathbf{r}') \qquad (65)$$

and the second term gives the correlation part of the self-energy which has the following spectral representation:

$$\Sigma^c(\mathbf{r},\mathbf{r}',\omega) = \int_{-\infty}^{\mu} d\omega'\, \frac{\Gamma(\mathbf{r},\mathbf{r}',\omega')}{\omega-\omega'-i\delta} + \int_{\mu}^{\infty} d\omega'\, \frac{\Gamma(\mathbf{r},\mathbf{r}',\omega')}{\omega-\omega'+i\delta} \qquad (66)$$

The spectral representation of the correlation part of the screened Coulomb interaction $W_c = W - v$ is

$$W_c(\mathbf{r},\mathbf{r}',\omega) = \int_{-\infty}^{0} d\omega'\, \frac{D(\mathbf{r},\mathbf{r}',\omega')}{\omega-\omega'-i\delta} + \int_{0}^{\infty} d\omega'\, \frac{D(\mathbf{r},\mathbf{r}',\omega')}{\omega-\omega'+i\delta} \qquad (67)$$

Using the spectral representations of G and W_c in Eq. (64) gives (Appendix B)

$$\begin{aligned}\Gamma(\mathbf{r},\mathbf{r}';\omega) &= -\mathrm{sgn}(\omega-\mu)\int_0^{\infty} d\omega'\theta(\mu-\omega-\omega')A(\mathbf{r},\mathbf{r}',\omega+\omega')D(\mathbf{r},\mathbf{r}',\omega) \\ &\quad +\mathrm{sgn}(\omega-\mu)\int_0^{\infty} d\omega'\theta(-\mu+\omega-\omega')A(\mathbf{r},\mathbf{r}',\omega-\omega')D(\mathbf{r},\mathbf{r}',\omega) \\ &= -\mathrm{sgn}(\omega-\mu)\sum_i f_i(\mathbf{r})f_i^*(\mathbf{r}')\; D(\mathbf{r},\mathbf{r}',\alpha_i(\omega))\;\theta[\alpha_i(\omega)] \\ &\quad +\mathrm{sgn}(\omega-\mu)\sum_i g_i(\mathbf{r})g_i^*(\mathbf{r}')\; D(\mathbf{r},\mathbf{r}',\beta_i(\omega))\;\theta[\beta_i(\omega)] \qquad (68)\end{aligned}$$

where

$$\alpha_i(\omega) = -\omega + \mu - e(N-1,i) \quad (69)$$
$$\beta_i(\omega) = \omega - \mu - e(N+1,i) \quad (70)$$

The real part of the self-energy can be obtained from the Hilbert transform in Eq. (66).

4.2 Σ_{GW} with a noninteracting G_0

In practice, most GW calculations have been performed using a noninteracting Green's function. The issue of self-consistency will be discussed at a later section. For a non-interacting system, we have

$$\begin{aligned} e(N-1,i) &= E(N-1,i) - E(N-1) \\ &= -\varepsilon_i + E(N) - E(N-1) \\ &= -\varepsilon_i + \mu \quad (71) \\ e(N+1,j) &= \varepsilon_j - \mu \quad (72) \end{aligned}$$

$$f_i = \psi_i, \quad \text{with } \varepsilon_i \le \mu \quad (73)$$
$$g_j = \psi_j^*, \quad \text{with } \varepsilon_j > \mu \quad (74)$$

where ε_i is the single-particle eigenvalue of the the single-particle state ψ_i of the non-interacting system. The spectral function of G_0 is

$$A^0(\mathbf{r},\mathbf{r}',\omega) = \sum_i \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')\delta(\omega - \varepsilon_i) \quad (75)$$

since A is real. For a non-interacting system, the spectral function consists of delta function peaks at the single-particle eigenvalues whereas in the interacting system, apart from the quasi-particle peak, the spectral function may have a satellite structure.

The spectral function of P for the non-interacting system is

$$S^0(\mathbf{r},\mathbf{r}',\omega) = \sum_{i\le\mu}\sum_{j>\mu} \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')\psi_j^*(\mathbf{r})\psi_j(\mathbf{r}')\ \delta[\omega - (\varepsilon_j - \varepsilon_i)] \quad (76)$$

and the spectral function for Σ_{GW} is (Appendix B)

$$\Gamma(\mathbf{r},\mathbf{r}',\omega) = \begin{cases} \sum_{i\le\mu} \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')D(\mathbf{r},\mathbf{r}',\varepsilon_i - \omega)\theta(\varepsilon_i - \omega) & \text{for } \omega \le \mu \\ \sum_{i>\mu} \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}')D(\mathbf{r},\mathbf{r}',\omega - \varepsilon_i)\theta(\omega - \varepsilon_i) & \text{for } \omega > \mu \end{cases} \quad (77)$$

The bare exchange part becomes

$$\Sigma^x(\mathbf{r},\mathbf{r}') = -\sum_{i\le\mu} \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')v(\mathbf{r}-\mathbf{r}') \quad (78)$$

4.3 Coulomb hole and screened exchange (COHSEX)

A physically appealing way of expressing the self-energy is by dividing it into a screened-exchange term Σ_{SEX} and a Coulomb-hole term Σ_{COH} (COHSEX) [46, 47]. It is straightforward to verify that the real part of the self-energy can be written as

$$\mathrm{Re}\Sigma_{\mathrm{SEX}}(\mathbf{r},\mathbf{r}',\omega) = -\sum_i^{\mathrm{occ}} \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')\mathrm{Re}W(\mathbf{r},\mathbf{r}',\omega - \varepsilon_i) \quad (79)$$

$$\mathrm{Re}\Sigma_{\mathrm{COH}}\left(\mathbf{r},\mathbf{r}',\omega\right)=\sum_i\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')\mathrm{P}\int_0^\infty d\omega'\frac{D\left(\mathbf{r},\mathbf{r}',\omega'\right)}{\omega-\varepsilon_i-\omega'} \tag{80}$$

The physical interpretation of Σ_{COH} becomes clear in the static approximation due to Hedin (1965a). If we are interested in a state with energy E close to the Fermi level, the matrix element $\langle\psi|\mathrm{Re}\Sigma_{\mathrm{COH}}(E)|\psi\rangle$ picks up most of its weight from states with energies ε_i close to E in energy. We may then assume that $E-\varepsilon_i$ is small compared to the main excitation energy of D, which is at the plasmon energy. If we set $E-\varepsilon_i=0$, we get

$$\mathrm{Re}\Sigma_{\mathrm{COH}}\left(\mathbf{r},\mathbf{r}'\right)=\frac{1}{2}\delta\left(\mathbf{r}-\mathbf{r}'\right)W^c\left(\mathbf{r},\mathbf{r}',0\right) \tag{81}$$

This is simply the interaction energy of the quasiparticle with the induced potential due to the screening of the electrons around the quasiparticle. The factor of $1/2$ arises from the adiabatic growth of the interaction. In this static COHSEX approximation, Σ_{COH} becomes local.

4.4 Generalization of the HFA

The GWA may be regarded as a generalization of the HFA. ¿From Eq. (64) and using the spectral representation of G_0, the correlated part of the self-energy can be written as

$$\begin{aligned}\Sigma_c(\mathbf{r},\mathbf{r}';\omega) &= \sum_{i\le\mu}\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')W_c^-(\mathbf{r},\mathbf{r}';\omega-\epsilon_i)\\ &+ \sum_{i>\mu}\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')W_c^+(\mathbf{r},\mathbf{r}';\omega-\epsilon_i)\end{aligned} \tag{82}$$

where

$$W_c^\pm(\mathbf{r},\mathbf{r}';\omega) = \frac{i}{2\pi}\int_{-\infty}^\infty d\omega'\frac{W_c(\mathbf{r},\mathbf{r}';\omega')}{\omega+\omega'\pm i\delta} \tag{83}$$

In short we can write

$$\Sigma(\mathbf{r},\mathbf{r}';\omega)=-\sum_i\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')W_0(\mathbf{r},\mathbf{r}';\omega-\epsilon_i) \tag{84}$$

where

$$\begin{aligned}W_0(\mathbf{r},\mathbf{r}';\omega-\epsilon_i) \equiv\ & \left\{v(\mathbf{r}-\mathbf{r}')-W_c^-(\mathbf{r},\mathbf{r}';\omega-\epsilon_i)\right\}\,\theta(\mu-\epsilon_i)\\ & -W_c^+(\mathbf{r},\mathbf{r}';\omega-\epsilon_i)\,\theta(\epsilon_i-\mu)\end{aligned} \tag{85}$$

Thus, the self-energy in the GWA has the same form as in the HFA except that it depends on energy and contains a term which depends on unoccupied states as a consequence of correlation effects. Thus the GWA can be interpreted as a generalization of the Hartree-Fock approximation (HFA) with a potential W_0 which contains dynamical screening of the Coulomb potential. Note, however, that W_0 is not the same as the dynamically screened potential W.

As previously mentioned, the HFA does not take into account the effects of screening which, for insulators, results in too large band gaps. It can be shown that the GWA gives the correct band gap, at least for localized states which are well isolated from the other states. Consider the correlated part of the self-energy for an occupied core-like state ψ_d.

$$\begin{aligned}\langle\psi_d|\Sigma_c(\epsilon_d)|\psi_d\rangle &= \langle\psi_d\psi_d|W_c^-(0)|\psi_d\psi_d\rangle\\ &+ \sum_{i\neq d}^{occ}\langle\psi_d\psi_i|W_c^-(\epsilon_d-\epsilon_i)|\psi_i\psi_d\rangle\\ &+ \sum_i^{unocc}\langle\psi_d\psi_i|W_c^+(\epsilon_d-\epsilon_i)|\psi_i\psi_d\rangle\end{aligned} \tag{86}$$

Strictly speaking, the self-energy should be evaluated at the new energy $E_d = \epsilon_d +$ self-energy correction and it is understood to be the case here. If ψ_d is localized and well separated in energy from other states, then the first term is evidently much larger than the rest. Thus we may make the following approximation:

$$\begin{aligned} \langle\psi_d|\Sigma_c(\epsilon_d)|\psi_d\rangle &\approx \langle\psi_d\psi_d|W_c^-(0)|\psi_d\psi_d\rangle \\ &= -\frac{1}{2}\langle\psi_d\psi_d|W_c(0)|\psi_d\psi_d\rangle \end{aligned} \tag{87}$$

The last step is shown in Appendix A. This is a correction due to the work done on the electron by the polarization field from zero to $W_c(0)$ [46, 47]. A similar result,

$$+\frac{1}{2}\langle\psi_d\psi_d|W_c(0)|\psi_d\psi_d\rangle$$

is obtained for an unoccupied core-like state of the same character so that the energy separation of the states is

$$\begin{aligned} \Delta &= \epsilon_2^{HF} - \epsilon_1^{HF} + \langle\psi_d\psi_d|W_c(0)|\psi_d\psi_d\rangle \\ &= \langle\psi_d\psi_d|v|\psi_d\psi_d\rangle + \langle\psi_d\psi_d|W_c(0)|\psi_d\psi_d\rangle \\ &= \langle\psi_d\psi_d|W(0)|\psi_d\psi_d\rangle \end{aligned} \tag{88}$$

which agrees with the intuitive result that the "gap" is given by the screened Coulomb interaction: $\Delta = U \approx W(0)$. Since $W_c(0)$ is negative, we see that the self-energy correction to the HFA raises an occupied state and lowers an unoccupied state. This is still true also for states which are not so localized.

5 Vertex corrections: Beyond *GW*

The GWA has proven to be very successful for describing quasiparticle energies for *sp* and even 3*d* systems. Its description for satellite structures, however, is less satisfactory. A number of cases which reveal the shortcomings of the GWA in describing the satellite structure are

- The valence as well as the core photoemission spectra of the alkalis show a series of plasmon satellites which are located at multiple of plasmon energies below the main quasiparticle peak. The GWA gives only one plasmon satellite and its position is typically 1.5 plasmon energy below the quasiparticle peak [48].
- The valence photoemission spectrum of Ni shows the presence of a satellite at 6 eV which is not obtained within the GWA. Similarly, the GWA appears to be insufficient for describing the satellite structure in transition metal oxides and the situation becomes worse in the f-systems.

The presence of only one plasmon satellite in the GWA follows from the fact that the self-energy is of first order in W which contains one plasmon excitation through the RPA dielectric function. From the diagram, it is clear that a hole or an electron interacts with its surrounding by exchanging a plasmon. The too large plasmon energy can be understood qualitatively as an average of the first and the second plasmon energies since they carry most of the satellite weight. Because the peak in $\mathrm{Im}\,\Sigma^c$ is at one plasmon energy below the quasiparticle peak, the plasmon satellite in the spectral function ends up at $\approx$1.5 plasmon energy To get the correct satellite energy the peak in $\mathrm{Im}\,\Sigma^c$ should therefore lie somewhat closer to the quasiparticle energy.

The problem with the satellite structure in Ni or transition metal oxides is of a different nature since the satellite energy is much lower than the plasmon energy. In the atomic picture, the

ground state of Ni is a mixture of the configurations $3d^9 4s$ and $3d^8 4s^2$. The final configurations after photoemission are $3d^8 4s$ and $3d^7 s^2$. The former corresponds to the main line (quasiparticle) and the latter to the satellite. The two configurations are separated in energy by 6 eV which is essentially the Coulomb energy of two d holes. This value is reduced by screening in the solid. The presence of satellite at a certain energy implies that the imaginary part of the self-energy should exhibit a peak at an energy slightly lower then the satellite energy. Such a two-hole excitation state is partly described by the GWA to second order in the bare interaction but the GWA mainly describes the coupling of the electrons to the plasmon excitation. A secondary plasmon-like excitation is possible within the RPA when the band structure gives rise to two well-defined peaks in Im ϵ^{-1}.

It is a major challenge to develop a theory beyond the GWA for real systems which can overcome the problems described above. Here we describe those approaches which are based on systematic diagrammatic expansions. One approach which has been applied with success to the alkalis is the cumulant expansion [66, 19, 49]. As described in the next section, this approach is suitable for dealing with systems which can be mapped into systems of electrons coupled to bosons (e.g. plasmons) where long-range correlations dominate. However, short-range correlations arising from multiple-hole interactions on the same site should be better treated within the T-matrix approach, described in a later section. In this contribution, we are concerned with vertex corrections related to the satellite description rather than quasiparticle description.

The effects of vertex corrections on the quasiparticles have been studied by a number of authors. These include a direct calculation of the second-order diagram [27, 28, 21] and vertex correction based on the LDA V^{xc} with application to Si [32]. Vertex corrections in the electron gas [77, 37] are discussed in a review article by Mahan [79] where it is emphasized that it is important to include vertex corrections in both the response function and the self-energy in a consistent way [107, 14, 15, 90].

5.1 The cumulant expansion

One of the first applications of the cumulant expansion method was in the problems of X-ray spectra of core-electrons in metals [84, 66]. The problem is modelled by a Hamiltonian consisting of a core electron interacting with electron-hole excitations and a set of plasmons:

$$\begin{aligned} H &= \varepsilon c^\dagger c + \sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}\sigma} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + \sum_{\mathbf{q}} \omega_{\mathbf{q}} b^\dagger_{\mathbf{q}} b_{\mathbf{q}} \\ &\quad + \sum_{\mathbf{k}\mathbf{k}'\sigma} V_{\mathbf{k}\mathbf{k}'} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma'} c c^\dagger + \sum_{\mathbf{q}} c c^\dagger g_{\mathbf{q}} \left(b_{\mathbf{q}} + b^\dagger_{\mathbf{q}} \right) \end{aligned} \tag{89}$$

where c is the annihilation operator for the core electron, $c^\dagger_{\mathbf{k}\sigma}$ is the creation operator for a conduction electron of wave vector $\mathbf{k}$, spin σ, and energy $\varepsilon_{\mathbf{k}\sigma}$ and $b^\dagger_{\mathbf{q}}$ is the creation operator for a plasmon of wave vector $\mathbf{q}$ and energy $\omega_{\mathbf{q}}$. The last two terms are the coupling of the core electron to the conduction electrons (electron-hole excitations) and to the plasmons respectively. The model Hamiltonian without the last term [75, 76] was solved exactly by Nozières and de Dominicis [84] and the full Hamiltonian was solved exactly by Langreth [66] who also showed that the cumulant expansion also gives the exact solution. We consider the case with coupling only to the plasmon field [72, 73] but not to the conduction electrons. The exact solution is given by [66]

$$A_\pm(\omega) = \sum_{n=0}^{\infty} \frac{e^{-a} a^n}{n!} \delta(\omega - \varepsilon - \Delta\varepsilon \mp n\omega_p) \tag{90}$$

where $+$ refers to absorption spectrum and $-$ to emission spectrum. $a = \sum_{\mathbf{q}} g^2_{\mathbf{q}}/\omega^2_p$ and $\Delta\varepsilon = a\omega_p$ is the shift in core energy due to the interaction with the plasmon field. The spectra consist

therefore of the main quasiparticle peak at $\omega = \varepsilon + \Delta\varepsilon$ and a series of plasmon excitations at multiples of the plasmon energy below the quasiparticle peak. It has been assumed that the plasmon excitations have no dispersion although this assumption is not necessary.

The physics of the cumulant expansion when applied to valence electrons was discussed in detail by Hedin [49, 1]. More recently, the cumulant expansion was calculated to higher order for a model Hamiltonian by Gunnarsson, Meden, and Schönhammer [43].

5.1.1 Theory

In the cumulant expansion approach, the Green function for the hole ($t < 0$) is written as

$$\begin{aligned} G(k,t) &= i\theta(-t)\langle N|\hat{c}_k^\dagger(0)\hat{c}_k(t)|N\rangle \\ &= i\theta(-t)\,\mathrm{e}^{-i\varepsilon_k t + C^h(k,t)} \end{aligned} \tag{91}$$

where k denotes all possible quantum labels. For $\omega \leq \mu$ only the first term in equation (26) contributes and the hole spectral function is

$$\begin{aligned} A(k,\omega \leq \mu) &= \frac{1}{\pi}\mathrm{Im}G(k,\omega) \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dt e^{i\omega t}\langle N|\hat{c}_k^\dagger(0)\hat{c}_k(t)|N\rangle \\ &= \frac{1}{\pi}\mathrm{Im}\, i\int_{-\infty}^{0} dt\,\mathrm{e}^{i\omega t}\mathrm{e}^{-i\varepsilon_k t + C^h(k,t)} \end{aligned} \tag{92}$$

$C^h(k,t)$ is called the cumulant. Expanding the exponential in powers of the cumulant we get

$$G(k,t) = G_0(k,t)\left[1 + C^h(k,t) + \frac{1}{2}[C^h(k,t)]^2 + \ldots\right] \tag{93}$$

where $G_0(k,t) = i\theta(-t)\exp(-i\varepsilon_k t)$. In terms of the self-energy, the Green function for the hole can be expanded as

$$G = G_0 + G_0\Sigma G_0 + G_0\Sigma G_0\Sigma G_0 + \ldots \tag{94}$$

We may group the cumulant into terms labelled by the order of the interaction n:

$$C^h = \sum_{n=1}^{\infty} C_n^h \tag{95}$$

The above equations for G may now be equated and the cumulant can be obtained by equating terms of the same order in the interaction. Thus to lowest order in the screened interaction W, the cumulant is obtained by equating

$$G_0 C^h = G_0\Sigma G_0 \equiv \Delta G^{(1)} \tag{96}$$

where $\Sigma = \Sigma_{GW} = iG_0W$. If G_0 corresponds to, e.g. the LDA G, then $\Sigma = \Sigma_{GW} - V^{xc}$. The explicit form of $\Delta G^{(1)}$ is

$$\Delta G^{(1)}(1,2) = \int d3d4 G_0(1,3)\,\Sigma(3,4)\,G_0(4,2) \tag{97}$$

Using

$$\begin{aligned} G_0(1,2) &= i\sum_n^{\mathrm{occ}} \phi_n(\mathbf{r}_1)\,\phi_n^*(\mathbf{r}_2)\,\mathrm{e}^{-i\varepsilon_n(t_1-t_2)}\theta(t_2-t_1) \\ &\quad -i\sum_n^{\mathrm{unocc}} \phi_n(\mathbf{r}_1)\,\phi_n^*(\mathbf{r}_2)\,\mathrm{e}^{-i\varepsilon_n(t_1-t_2)}\theta(t_1-t_2) \end{aligned} \tag{98}$$

and taking matrix element with respect to an occupied state ϕ_k yields

$$\Delta G^{(1)}(k,t_1-t_2) = -\mathrm{e}^{-i\varepsilon_k(t_1-t_2)}\int_{t_1}^{\infty} dt_3 \int_{-\infty}^{t_2} dt_4\, \mathrm{e}^{i\varepsilon_k(t_3-t_4)}\Sigma(k,t_3-t_4) \tag{99}$$

Without loss of generality we can set $t_2 = 0$ and $t_1 = t < 0$:

$$\Delta G^{(1)}(k,t) = -\mathrm{e}^{-i\varepsilon_k t}\int_{t}^{\infty} dt_3 \int_{t_3}^{\infty} d\tau\, \mathrm{e}^{i\varepsilon_k\tau}\Sigma(k,\tau) \tag{100}$$

The first-order cumulant is therefore

$$\begin{aligned} C^h(k,t) &= i\int_t^{\infty} dt' \int_{t'}^{\infty} d\tau\, \mathrm{e}^{i\varepsilon_k\tau}\Sigma(k,\tau) \\ &= i\int_t^{0} dt' \int_{t'}^{\infty} d\tau\, \mathrm{e}^{i\varepsilon_k\tau}\Sigma(k,\tau) + C^h(k,0) \end{aligned} \tag{101}$$

$C^h(k,0)$ is a constant which contributes to an asymmetric line shape of the quasiparticle.

The cumulant can be expressed as an integral over frequency by defining a Fourier transform

$$\begin{aligned} \Sigma(k,\tau) &= \int \frac{d\omega}{2\pi}\mathrm{e}^{-i\omega\tau}\Sigma(k,\omega) \\ &= \int \frac{d\omega}{2\pi}\mathrm{e}^{-i\omega\tau}\left\{\int_{-\infty}^{\mu} d\omega' \frac{\Gamma(k,\omega')}{\omega-\omega'-i\delta} + \int_{\mu}^{\infty} d\omega' \frac{\Gamma(k,\omega')}{\omega-\omega'+i\delta}\right\} \\ &= i\theta(-\tau)\,\mathrm{e}^{\delta\tau}\int_{-\infty}^{\mu} d\omega' \mathrm{e}^{-i\omega'\tau}\Gamma(k,\omega') \\ &\quad -i\theta(\tau)\,\mathrm{e}^{-\delta\tau}\int_{\mu}^{\infty} d\omega' \mathrm{e}^{-i\omega'\tau}\Gamma(k,\omega') \end{aligned} \tag{102}$$

The second line uses the spectral representation of Σ in equation (66) and the last line is obtained from

$$\begin{aligned} \int \frac{d\omega}{2\pi}\frac{\mathrm{e}^{-i\omega\tau}}{\omega-\omega'-i\delta} &= \mathrm{e}^{-i\omega'\tau}\int \frac{d\omega''}{2\pi}\frac{\mathrm{e}^{-i\omega''\tau}}{\omega''-i\delta}, \quad \omega'' = \omega-\omega' \\ &= i\theta(-\tau)\,\mathrm{e}^{-i\omega'\tau}\mathrm{e}^{\delta\tau} \end{aligned} \tag{103}$$

Similarly for the other integral. The self-energy becomes

$$\Sigma(k,\tau) = i\theta(-\tau)\,\mathrm{e}^{\delta\tau}\int_{-\infty}^{\mu} d\omega\, \mathrm{e}^{-i\omega\tau}\Gamma(k,\omega) - i\theta(\tau)\,\mathrm{e}^{-\delta\tau}\int_{\mu}^{\infty} d\omega\, \mathrm{e}^{-i\omega\tau}\Gamma(k,\omega) \tag{104}$$

The cumulant can now be expressed as, keeping in mind that $t < 0$,

$$\begin{aligned} C_1^h(k,t) &= i\int_t^{0} dt' \int_{t'}^{\infty} d\tau\, \mathrm{e}^{i\varepsilon_k\tau}\Sigma(k,\tau) \\ &= i\int_t^{0} dt' \int_{t'}^{\infty} d\tau\, \mathrm{e}^{i\varepsilon_k\tau}\left\{i\theta(-\tau)\,\mathrm{e}^{\delta\tau}\int_{-\infty}^{\mu} d\omega\, \mathrm{e}^{-i\omega\tau}\Gamma(k,\omega)\right. \\ &\qquad \left. -i\theta(\tau)\,\mathrm{e}^{-\delta\tau}\int_{\mu}^{\infty} d\omega\, \mathrm{e}^{-i\omega\tau}\Gamma(k,\omega)\right\} \\ &= -\int_t^{0} dt' \int_{t'}^{0} d\tau \int_{-\infty}^{\mu} d\omega\, \mathrm{e}^{i(\varepsilon_k-\omega-i\delta)\tau}\Gamma(k,\omega) \\ &\quad +\int_t^{0} dt' \int_{0}^{\infty} d\tau \int_{\mu}^{\infty} d\omega\, \mathrm{e}^{i(\varepsilon_k-\omega+i\delta)\tau}\Gamma(k,\omega) \end{aligned}$$

$$
\begin{aligned}
&= i\int_t^0 dt' \int_{-\infty}^{\mu} d\omega \frac{1-\mathrm{e}^{i(\varepsilon_k-\omega-i\delta)t'}}{\varepsilon_k-\omega-i\delta}\Gamma(k,\omega) \\
&\quad +i\int_t^0 dt' \int_{\mu}^{\infty} d\omega \frac{\Gamma(k,\omega)}{\varepsilon_k-\omega+i\delta} \\
&= i\int_t^0 dt' \left\{ \int_{-\infty}^{\mu} d\omega \frac{\Gamma(k,\omega)}{\varepsilon_k-\omega-i\delta} + \int_{\mu}^{\infty} d\omega \frac{\Gamma(k,\omega)}{\varepsilon_k-\omega+i\delta} \right\} \\
&\quad + \int_{-\infty}^{\mu} d\omega \frac{\mathrm{e}^{i(\varepsilon_k-\omega-i\delta)t}-1}{(\varepsilon_k-\omega-i\delta)^2}\Gamma(k,\omega) \\
&= -i\Sigma(k,\varepsilon_k)t + \left.\frac{\partial\Sigma^h(k,\omega)}{\partial\omega}\right|_{\omega=\varepsilon_k} \\
&\quad + \int_{-\infty}^{\mu} d\omega \frac{\mathrm{e}^{i(\varepsilon_k-\omega-i\delta)t}}{(\varepsilon_k-\omega-i\delta)^2}\Gamma(k,\omega)
\end{aligned} \tag{105}
$$

The last line is obtained from

$$
\begin{aligned}
\int_{-\infty}^{\mu} d\omega' \frac{\Gamma(k,\omega')}{(\varepsilon_k-\omega'-i\delta)^2} &= -\frac{\partial}{\partial\omega}\int_{-\infty}^{\mu} d\omega' \frac{\Gamma(k,\omega')}{(\omega-\omega'-i\delta)}\bigg|_{\omega=\varepsilon_k} \\
&= -\left.\frac{\partial\Sigma^h(k,\omega)}{\partial\omega}\right|_{\omega=\varepsilon_k}
\end{aligned} \tag{106}
$$

where

$$
\Sigma^h(k,\omega) = \int_{-\infty}^{\mu} d\omega' \frac{\Gamma(k,\omega')}{\omega-\omega'-i\delta} \tag{107}
$$

We can also evaluate $C^h(k,0)$:

$$
\begin{aligned}
C^h(k,0) &= i\int_0^{\infty} dt' \int_{t'}^{\infty} d\tau\, \mathrm{e}^{i\varepsilon_k\tau} \left\{ i\theta(-\tau)\,\mathrm{e}^{\delta\tau} \int_{-\infty}^{\mu} d\omega\, \mathrm{e}^{-i\omega\tau}\Gamma(k,\omega) \right. \\
&\qquad \left. -i\theta(\tau)\,\mathrm{e}^{-\delta\tau} \int_{\mu}^{\infty} d\omega\, \mathrm{e}^{-i\omega\tau}\Gamma(k,\omega) \right\} \\
&= \int_0^{\infty} dt' \int_{t'}^{\infty} d\tau \int_{\mu}^{\infty} d\omega\, \mathrm{e}^{i(\varepsilon_k-\omega+i\delta)\tau}\Gamma(k,\omega) \\
&= i\int_0^{\infty} dt' \int_{\mu}^{\infty} d\omega \frac{\mathrm{e}^{i(\varepsilon_k-\omega+i\delta)t'}}{\varepsilon_k-\omega+i\delta}\Gamma(k,\omega) \\
&= -\int_{\mu}^{\infty} d\omega \frac{\Gamma(k,\omega)}{(\varepsilon_k-\omega+i\delta)^2} \\
&= \left.\frac{\partial\Sigma^p(k,\omega)}{\partial\omega}\right|_{\omega=\varepsilon_k}
\end{aligned} \tag{108}
$$

where

$$
\Sigma^p(k,\omega) = \int_{\mu}^{\infty} d\omega' \frac{\Gamma(k,\omega')}{\omega-\omega'+i\delta} \tag{109}
$$

The total cumulant can be conveniently divided into a quasiparticle part which is linear in t and a satellite part:

$$
C^h = C^h_{QP} + C^h_S \tag{110}
$$

Thus collecting the linear terms in t from equations (105) and (108) we get the cumulant contribution to the quasiparticle:

$$
C^h_{QP}(k,t) = (i\alpha_k + \gamma_k) + (-i\Delta\varepsilon_k + \eta_k)t \tag{111}
$$

where

$$i\alpha_k + \gamma_k = \left.\frac{\partial \Sigma(k,\omega)}{\partial \omega}\right|_{\omega=\varepsilon_k}, \quad \Delta\varepsilon_k = \mathrm{Re}\Sigma(\mathrm{k},\varepsilon_\mathrm{k}), \quad \eta_\mathrm{k} = \mathrm{Im}\Sigma(\mathrm{k},\varepsilon_\mathrm{k}) \tag{112}$$

The contribution to the satellite is

$$C_S^h(k,t) = \int_{-\infty}^{\mu} d\omega \frac{\mathrm{e}^{i(\varepsilon_k-\omega-i\delta)t}}{(\varepsilon_k-\omega-i\delta)^2}\Gamma(k,\omega) \tag{113}$$

A similar derivation can be carried out for the particle Green function

$$G(k,t>0) = -i\theta(t)\,\mathrm{e}^{-i\varepsilon_k t + C^p(k,t)} \tag{114}$$

where k labels an unoccupied state. The result is

$$C_{QP}^p(k,t) = (i\alpha_k + \gamma_k) + (-i\Delta\varepsilon_k - \eta_k)t \tag{115}$$

$$C_S^p(k,t) = \int_{\mu}^{\infty} d\omega \frac{\mathrm{e}^{i(\varepsilon_k-\omega+i\delta)t}}{(\varepsilon_k-\omega+i\delta)^2}\Gamma(k,\omega) \tag{116}$$

It is physically appealing to extract the quasiparticle part from the Green function:

$$\begin{aligned} G_{QP}^h(k,t) &= i\theta(-t)\,\mathrm{e}^{-i\varepsilon_k t + C_{QP}^h(k,t)} \\ &= i\theta(-t)\,\mathrm{e}^{i\alpha_k+\gamma_k}\mathrm{e}^{(-iE_k+\eta_k)t}, \quad E_k = \varepsilon_k + \Delta\varepsilon_k \end{aligned} \tag{117}$$

The spectral function for this quasiparticle can be calculated analytically:

$$A_{QP}(k,\omega<\mu) = \frac{\mathrm{e}^{-\gamma_k}}{\pi}\frac{\eta_k\cos\alpha_k - (\omega-E_k)\sin\alpha_k}{(\omega-E_k)^2+\eta_k^2} \tag{118}$$

Thus we can see that the quasiparticle peak is essentially determined by the GW value.

¿From equation (91) we have for $t<0$

$$\langle N|\hat{c}_k^\dagger(0)\hat{c}_k(t)|N\rangle = \mathrm{e}^{-i\varepsilon_k t + C^h(k,t)} \tag{119}$$

By analytical continuation to $t>0$ and using equation (92) the spectral function can be rewritten as

$$A(k,w) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dt\,\mathrm{e}^{i\omega t}\mathrm{e}^{-i\varepsilon_k t + C^h(k,t)} \tag{120}$$

where C^h for positive t is obtained from $C^{h*}(-t) = C^h(t)$ since $A(k,\omega)$ must be real.

The total spectra can be written as a sum of A_{QP} and a convolution between the quasiparticle and the satellite part:

$$\begin{aligned} A(k,\omega) &= A_{QP}(k,\omega) + \frac{1}{2\pi}\int_{-\infty}^{\infty} dt\,\mathrm{e}^{i\omega t}\mathrm{e}^{(-iE_k+\Gamma_k+C^h(k,0))t}\left[\mathrm{e}^{C_S^h(k,t)}-1\right] \\ &= A_{QP}(k,\omega) + A_{QP}(k,\omega) * A_S(k,\omega) \end{aligned} \tag{121}$$

where

$$\begin{aligned} A_S(k,\omega) &= \frac{1}{2\pi}\int dt\,\mathrm{e}^{i\omega t}\left\{\mathrm{e}^{C_S^h(k,t)}-1\right\} \\ &= \frac{1}{2\pi}\int dt\,\mathrm{e}^{i\omega t}\left\{C_S^h(k,t) + \frac{1}{2!}\left[C_S^h(k,t)\right]^2 + \ldots\right\} \end{aligned} \tag{122}$$

The second term $A_{QP} * A_S$ is responsible for the satellite structure. The Fourier transform of C_S^h can be done analytically

$$\begin{aligned} C_S^h(k,\omega<0) &= \int_{-\infty}^{\infty}\frac{dt}{2\pi}e^{i\omega t}C_S^h(k,t) \qquad (123)\\ &= \int_{-\infty}^{0}\frac{dt}{2\pi}\left\{e^{i\omega t}C_S^h(k,t)+e^{-i\omega t}C_S^h(k,-t)\right\}\\ &= \int_{-\infty}^{0}\frac{dt}{2\pi}\left\{e^{i\omega t}C_S^h(k,t)+\text{c.c.}\right\}\\ &= \int_{-\infty}^{0}\frac{dt}{2\pi}\int_{-\infty}^{\mu}d\omega'\frac{e^{i(\varepsilon_k+\omega-\omega'-i\delta)t}}{(\varepsilon_k-\omega'-i\delta)^2}\Gamma(k,\omega')+\text{c.c.} \qquad (124)\end{aligned}$$

Integrating over t gives

$$\begin{aligned} C_S^h(k,\omega<0) &= \frac{1}{\pi}\text{Im}\int_{-\infty}^{\mu}d\omega'\frac{\Gamma(k,\omega')}{(\varepsilon_k-\omega'-i\delta)^2(\varepsilon_k+\omega-\omega'-i\delta)}\\ &= -\frac{1}{\pi}\frac{\partial}{\partial\omega''}\text{Im}\int_{-\infty}^{\mu}d\omega'\frac{\Gamma(k,\omega')}{(\omega''-\omega'-i\delta)(\varepsilon_k+\omega-\omega'-i\delta)}\Big|_{\omega''=\varepsilon_k}\\ &= -\frac{\partial}{\partial\omega''}\left\{\frac{\Gamma(k,\varepsilon_k+\omega)}{\omega''-\omega-\varepsilon_k}+\frac{\Gamma(k,\omega'')}{\varepsilon_k+\omega-\omega''}\right\}\Big|_{\omega''=\varepsilon_k}\\ &= \frac{\Gamma(k,\varepsilon_k+\omega)-\Gamma(k,\varepsilon_k)-\omega\Gamma'(k,\varepsilon_k)}{\omega^2} \qquad (125)\end{aligned}$$

As follows from equations 118 and 121, the quasiparticle energy in the cumulant expansion is essentially determined by E_k, which is the quasiparticle energy in the GWA.

Let us apply the cumulant method to the Hamiltonian

$$H=\varepsilon c^\dagger c+\omega_p b^\dagger b+gcc^\dagger\left(b^\dagger+b\right) \qquad (126)$$

which is a simplified version of the model Hamiltonian discussed previously. First we must calculate the self-energy to first order in the plasmon propagator

$$D(\omega)=\frac{1}{\omega-\omega_p+i\delta}-\frac{1}{\omega+\omega_p-i\delta} \qquad (127)$$

The effective interaction between the core electron and its surrounding is g^2D (see, e.g., Inkson [59]). The self-energy of the core electron as a result of the coupling to the plasmon is then

$$\begin{aligned} \Sigma(\omega) &= \frac{i}{2\pi}\int d\omega' G(\omega+\omega')\,g^2D(\omega')\\ &= \frac{ig^2}{2\pi}\int d\omega'\frac{1}{\omega+\omega'-\varepsilon-i\delta}\left\{\frac{1}{\omega'-\omega_p+i\delta}-\frac{1}{\omega'+\omega_p-i\delta}\right\}\\ &= \frac{g^2}{\omega+\omega_p-\varepsilon-i\delta} \qquad (128)\end{aligned}$$

The spectral function for Σ is given by

$$\Gamma(\omega)=g^2\delta(\omega-\varepsilon+\omega_p) \qquad (129)$$

and the self-energy correction to the core eigenvalue is

$$\Delta\varepsilon=\frac{g^2}{\omega_p} \qquad (130)$$

The derivative of Σ at $\omega = \varepsilon$ is

$$\left.\frac{\partial \Sigma}{\partial \omega}\right|_{\omega=\varepsilon} = -\left(\frac{g}{\omega_p}\right)^2 \tag{131}$$

$C^h(k,0) = 0$ and $C_S^h(k,t)$ becomes

$$\int_{-\infty}^{\mu} d\omega \frac{\mathrm{e}^{i(\varepsilon-\omega-i\delta)t}}{(\varepsilon-\omega-i\delta)^2}\Gamma(\omega) = \left(\frac{g}{\omega_p}\right)^2 \mathrm{e}^{i\omega_p t} \tag{132}$$

The spectral function is thus given by

$$\begin{aligned} A(\omega) &= \frac{1}{\pi}\mathrm{Im}\, i\int_{-\infty}^{0} dt\, \mathrm{e}^{i\omega t}\mathrm{e}^{-i\varepsilon t + C^h(t)} \\ &= \frac{1}{\pi}\mathrm{Im}\, i\mathrm{e}^{-(g/\omega_p)^2}\int_{-\infty}^{0} dt\, \mathrm{e}^{i(\omega-\varepsilon-\Delta\varepsilon)t}\sum_{n=0}^{\infty}\frac{[C^h(t)]^n}{n!} \\ &= \mathrm{e}^{-(g/\omega_p)^2}\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{g}{\omega_p}\right)^{2n}\delta(\omega-\varepsilon-\Delta\varepsilon+n\omega_p) \end{aligned} \tag{133}$$

which is precisely the exact solution [66].

5.1.2 Comparison between the cumulant expansion and the GWA

To identify vertex corrections contained in the cumulant expansion, we compare it with the GWA. Direct comparison in the self-energy diagrams is, however, difficult if not impossible. This is because the cumulant expansion is an expansion in the Green function, rather than in the self-energy. It is therefore more appropriate to compare the Green functions in the two approximations. In figure 10 the Green function diagrams are shown to second order in the screened interaction, which should be sufficient for our purpose. The cumulant expansion diagrams are obtained by considering the three possible time-orderings of the integration time variables t' in $C^2(k,t)$ with $C(k,t)$ given by equation (122). As can be seen in the figure, the cumulant expansion contains second-order diagrams which are not included in the GWA. It is these additional diagrams that give rise to the second plasmon satellite and they are quite distinct from the second-order diagram common to both approximations. The interpretation of the latter diagram is that a hole emits a plasmon which is reabsorbed at a later time and the hole returns to its original state before plasmon emission. This process is repeated once at a later time. Thus there is only one plasmon coupled to the hole at one time. In contrast, the other two diagrams, not contained in the GWA, describe an additional plasmon emission before the first one is reabsorbed, giving two plasmons coupled to the hole simultaneously. Similar consideration can be extended to the higher-order diagrams.

The cumulant expansion contains only boson-type diagrams describing emission and reabsorbtion of plasmons but it does not contain diagrams corresponding to interaction between a hole and particle-hole pairs. This type of interaction is described by the ladder diagrams. For this reason, the cumulant expansion primarily corrects the satellite description whereas the quasiparticle energies are to a large extent determined by the GWA as mentioned before.

5.1.3 Applications

The cumulant expansion was applied recently to calculate the photoemission spectra in Na and Al [13]. The experimental spectra consist of a quasiparticle peak with a set of plasmon satellites separated from the quasiparticle by multiples of the plasmon energy (figure 11a). The spectra in the GWA shows only one plasmon satellite located at a too high energy, approximately 1.5

ω_p below the quasiparticle (11b)) which is similar to the core electron case. The cumulant expansion method remedies this problem and yields spectra in good agreement with experiment regarding the position of the satellites as can be seen in Fig. 11b. The relative intensities of the satellites with respect to that of the quasiparticle are still in discrepancy. This is likely due to extrinsic effects corresponding to the interaction of the photoemitted electron with the bulk and the surface on its way out of the solid resulting in energy loss. These are not taken into account in the sudden approximation.

When applied to valence electrons with band dispersions the cumulant expansion does not yield the exact result anymore as in the core electron case. Surprisingly, the numerical results show that the cumulant expansion works well even in Al with a band width of $\approx$11 eV. Considering its simplicity, it is a promising approach for describing plasmon satellites.

5.2 T-matrix approximation

In many strongly correlated systems, such as transition metal oxides, the photoemission spectra often show a satellite structure a few eV below the main peak. The origin of this satellite is different from that of the plasmon-related satellite which is also found in sp-systems and which usually has a much higher energy. The additional satellite found in strongly correlated systems is due to the presence of two or more holes in a narrow band after a photoelectron is emitted, i.e. it is due to short-range rather than long-range correlations. An illustrative example is provided by Cu and Ni. In Cu, the $3d$-band is fully occupied so that after photoemission there is only one d-hole in the system corresponding to the $3d^9$ configuration resulting in no hole-hole correlation. Consequently there is no satellite either and a single-particle theory is sufficient to describe the electronic structure of Cu. On the other hand, the ground state of Ni already contains a configuration with one d-hole so that after photoemission, the final state contains a configuration with two d holes. Since the two holes are localized on the same atomic site, there will be a strong $d-d$ interaction resulting in the well-known 6 eV satellite which in the atomic picture corresponds to the energy of the $3d^8$ configuration.

So far there is no good *ab initio* scheme for dealing with short-range correlations. Most works have been based on model Hamiltonians which have given important insights into the underlying physics but which contain adjustable parameters, preventing direct quantitative comparison with experiment. The GWA takes into account long-range correlations through the RPA screening which determines to a large extent the quasiparticle energies. It is known that the GWA works well for quasiparticle energies but from a number of calculations [8, 11, 12] it is clear that the GWA has shortcomings in describing satellite structures. Even the plasmon-related satellites are not well described as discussed in the previous section. A natural extension of the GWA is to include short-range hole-hole correlations. This type of interactions seems to be suitably described by the T-matrix approximation [60]. Previous calculations based on the Hubbard model showed qualitatively that the T-matrix theory was capable of yielding a satellite structure in Ni [69, 70, 86]. Most other works in Ni have also been based on model Hamiltonians [104, 98, 24, 58]. T-matrix calculation on a two-dimensional Hubbard model is also found to improve the satellite description of the GWA [105]. In this section we develop a T-matrix theory for performing *ab initio* calculations on real systems [96, 97]. The method is applied to calculate the spectral function of Ni for which 6 eV satellite is not obtained in the GWA [8].

An extension of the T-matrix theory including Faddeev's three-body interaction [33] was made by Igarashi [56, 57] and by Calandra and Manghi [24, 80]. The theory has been applied to Ni [58] and NiO [80] within the Hubbard model.

5.2.1 T-matrix

The T-matrix is defined by the following Bethe-Salpeter equation:

$$\begin{aligned} T &= U + UKT \\ &= [1 - UK]^{-1}U \end{aligned} \tag{134}$$

where U is a screened Coulomb interaction and K is a two-particle propagator. Diagrammatically, the multiple scattering processes are shown in figure 12. It is evident from the figure that the T-matrix describes multiple scattering between two holes or electrons.

The full expression for equation (134) is

$$\begin{aligned} T_{\sigma\sigma'}(1,2|3,4) &= U(1-2)\delta(1-3)\delta(2-4) \\ &+ U(1-2)\int d1'd2'\, K_{\sigma\sigma'}(1,2|1',2')T_{\sigma\sigma'}(1',2'|3,4) \end{aligned} \tag{135}$$

where we have used a short-hand notation $1 \equiv (\mathbf{r}_1, t_1)$ and σ labels the spin. The kernel K or the two-particle propagator is given by

$$K_{\sigma\sigma'}(1,2|1',2') = iG_{\sigma}(1',1)G_{\sigma'}(2',2) \tag{136}$$

where G_σ is a time-ordered single-particle Green function

$$iG_{\sigma}(1',1) = \langle N|T[\hat{\phi}_{\sigma}(1')\hat{\phi}^{\dagger}_{\sigma}(1)]|N\rangle \tag{137}$$

We have assumed that U is an instantaneous interaction which means that $t_1 = t_2$, $t_3 = t_4$, and $t_{1'} = t_{2'}$. Without loss of generality, we may set $t_1 = t_2 = 0$ and $t_3 = t_4 = t$. Equation (135) then becomes

$$\begin{aligned} &T_{\sigma\sigma'}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}_3, \mathbf{r}_4; t) = \\ &\quad U(\mathbf{r}_1 - \mathbf{r}_2)\delta(\mathbf{r}_1 - \mathbf{r}_3)\delta(\mathbf{r}_2 - \mathbf{r}_4)\delta(t) \\ &\quad + U(\mathbf{r}_1 - \mathbf{r}_2)\int d^3r'_1 d^3r'_2 \int dt' K_{\sigma\sigma'}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}'_1, \mathbf{r}'_2; t') \\ &\qquad\qquad \times T_{\sigma\sigma'}(\mathbf{r}'_1, \mathbf{r}'_2|\mathbf{r}_3, \mathbf{r}_4; t - t') \end{aligned} \tag{138}$$

Fourier transformation of the above equation yields

$$\begin{aligned} &T_{\sigma\sigma'}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}_3, \mathbf{r}_4; \omega) = \\ &\quad U(\mathbf{r}_1 - \mathbf{r}_2)\delta(\mathbf{r}_1 - \mathbf{r}_3)\delta(\mathbf{r}_2 - \mathbf{r}_4) \\ &\quad + U(\mathbf{r}_1 - \mathbf{r}_2)\int d^3r'_1 d^3r'_2 \int K_{\sigma\sigma'}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}'_1, \mathbf{r}'_2; \omega) \\ &\qquad\qquad \times T_{\sigma\sigma'}(\mathbf{r}'_1, \mathbf{r}'_2|\mathbf{r}_3, \mathbf{r}_4; \omega) \end{aligned} \tag{139}$$

where

$$K_{\sigma\sigma'}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}'_1, \mathbf{r}'_2; \omega) = i\int \frac{d\omega'}{2\pi} G_{\sigma}(\mathbf{r}'_1, \mathbf{r}_1; \omega - \omega')G_{\sigma'}(\mathbf{r}'_2, \mathbf{r}_2; \omega') \tag{140}$$

The Fourier transforms are defined by

$$F(\omega) = \int_{-\infty}^{\infty} dt\, e^{-i\omega t} f(t) \tag{141}$$

$$f(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, e^{i\omega t} F(\omega) \tag{142}$$

Using the spectral representation of G_σ

$$G_\sigma(\omega) = \int_{-\infty}^{\mu} d\omega' \frac{A_\sigma(\omega')}{\omega - \omega' - i\delta} + \int_{\mu}^{\infty} d\omega' \frac{A_\sigma(\omega')}{\omega - \omega' + i\delta} \tag{143}$$

with a non-interacting A_σ

$$A_\sigma(\mathbf{r}, \mathbf{r}'; \omega) = \sum_i \psi_{i\sigma}(\mathbf{r}) \psi^*_{i\sigma}(\mathbf{r}') \delta(\omega - \varepsilon_{i\sigma}) \tag{144}$$

an explicit form for the kernel K is given by

$$\begin{aligned} K_{\sigma\sigma'}(\mathbf{r}_1, \mathbf{r}_2 | \mathbf{r}'_1, \mathbf{r}'_2; \omega) = \\ - \sum_{ij}^{\text{occ}} \frac{\psi_{i\sigma}(\mathbf{r}'_1) \psi^*_{i\sigma}(\mathbf{r}_1) \psi_{j\sigma'}(\mathbf{r}'_2) \psi^*_{j\sigma'}(\mathbf{r}_2)}{\omega - \varepsilon_{i\sigma} - \varepsilon_{j\sigma'} - i\delta} \\ + \sum_{ij}^{\text{unocc}} \frac{\psi_{i\sigma}(\mathbf{r}'_1) \psi^*_{i\sigma}(\mathbf{r}_1) \psi_{j\sigma'}(\mathbf{r}'_2) \psi^*_{j\sigma'}(\mathbf{r}_2)}{\omega - \varepsilon_{i\sigma} - \varepsilon_{j\sigma'} + i\delta} \end{aligned} \tag{145}$$

The first term on the right hand side is due to hole-hole scattering and the second to particle-particle scattering. This expression is similar to the RPA polarization propagator but the states are either both occupied or unoccupied.

5.2.2 T-matrix self-energy

The self-energy can be obtained from the T-matrix which consists of the direct term

$$\Sigma^d(4,2) = -i \int d1 d3 G(1,3) T(1,2|3,4) \tag{146}$$

and the exchange term

$$\Sigma^x(3,2) = i \int d1 d4 G(1,4) T(1,2|3,4) \tag{147}$$

Fourier transformation gives

$$\begin{aligned} \Sigma^d_\sigma(\mathbf{r}_4, \mathbf{r}_2; \omega) = -i \int d^3 r_1 d^3 r_3 \\ \int \frac{d\omega'}{2\pi} \sum_{\sigma'} G_{\sigma'}(\mathbf{r}_1, \mathbf{r}_3; \omega' - \omega) T_{\sigma'\sigma}(\mathbf{r}_1, \mathbf{r}_2 | \mathbf{r}_3, \mathbf{r}_4; \omega') \end{aligned} \tag{148}$$

$$\begin{aligned} \Sigma^x_\sigma(\mathbf{r}_3, \mathbf{r}_2; \omega) = i \int d^3 r_1 d^3 r_4 \\ \int \frac{d\omega'}{2\pi} G_\sigma(\mathbf{r}_1, \mathbf{r}_4; \omega' - \omega) T_{\sigma\sigma}(\mathbf{r}_1, \mathbf{r}_2 | \mathbf{r}_3, \mathbf{r}_4; \omega') \end{aligned} \tag{149}$$

We note that for the exchange term there is no summation over the spin since exchange between particles of opposite spins is zero.

The spectral representation of $T_{\sigma\sigma'}$ is given by

$$T_{\sigma\sigma'}(\omega) = \int_{-\infty}^{2\mu} d\omega' \frac{Q_{\sigma\sigma'}(\omega')}{\omega - \omega' - i\delta} + \int_{2\mu}^{\infty} d\omega' \frac{Q_{\sigma\sigma'}(\omega')}{\omega - \omega' + i\delta} \tag{150}$$

where

$$Q_{\sigma\sigma'}(\omega) = -\frac{1}{\pi} \text{Im}\, T_{\sigma\sigma'}(\omega)\, \text{sgn}(\omega - 2\mu) \tag{151}$$

The analytic structure of T is determined by K in equation (145) which can be seen by considering equation (134). The first order term in the T-matrix is the Hartree-Fock term which is independent of frequency and therefore it does not influence the analytic structure. The T-matrix without the first order term is then given by

$$\begin{aligned} T &= UKU + UKT \\ &= [1 - UK]^{-1} UKU \end{aligned} \tag{152}$$

Thus, the analytic structure is determined by K because it can be shown that $[1 - UK]^{-1}$ has no poles in the first and third quadrants (the poles of K lie in the second and fourth quadrants).

Using the spectral representations of G and T the self-energy can be written explicitly as

$$\begin{aligned} &\text{Im}\, \Sigma^d_\sigma(\mathbf{r}_4, \mathbf{r}_2; \omega > \mu) = \\ &\quad \int d^3r_1 d^3r_3 \sum_{\mathbf{k}'n'\sigma'}^{\text{occ}} \psi_{\mathbf{k}'n'\sigma'}(\mathbf{r}_1) \psi^*_{\mathbf{k}'n'\sigma'}(\mathbf{r}_3) \\ &\times \quad \text{Im}\, T_{\sigma'\sigma}(\mathbf{r}_1, \mathbf{r}_2 | \mathbf{r}_3, \mathbf{r}_4; \omega + \varepsilon_{\mathbf{k}'n'\sigma'})\, \theta(\omega + \varepsilon_{\mathbf{k}'n'\sigma'} - 2\mu) \end{aligned} \tag{153}$$

$$\begin{aligned} &\text{Im}\, \Sigma^d_\sigma(\mathbf{r}_4, \mathbf{r}_2; \omega \leq \mu) = \\ &\quad - \int d^3r_1 d^3r_3 \sum_{\mathbf{k}'n'\sigma'}^{\text{unocc}} \psi_{\mathbf{k}'n'\sigma'}(\mathbf{r}_1) \psi^*_{\mathbf{k}'n'\sigma'}(\mathbf{r}_3) \\ &\times \quad \text{Im}\, T_{\sigma'\sigma}(\mathbf{r}_1, \mathbf{r}_2 | \mathbf{r}_3, \mathbf{r}_4; \omega + \varepsilon_{\mathbf{k}'n'\sigma'})\, \theta(-\omega - \varepsilon_{\mathbf{k}'n'\sigma'} + 2\mu) \end{aligned} \tag{154}$$

A similar expression for the exchange part can be easily derived by interchanging $\mathbf{r_3}$ and $\mathbf{r_4}$ in the T-matrix. The real part of Σ is obtained from the spectral representation

$$\Sigma_\sigma(\omega) = \int_{-\infty}^{\mu} d\omega' \frac{\Gamma_\sigma(\omega')}{\omega - \omega' - i\delta} + \int_{\mu}^{\infty} d\omega' \frac{\Gamma_\sigma(\omega')}{\omega - \omega' + i\delta} \tag{155}$$

where

$$\Gamma_\sigma(\omega) = -\frac{1}{\pi} \text{Im}\, \Sigma_\sigma(\omega)\, \text{sgn}(\omega - \mu) \tag{156}$$

The screened potential U is in general frequency dependent. For a narrow band of width Δ, the time scale for a hole to hop from one site to a neighbouring site is determined by $1/\Delta$ which is large and in frequency space it means that the largest contribution to the T-matrix comes from $\omega \approx 0$. Physically, the interaction is essentially instantaneous within the same site and therefore it is justified to use a static screened interaction. This static approximation will be used in the present work.

The one- electron spectral functions are obtained from

$$A_i(\omega) = \frac{1}{\pi} \frac{|\text{Re}\, \Sigma_i(\omega)|}{|\omega - \varepsilon_i - \text{Im}\, \Sigma_i(\omega)|^2 + |\text{Re}\, \Sigma_i(\omega)|^2} \tag{157}$$

with the shorthand notation $i \equiv \mathbf{k}n\sigma$. We have assumed in the above expression that the self-energy is diagonal in the LDA wavefunctions which are used to construct G.

5.2.3 Double counting and the total self-energy

To calculate the total self-energy, we add the T-matrix self-energy to the GW self-energy. However, this leads to double counting since the second order term in the T-matrix is already included in the GW self-energy. A straightforward subtraction of the second order term leads however to wrong analytic properties of the self-energy and consequently gives some negative

spectral weight. To solve this double-counting problem, we proceed as follows. We divide the bare Coulomb interaction into the short-range screened potential U and a long-range part v_L ((von Barth 1995):

$$v = U + v_L \tag{158}$$

The correlation part of the GW self-energy is schematically given by

$$\Sigma^c_{GW} = GvRv \tag{159}$$

where R is the total RPA response function

$$R = (1 - Pv)^{-1}P \tag{160}$$

and P is the RPA polarization function. Using the above division of the Coulomb potential we obtain

$$\Sigma^c_{GW} = GURv_L + Gv_LRU + Gv_LRv_L + GURU \tag{161}$$

The last term is then subtracted out to avoid double counting. It has the same analytic structure as the GW self-energy, but since U is smaller than the bare v, this term is always smaller than the GW self-energy for all frequencies. This guarantees that the resulting self-energy has the correct analytic properties as in the GWA. Numerically this term turns out to be very small.

Thus, according to the above scheme the total correlated self-energy becomes:

$$\Sigma^c_{GWT} = \Sigma^c_{GW} + \Sigma^c_T - GURU \tag{162}$$

where Σ^c_T is given by

$$\Sigma^c_T = \Sigma^d + \Sigma^x \tag{163}$$

with Σ^d and Σ^x given by equations (148) and (149) respectively.

5.2.4 Numerical procedure

Since the T-matrix describes scattering between localized holes or particles, it is suitable to work with basis functions which are also localized such as the linear muffin-tin orbitals (LMTO). In the atomic sphere approximation, the LMTO basis functions are of the form (central cell)

$$\chi^\sigma_{RL}(\mathbf{r},\mathbf{k}) = \varphi^\sigma_{RL}(\mathbf{r}) + \sum_{R'L'} \dot{\varphi}^\sigma_{R'L'}(\mathbf{r})\, h^\sigma_{R'L',RL}(\mathbf{k}) \tag{164}$$

ϕ^σ_{RL} is a solution to the Schrödinger equation in atomic sphere R at an energy in the center of the L-band region and $\dot{\phi}^\sigma_{RL}$ is the corresponding energy derivative. $h^\sigma_{R'L',RL}$ is a constant matrix depending on the crystal structure as well as the potential. The Bloch states are expanded in the LMTO basis.

$$\psi_{\mathbf{k}n\sigma}(\mathbf{r}) = \sum_{RL} \chi^\sigma_{RL}(\mathbf{r},\mathbf{k})\, b_{RL}(\mathbf{k}n\sigma) \tag{165}$$

Inserting $\psi_{\mathbf{k}n\sigma}$ in equation (145), we obtain

$$\begin{aligned} K_{\sigma\sigma'}(\mathbf{r}_1,\mathbf{r}_2|\mathbf{r'}_1,\mathbf{r'}_2;\omega) = & -\sum_{\mathbf{k}n,\mathbf{k}'n'}^{\text{occ}} \sum_{R_1L_1} \sum_{R_2L_2} \sum_{R_3L_3} \sum_{R_4L_4} \\ & \chi^\sigma_{R_1L_1}(\mathbf{r'}_1,\mathbf{k})\chi^{\sigma *}_{R_2L_2}(\mathbf{r}_1,\mathbf{k})\chi^{\sigma'}_{R_3L_3}(\mathbf{r'}_2,\mathbf{k}')\chi^{\sigma' *}_{R_4L_4}(\mathbf{r}_2,\mathbf{k}') \\ & \times \frac{b_{R_1L_1}(\mathbf{k}n\sigma)b^*_{R_2L_2}(\mathbf{k}n\sigma)b_{R_3L_3}(\mathbf{k}'n'\sigma')b^*_{R_4L_4}(\mathbf{k}'n'\sigma')}{\omega - \varepsilon_{\mathbf{k}n\sigma} - \varepsilon_{\mathbf{k}'n'\sigma'} - i\delta} \\ & + \text{unoccupied term} \end{aligned} \tag{166}$$

So far the above expression is exact. The largest contribution to K arises from the onsite terms, i.e. within the same unit cell. Furthermore, since U is short-range, the largest contribution to the T-matrix also arises from the onsite terms. For practical purpose, we then neglect the $\mathbf{k}$ dependence of χ so that

$$K_{\sigma\sigma'}(\mathbf{r}_1,\mathbf{r}_2|\mathbf{r'}_1,\mathbf{r'}_2;\omega) =$$

$$\sum_{\alpha\beta\gamma\delta} \varphi^{\sigma}_{\alpha}(\mathbf{r'}_1)\varphi^{\sigma *}_{\beta}(\mathbf{r}_1)\; K_{\sigma\sigma'}(\alpha\beta|\gamma\delta;\;\omega)\; \varphi^{\sigma'}_{\gamma}(\mathbf{r'}_2)\varphi^{\sigma' *}_{\delta}(\mathbf{r}_2) \tag{167}$$

where

$$K_{\sigma\sigma'}(\alpha\beta|\gamma\delta;\;\omega) \;=\; -\sum_{\mathbf{k}n,\mathbf{k'}n'}^{\mathrm{occ}} \frac{b_{\alpha}(\mathbf{k}n\sigma)b^{*}_{\beta}(\mathbf{k}n\sigma)b_{\gamma}(\mathbf{k'}n'\sigma')b^{*}_{\delta}(\mathbf{k'}n'\sigma')}{\omega-\varepsilon_{\mathbf{k}n\sigma}-\varepsilon_{\mathbf{k'}n'\sigma'}-i\delta}$$

$$+\text{unoccupied term} \tag{168}$$

We have used a short-hand notation $\alpha \equiv RL$

Using this expression for K in equation (139) and taking matrix element with respect to $\varphi^{\sigma}_{R_1L_1}(\mathbf{r}_1)\;\varphi^{\sigma}_{R_2L_2}(\mathbf{r}_2)$ and $\varphi^{\sigma'}_{R_3L_3}(\mathbf{r}_3)\;\varphi^{\sigma'}_{R_4L_4}(\mathbf{r}_4)$ we obtain a matrix equation for T

$$T_{\sigma\sigma'}(\alpha\beta|\gamma\delta;\;\omega) = U(\alpha\beta|\gamma\delta)$$

$$+\sum_{\eta\nu,\lambda\rho} U(\alpha\beta|\eta\nu)\; K_{\sigma\sigma'}(\eta\nu|\lambda\rho;\;\omega)\; T_{\sigma\sigma'}(\lambda\rho|\gamma\delta;\;\omega) \tag{169}$$

which can be easily solved by inversion.

$$U(\alpha\beta|\eta\nu) = \int d1d2\varphi_{\alpha}(1)\varphi_{\beta}(2)v(1-2)\varphi_{\eta}(1)\varphi_{\nu}(2) \tag{170}$$

We note that although U depends on the spin through the basis, it has no physical spin dependence.

Having obtained T, it is straightforward to calculate the self-energy. From equations (153) and equations (154), the matrix element of the self-energy in a Bloch state $\psi_{\mathbf{k}n\sigma}$ is given by

$$\mathrm{Im}\;\Sigma^{d}_{\sigma}(\mathbf{k}n;\omega>\mu) =$$

$$\sum_{\mathbf{k'}n'\sigma'}^{\mathrm{occ}}\sum_{\alpha\beta,\gamma\delta} b_{\alpha}(\mathbf{k}n\sigma)b^{*}_{\beta}(\mathbf{k}n\sigma)b_{\gamma}(\mathbf{k'}n'\sigma')b^{*}_{\delta}(\mathbf{k'}n'\sigma')$$

$$\times\;\; \mathrm{Im}\;T_{\sigma'\sigma}(\alpha\beta|\gamma\delta;\;\omega+\varepsilon_{\mathbf{k'}n'\sigma'})\;\theta(\omega+\varepsilon_{\mathbf{k'}n'\sigma'}-2\mu) \tag{171}$$

$$\mathrm{Im}\;\Sigma^{d}_{\sigma}(\mathbf{k}n;\omega\le\mu) =$$

$$-\sum_{\mathbf{k'}n'\sigma'}^{\mathrm{unocc}}\sum_{\alpha\beta,\gamma\delta} b_{\alpha}(\mathbf{k}n\sigma)b^{*}_{\beta}(\mathbf{k}n\sigma)b_{\gamma}(\mathbf{k'}n'\sigma')b^{*}_{\delta}(\mathbf{k'}n'\sigma')$$

$$\times\;\; \mathrm{Im}\;T_{\sigma'\sigma}(\alpha\beta|\gamma\delta;\;\omega+\varepsilon_{\mathbf{k'}n'\sigma'})\;\theta(-\omega-\varepsilon_{\mathbf{k'}n'\sigma'}+2\mu) \tag{172}$$

The real part of Σ is obtained from equation (155). It is straightforward to derive the corresponding expressions for the exchange part.

5.2.5 Application to Ni with a model screened interaction

As a first application, we use the T-matrix theory to calculate the self-energy of ferromagnetic Ni. In the case of Ni, the T-matrix should describe repeated scattering of two 3d holes on the same site, corresponding to the atomic $3d^8$ configuration. The mechanism for the origin of the 6 eV satellite is discussed in the previous section on Ni.

A model potential of the form

$$U(|\mathbf{r}-\mathbf{r}'|) = \frac{\mathrm{erfc}(\alpha|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} \tag{173}$$

is used to calculate the T-matrix. For the T-matrix self-energy, the position of the peaks in Im Σ is rather insensitive to the screening parameter. However, the intensity varies with α and consequently the satellite position, which is determined by Re Σ, is modified accordingly. Thus, if we were allowed to adjust the screening parameter α, the position of the satellite could be shifted arbitrarily. It is therefore crucial to determine the screened interaction U from a parameter-free scheme. According to constrained LDA calculations [42] $U_{dd} \approx 5.5$ eV [98]. Consequently in our calculations we choose $\alpha = 1.2$ which is equivalent to $U_{dd} \approx 5.5$ eV.

The calculated spectral functions are displayed in figure 13 and compared with the GW spectra. The main peak ≈ 3 eV below the Fermi level is the quasiparticle peak. Two satellite structures originating from the T-matrix self-energy, absent in the GWA, can be observed below the main peak and just above the Fermi level. The position of the satellite below the Fermi level is, however, somewhat too low. This is likely due to the difficulties in determining the correct screened interaction in the T-matrix and the neglect of particle-hole interaction in the present scheme.

A new interesting feature is the presence of a peak structure just above the Fermi energy which arises from particle-particle scattering. At first sight, these scattering processes are expected to be insignificant, since the number of unoccupied states is small which leads to a small T-matrix for positive energies. However, our results point to the importance of matrix element and bandstructure effects, usually neglected in the Hubbard model. As may be seen by an examination of equation (153) there is a sum over occupied states which amplifies the small contribution from the T-matrix. It may be possible to measure the satellite structure above the Fermi level in an angular resolved inverse photoemission experiment by choosing certain k-vectors, where the quasiparticle peak is well separated from the satellite.

There is a significant difference between the majority and minority self-energy as reflected in the spectral function in figure 13. The probability of creating a hole in the majority channel is larger than in the minority channel, since the former is fully occupied. The virtual hole can mainly be created in the minority channel due to the presence of $3d$ unoccupied states just above the Fermi level. This implies that the majority $3d$ self-energy will be larger than the minority one which reduces the exchange splitting and the bandwidth, improving the GW result. It is found that the T-matrix self-energy reduces the exchange splitting by ≈ 0.3 eV [96, 97] for states at the bottom of the 3d band and thus also improving the GW bandwidth. But contrary to model calculations [69, 70], the T-matrix has a significantly smaller effect on the bandwidth.

6 Self-consistency and total energies

6.1 Self-consistency

The current practice of performing GW calculations is to choose a non-interacting Hamiltonian (usually the LDA) and to construct the Green function G_0 from the eigenstates of the Hamiltonian and to calculate the response function. These are then used to calculate the self-energy

which is taken as the final result. This procedure is non-self-consistent since the interacting Green function G obtained from the Dyson equation $G = G_0 + G_0\Sigma_0 G$ is not necessarily the same as G_0. To achieve self-consistency, the interacting Green function should be used to form a new polarization function $P = -iGG$, a new screened interaction W, and a new self-energy Σ. This process is repeated until G obtained from the Dyson equation is the same as G used to calculate the self-energy. Self-consistency guarantees that the final result is independent of the starting Green function G_0. Moreover, a self-consistent GW scheme ensures conservation of particle number and energy. This is a consequence of a general theorem due to Baym and Kadanoff [14] and Baym [15] who proved that approximations for Σ are conserving if Σ is obtained as a functional derivative of an energy functional Φ with respect to G :

$$\Sigma = \frac{\delta\Phi[G]}{\delta G} \tag{174}$$

Conservation of particle number means that the continuity equation

$$-\partial_t n(\mathbf{r},t) = \nabla\cdot\mathbf{j}(\mathbf{r},t) \tag{175}$$

is satisfied when n and $\mathbf{j}$ is obtained from the self-consistent Green function. Furthermore

$$N = \frac{1}{\pi} tr \int_{-\infty}^{\mu} d\omega \,\mathrm{Im}\, G(\omega) \tag{176}$$

gives the correct total number of particles. Conservation of energy means that the energy change when an external potential is applied to the system is equal to the work done by the system against the external potential when calculated using the self-consistent G.

The first self-consistent calculation GW calculation was probably by [29] for a model quasi-one dimensional semiconducting wire. The relevance of this model to real solids is, however, unclear. Self-consistent calculations for the electron gas were performed recently by von Barth and Holm [106] and by Shirley [91]. The calculations were done for two cases: in the first case the screened interaction W is fixed at the RPA level, $W = W_0$, (calculated using the non-interacting G_0) and only the Green function is allowed to vary to self-consistency and in the other case both G and W are allowed to vary (full self-consistency case) [54]. The results of these studies are

- The band width is increased from its non-self-consistent value, worsening the agreement with experiment (figure 14).
- The weight of the quasiparticles is increased, reducing the weight in the plasmon satellite.
- The quasiparticles are slightly narrowed, increasing their life-time.
- The plasmon satellite is broadened and shifted towards the Fermi level (figure 15). In the full self-consistent case, the plasmon satellite almost disappears (figure 16).

The main effects of self-consistency are mainly due to allowing the quasiparticle weight Z to vary. The increase in band width is disturbing and can be understood as follows. We consider the first case with fixed $W = W_0$ for simplicity. First we note that the GW result for the band width after one iteration is close to the free electron one. This means that there is almost a complete cancellation between exchange and correlation. After one iteration the quasiparticle weight is reduced to typically 0.7 and the rest of the weight goes to the plasmon satellite. The new Im Σ, calculated from G obtained from the first iteration has its weight reduced at low energy and increased at high energy compared to the non-self-consistent result (figure 17). This is due to the sum-rule [106]

$$\int_{-\infty}^{\infty} d\omega |\mathrm{Im}\Sigma^c(\mathbf{k},\omega)| = \sum_{\mathbf{q}} \int_0^{\infty} d\omega |\mathrm{Im}W(\mathbf{q},\omega)| \tag{177}$$

which shows that the left hand side is a constant depending only on the prechosen W_0 but independent of $\mathbf{k}$ and self-consistency. For a state at the Fermi level, self-consistency has little effects since $\mathrm{Im}\,\Sigma$ has almost equal weights for the hole ($\omega \leq \mu$) and the particle part ($\omega > \mu$) which cancel each other when calculating $\mathrm{Re}\,\Sigma_c$, as can be seen in equation (66) and as illustrated in figure 18. But for the state at the bottom of the valence band, $\mathrm{Im}\,\Sigma$ has most of its weight in the hole part (see e.g. figure 6) so that the shifting of the weight in $\mathrm{Im}\,\Sigma$ to higher energy causes $\mathrm{Re}\,\Sigma_c$ to be less positive than its non-self-consistent value. A similar effect is found for the exchange part which becomes less negative but because the bare Coulomb interaction has no frequency dependence, the renormalization factor has a smaller effect on Σ^x so that the reduction in Σ^x is less than the reduction in Σ^c. The net effect is then an increase in the band width. The shifting of the weight in $\mathrm{Im}\Sigma$ to higher energy has immediate consequences of increasing the life-time and the renormalization weight Z (through a decrease in $\partial \mathrm{Re}\Sigma_c/\partial\omega|$) of of the quasiparticles and of broadening the plasmon satellite, compared to the results of one iteration.

When W is allowed to vary (full self-consistency) the band width becomes even more widened and the plasmon satellite becomes broad and featureless, in contradiction to experiment. The quasiparticle weight is increased further [54]. These can be traced back to the disappearance of a well-defined plasmon excitation in W. The reason for this is that the quantity $P = -iGG$ no longer has a physical meaning of a response function, rather it is an auxiliary quantity needed to construct W. Indeed, it does not satisfy the usual $f-$ sum rule. The equation $\epsilon(\mathbf{q}, \omega_p) = 0$ determining the plasmon energy is not satisfied any more since the Green function now always has weight around $\omega = \omega_p$. This has the effect of transferring the weight in $\mathrm{Im}\,\Sigma$ even further to higher energy with the consequences discussed in the previous paragraph.

A self-consistent calculation for the electron gas within the cumulant expansion has been performed by [53]. The result for the quasiparticle energy is similar to the self-consistent GWA since the quasiparticle energy in the cumulant expansion is essentially determined by that of the GWA as discussed in the cumulant expansion section. The satellite part is little affected by self-consistency.

6.2 Total energies

So far, the GWA has been applied mainly to calculate single-particle excitation spectra, but it is also possible to calculate the total energy [54, 35]. Holm and von Barth [54] calculated the total energy using the self-consistent Green function and self-energy in the Galitskii-Migdal formulation [38]:

$$E = -\frac{i}{2}\int d^3\dot{r} \lim_{t'\to t^+}\lim_{\mathbf{r}'\to\mathbf{r}}\left[i\frac{\partial}{\partial t} + h_0(\mathbf{r})\right] G(\mathbf{r}t, \mathbf{r}'t') \tag{178}$$

where h_0 is the kinetic energy operator plus a local external potential. The total energy turns out to be very accurate in comparison with the quantum Monte Carlo (QMC) results of Ceperley and Alder [25]. For $r_s = 2$ and 4 QMC gives 0.004 and -0.155 Rydberg respectively while self-consistent GW gives 0.005 and -0.156 Rydberg. This is rather surprising since the GWA represents only the first term in the self-energy expansion. The reason for the accurate results is not fully understood. Applications to other more realistic systems are necessary to show if the results are of general nature. It is probably related to the fact that the self-consistent GW scheme is energy conserving and it is partly explained by consideration of the so called Luttinger-Ward energy functional [74]:

$$\Omega[G] = \mathrm{Tr}\left\{\ln(\Sigma[G] - G_0^{-1}) + \Sigma[G]G\right\} - \Phi[G] \tag{179}$$

which is variational with respect to G and it is stationary when G is equal to the self-consistent G which obeys the Dyson equation:

$$\frac{\delta\Omega}{\delta G} = 0 \;\text{ when }\; G = G_0 + G_0\Sigma G \tag{180}$$

$\Omega = E - \mu N$ is the grand canonical potential whose stationary value corresponds to the energy (minus μN) obtained from the Galitskii-Migdal formula and Φ is an energy functional consisting of irreducible energy diagrams. It is not clear, however, why the first order energy diagram (giving the GW self-energy upon taking functional derivative with respect to G) appears to represent a very good energy functional. Furthermore, the chemical potential calculated from $\mu = \partial E/\partial N$ is in agreement with the value obtained from $\mu = k_F^2/2 + \Sigma(k_F, 0)$. The particle density $n = 2\sum_{\mathbf{k}} \int_{-\infty}^{\mu} d\omega A(\mathbf{k}, \omega)$ yields $n = k_F^3/(3\pi^2)$, i.e. particle number is conserved, as proven by Baym. It can also be shown that with a fixed $W = W_0$ particle number is also conserved [54]. The conclusion is that fully self-consistent GW calculations for quasiparticle energies should be avoided. It is more fruitful to construct vertex corrections (beyond GW) or to perform partially self-consistent GW calculations where the choice of G and W is physically motivated. For instance, one can fix W at the RPA level and modify G by using quasiparticle energies but keeping the renormalization factor equal to one. Or one could choose a single-particle Hamiltonian such that the GW quasiparticle energies are consistent with the single-particle eigenvalues.

Recently, Almbladh and Hindgren [51] discovered that the GW total energies of the electron gas calculated using the Luttinger-Ward formalism [74] are in very good agreement with the QMC results of Ceperley and Alder and the self-consistent results of Holm and von Barth. Due to the variational nature of the Luttinger-Ward functional, the Green function need not be the self-consistent one. A generalization of the Luttinger-Ward formalism by treating both G and W as independent parameters has also been developed [2]. Using the non-interacting $G = G_0$ and the plasmon-pole approximation for W gives almost as good results [51] as the self-consistent results of Holm and von Barth [54] or the QMC results of Ceperley and Alder [25]. These results are very encouraging and applications to real inhomogeneous systems are now in progress to test the validity of the scheme.

7 Computational method

The calculation of the self-energy involves the calculations of the following quantities:

1. A self-consistent bandstructure which is the input to the self-energy calculation. In principle we may use any non-interacting Hamiltonian but we use the LDA in practice.
2. The bare Coulomb matrix v.
3. The polarization function P
4. The response function R and the screened Coulomb matrix W.
5. The self-energy Σ.

7.1 Basis functions

To construct a minimal basis, let us consider the polarization P. Due to the symmetry of P with respect to a lattice translation $\mathbf{T}$,

$$P(\mathbf{r} + \mathbf{T}, \mathbf{r}' + \mathbf{T}, \omega) = P(\mathbf{r}, \mathbf{r}', \omega), \tag{181}$$

it follows that P may be expanded as follows:

$$P(\mathbf{r}, \mathbf{r}', \omega) = \sum_{\mathbf{k}ij} B_{\mathbf{k}i}(\mathbf{r}) P_{ij}(\mathbf{k}, \omega) B^*_{\mathbf{k}j}(\mathbf{r}') \tag{182}$$

where $B_{\mathbf{k}i}$ are any basis functions satisfying Bloch's theorem and they are normalized to unity in the unit cell with volume Ω. All other quantities depending on two space variables such as G, W, and Σ can be expanded in a similar fashion. The polarization matrix P_{ij} is given by

$$\begin{aligned} P_{ij}(\mathbf{k},\omega) &= \int d^3r \int d^3r'\, B^*_{\mathbf{k}i}(\mathbf{r}) P(\mathbf{r},\mathbf{r}',\omega) B_{\mathbf{k}j}(\mathbf{r}') \\ &= \sum_{\mathbf{T},\mathbf{T}'} \int_\Omega d^3r \int_\Omega d^3r'\, B^*_{\mathbf{k}i}(\mathbf{r}+\mathbf{T}) P(\mathbf{r}+\mathbf{T},\mathbf{r}'+\mathbf{T}',\omega) B_{\mathbf{k}j}(\mathbf{r}'+\mathbf{T}') \\ &= \sum_{\mathbf{T},\mathbf{T}'} \int_\Omega d^3r \int_\Omega d^3r'\, e^{-i\mathbf{k}\cdot(\mathbf{T}-\mathbf{T}')} B^*_{\mathbf{k}i}(\mathbf{r}) P(\mathbf{r},\mathbf{r}'+\mathbf{T}'-\mathbf{T},\omega) B_{\mathbf{k}j}(\mathbf{r}') \\ &= N^2 \int_\Omega d^3r \int_\Omega d^3r'\, B^*_{\mathbf{k}i}(\mathbf{r}) P(\mathbf{r},\mathbf{r}',\omega) B_{\mathbf{k}j}(\mathbf{r}') \end{aligned} \tag{183}$$

Thus the calculation of P for the crystal is reduced to a unit cell. ¿From Eq. (76) we then get

$$\begin{aligned} S^0_{ij}(\mathbf{q},\omega) &= \sum_{\mathbf{k}} \sum_{n\le\mu} \sum_{n'>\mu} \langle B_{\mathbf{q}i}\psi_{\mathbf{k}n}|\psi_{\mathbf{k}+\mathbf{q},n'}\rangle\langle\psi_{\mathbf{k}+\mathbf{q},n'}|\psi_{\mathbf{k}n}B_{\mathbf{q}j}\rangle \\ &\qquad \times\delta[\omega-(\varepsilon_{\mathbf{k}+\mathbf{q},n'}-\varepsilon_{\mathbf{k}n})] \end{aligned} \tag{184}$$

In the above expression, the wave function $\psi_{\mathbf{k}n}$ is normalized to unity in the unit cell. The matrix elements are given by

$$\langle\psi_{\mathbf{k}+\mathbf{q}n'}|\psi_{\mathbf{k}n}B_{\mathbf{q}j}\rangle = \int_\Omega d^3r\, \psi^*_{\mathbf{k}+\mathbf{q}n'}\psi_{\mathbf{k}n}B_{\mathbf{q}j} \tag{185}$$

The problem is to choose a minimal number of basis functions needed to describe the response function. It is clear from the above expression that that the basis functions must span the space formed by products of the wave functions.

If the wave functions are expanded in plane waves, as in pseudopotential theory, then the basis functions will be products of plane waves which are also plane waves, in which case

$$B_{\mathbf{k}j} \Longrightarrow \frac{exp[i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}]}{\sqrt{\Omega}}, \quad S^0_{ij} \Longrightarrow S^0_{\mathbf{G}\mathbf{G}'}$$

One simply needs to have a large enough $\mathbf{G}$ vector in order to have a complete basis. The advantages of plane-wave basis are that matrix elements can be easily calculated and the Coulomb matrix is simple since it is diagonal, $v_{\mathbf{G}\mathbf{G}'}(\mathbf{k}) = 4\pi\delta_{\mathbf{G}\mathbf{G}'}/|\mathbf{k}+\mathbf{G}|^2$. Other advantages are good control over convergence and programming ease. There are, however, serious drawbacks:

1. It is not feasible to do all electron calculations. In many cases, it is essential to include core electrons. For example, the exchange of a $3d$ valence state with the $3s-3p$ core states in the late $3d$ metals is overestimated by the LDA by as much as 1 eV which would lead to an error of the same order in the pseudopotential method.

2. The size of the response matrix becomes prohibitively large for narrow band systems due to a large number of plane waves.

Moreover, the plane waves have no direct physical interpretation.

To overcome these drawbacks, we use the LMTO method which allows us to treat any system. The LMTO method uses a minimal number of basis functions and we carry over the concept of minimal basis in bandstructures to the dielectric matrix ϵ. Instead of a planewave basis, we use a "product basis" which consists of products of LMTO's. As will be clear later, the product basis constitutes a minimal basis for ϵ within the LMTO formalism. A method for inverting

the dielectric matrix using localized Wannier orbitals instead of planewaves has also been used in the context of local field and excitonic effects in the optical spectrum of covalent crystals [44, 45].

In the LMTO method within the atomic sphere approximation (ASA), the basis functions are given by

$$\chi_{RL\nu}(\mathbf{r}) = \phi_{RL\nu}(\mathbf{r}) + \sum_{R'L'\nu'} \dot{\phi}_{R'L'\nu'}(\mathbf{r}) h_{R'L'\nu',RL\nu} \tag{186}$$

The orbital ϕ is a solution to the Schrödinger equation for a given energy ε_ν and $\dot{\phi}$ is the energy derivative at ε_ν. The label R denotes atom type, $L = lm$ denotes angular momentum, and ν denotes the principal quantum number when there are more than one orbital per L channel [9]. Therefore a product of wave functions consists of products of the orbitals ϕ and $\dot{\phi}$ [10]:

$$\phi_{RL\nu}\phi_{RL'\nu'}, \quad \phi_{RL\nu}\dot{\phi}_{RL'\nu'}, \quad \dot{\phi}_{RL\nu}\dot{\phi}_{RL'\nu'} \tag{187}$$

This means that these product orbitals form a complete set of basis functions for the polarization function and also for the response function and the self-energy as discussed below. ¿From Eq. (60) we have

$$R = P + PvP + PvPvP + \ldots \tag{188}$$

Writing $P = \sum_{ij} |i\rangle P_{ij} \langle j|$, the second term can be written

$$PvP = \sum_{ij,kl} |i\rangle P_{ij} \langle j|v|k\rangle P_{kl} \langle l| \tag{189}$$

Similar expressions can be written down for the other terms and we can therefore write $R = \sum_{ij} |i\rangle R_{ij} \langle j|$, i.e. P and R span the same space.

The self-energy of a given state $\psi_{\mathbf{k}n}$ in the GWA schematically has the form

$$\begin{aligned} \Sigma(\mathbf{k}n, \omega) &= < \psi_{\mathbf{k}n}|GW|\psi_{\mathbf{k}n} > \\ &= < \psi_{\mathbf{k}n}\psi|v|\psi\psi_{\mathbf{k}n} > + < \psi_{\mathbf{k}n}\psi|vRv|\psi\psi_{\mathbf{k}n} > \end{aligned} \tag{190}$$

The two $\psi\psi$ come from the G. We see that Σ is sandwiched by products of two wave functions and it is therefore sufficient to have v expanded in the product orbitals. Thus, the product orbitals in Eq. (187) form a complete basis for a GW calculation.

We reduce the number of product orbitals in three steps:

1. We neglect terms containing $\dot{\phi}$ since they are small ($\dot{\phi}$ is typically 10% of ϕ). This reduces the number of product functions by a factor of 3. In some cases, it may be necessary to include $\dot{\phi}$.

2. If we are only interested in valence states, then there are no products between conduction states. Therefore, in sp systems, products of $\phi_d\phi_d$ can be neglected and similarly in d systems, products of $\phi_f\phi_f$ may be neglected.

3. The remaining product orbitals turn out to have a large number of linear dependencies, typically $30 - 50\%$. These linear dependencies can be eliminated systematically, which is described below.

A product orbital is defined by

$$\begin{aligned} b_i(\mathbf{r}) &\equiv \phi_{\mathbf{R}L\nu}(\mathbf{r})\phi_{\mathbf{R}L'\nu'}(\mathbf{r}) \\ &= \varphi_{\mathbf{R}l\nu}(r)\varphi_{\mathbf{R}l'\nu'}(r) y_L(\hat{\mathbf{r}}) y_{L'}(\hat{\mathbf{r}}) \end{aligned} \tag{191}$$

where $i = (\mathbf{R}, L\nu, L'\nu')$. Due to the ASA, this function is only non-zero inside a sphere centered on atom $\mathbf{R}$. Thus, there are no products between orbitals centered on different spheres. For a periodic system we need a Bloch basis and perform a Bloch sum

$$b_{\mathbf{k}i}(\mathbf{r}) = \sum_{\mathbf{T}} e^{i\mathbf{k}\cdot\mathbf{T}} \phi_{\mathbf{R}L\nu}(\mathbf{r} - \mathbf{R} - \mathbf{T}) \phi_{\mathbf{R}L'\nu'}(\mathbf{r} - \mathbf{R} - \mathbf{T}) \tag{192}$$

The $\mathbf{k}$ dependence is in some sense artificial because the function has no overlap with neighbouring spheres, similar to core states. After leaving out unnecessary products (step 2), we optimize the basis by eliminating linear dependencies (step 3). This is done by orthogonalizing the overlap matrix

$$O_{ij} = < b_i | b_j > \tag{193}$$

$$Oz = ez \tag{194}$$

and neglecting eigenvectors z with eigenvalues $\epsilon <$ tolerance. The resulting orthonormal basis is a linear combination of the product functions:

$$B_i = \sum_j b_j z_{ji}, \tag{195}$$

and typically we have $\sim 70 - 100$ functions per atom with *spdf* orbitals. The above procedure ensures that we have the smallest number of basis functions. Further approximations may be introduced to reduce the basis.

In Table 1 we show a completeness test for the basis. The slight discrepancy for high lying states is due to the neglect of $\dot{\phi}$ in the product basis which become more important for the broad high lying conduction states, and also because the optimization procedure puts less weight on those products which have smaller overlap. This is not crucial for two reasons: the matrix elements become smaller for the higher states, and in relation to *GW* calculations, there is a factor of $1/\omega$ which makes the higher states less important.

As applications of the product basis, we have calculated as examples the energy loss spectra of Ni (Fig. 2), Si (Fig. 3). and NiO (Fig. 8).

7.2 Special directions

When calculating matrix elements using the product basis, we encounter angular integrals of the form

$$\int d\Omega \, y_{L_1} y_{L_2} y_{L_3} y_{L_4} \tag{196}$$

Analytic evaluation of these integrals are computationally expensive. Instead, they are calculated by using special directions which are analogous to Gaussian quadratures in one dimension. In general, Gaussian integration over a unit sphere means that we try to find M directions Ω_i and weights w_i such that

$$\sum_{i=1}^{M} w_i y_L(\Omega_i) = \frac{\delta_{l,0}}{\sqrt{4\pi}} \quad for \;\; 0 \le l \le l_{max}, \;\; -l \le m \le l \tag{197}$$

Gaussian accuracy is achieved when the number of correctly integrated spherical harmonics is equal to the number of free parameters which is $3M - 2$, since the spherical harmonics transform among themselves under rotation so that one direction can thus be taken to be the z direction and the sum of the weights is one. In one dimension, Gaussian accuracy is always achieved and the mesh points are uniquely determined. In two dimensions, it is rarely achieved although

one usually comes rather close. There are often several sets of directions that yield the same accuracy. We have found by minimization of

$$\sum_L |\sum_{i=1}^{M} w_i y_L(\Omega_i) - \frac{\delta_{l,0}}{\sqrt{4\pi}}|^2 \tag{198}$$

a set of 62 cubic directions which correctly integrates all spherical harmonics up to and including $l = 11$ plus a few more for the $l = 12$ harmonics, a total of 168 functions. Full Gaussian accuracy would mean $3 \times 62 - 2 = 184$ functions and the cubic constraint only leads to a 9% less effective integration formula. The directions and the corresponding weights are

Direction	*Weight*/4π
(1, 0, 0)	0.130 612 244 897 931 /6
(1, 1, 1)/$\sqrt{3}$	0.128 571 428 571 554/8
(0.846 433 804 070 399,	0.740 816 326 530 515 /48
0.497 257 813 599 068, 0.190 484 860 662 438)	

(199)

plus all possible cubic variations of these (sign changes and permutations). We have also found a larger set with 114 cubic directions which integrate up to $l = 15$. The directions and weights are

Direction	*Weight*/4π
(1, 0, 0)	0.076 190 476 192 774 /6
(1, 1, 0)/$\sqrt{2}$	0.137 357 478 197 258 /12
(0.733 519 276 107 007,	0.344 086 737 167 612 /48
0.570 839 829 704 020, 0.368 905 625 333 822)	
(0.909 395 474 471 327,	0.442 365 308 442 356 /48
0.385 850 474 128 732, 0.155 303 839 700 451)	

(200)

Using these directions, any angular function can be integrated easily

$$\int d\Omega\, f(\Omega) = \sum_i w_i f(\Omega_i) \tag{201}$$

7.3 Evaluation of the Coulomb Matrix

We consider one atom per unit cell for simplicity. Extension to several atoms is straightforward. The Coulomb matrix is given by

$$v_{ij}(\mathbf{q}) = \frac{1}{N} \int d^3r \int d^3r' \, \frac{B^*_{\mathbf{q}i}(\mathbf{r}) B_{\mathbf{q}j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{202}$$

where $B_{\mathbf{q}i}$ is normalized to unity in the unit cell. The integrations over the whole space may be reduced to integrations over a unit cell Ω by using the property

$$B_{\mathbf{q}i}(\mathbf{r} + \mathbf{T}) = e^{i\mathbf{q}\cdot\mathbf{T}} B_{\mathbf{q}i}(\mathbf{r}) \tag{203}$$

and noting that the integration over $\mathbf{r}'$ is independent of the origin of $\mathbf{r}$. This gives

$$v_{ij}(\mathbf{q}) = \int_\Omega d^3r \int_\Omega d^3s \; B^*_{\mathbf{q}i}(\mathbf{s}) E_{\mathbf{q}}(\mathbf{s}, \mathbf{r}) B_{\mathbf{q}j}(\mathbf{r}) \tag{204}$$

where

$$
\begin{aligned}
E_{\mathbf{q}}(\mathbf{s},\mathbf{r}) &= \sum_{\mathbf{T}} \frac{e^{i\mathbf{q}\cdot\mathbf{T}}}{|\mathbf{s}-\mathbf{r}-\mathbf{T}|} \\
&= \frac{4\pi}{\Omega}\sum_{\mathbf{G}} \frac{e^{-(\mathbf{q}+\mathbf{G})^2/4\alpha^2}}{(\mathbf{q}+\mathbf{G})^2} e^{i(\mathbf{q}+\mathbf{G})\cdot(\mathbf{s}-\mathbf{r})} \\
&+ \alpha\sum_{\mathbf{T}} e^{i\mathbf{q}\cdot\mathbf{T}} \frac{erfc(\alpha|\mathbf{s}-\mathbf{r}-\mathbf{T}|)}{\alpha|\mathbf{s}-\mathbf{r}-\mathbf{T}|}
\end{aligned}
\tag{205}
$$

The Ewald method has been used to obtain the above decomposition into summations in the reciprocal and real space. $erfc$ is the complementary error function equal to $(1-erf)$, and α is an arbitrary constant whose value is chosen to give a fast convergence in the number of reciprocal lattice vectors and the number of neighbours. The essence of the Ewald method is to add and substract a Gaussian charge distribution which breaks the Coulomb potential from a point charge into a short and long range part. The short range part is done in real space and the long range part is done in reciprocal space. The main task is to calculate the potential

$$
\begin{aligned}
\Phi_{\mathbf{q}j}(\mathbf{s}) &= \int_{\Omega} d^3r\; E_{\mathbf{q}}(\mathbf{s},\mathbf{r}) B_{\mathbf{q}j}(\mathbf{r}) \\
&= \sum_i p_{\mathbf{q}i}(\mathbf{s}) z_{ij}
\end{aligned}
\tag{206}
$$

with z given by Eq. (194) and

$$
\begin{aligned}
p_{\mathbf{q}i}(\mathbf{s}) &= \int_{\Omega} d^3r\; b_{\mathbf{q}i}(\mathbf{r}) E_{\mathbf{q}}(\mathbf{s},\mathbf{r}) \\
&= \sum_{\mathbf{G}} p_{\mathbf{q}i}(\mathbf{s},\mathbf{G}) + \sum_{\mathbf{T}} p_{\mathbf{q}i}(\mathbf{s},\mathbf{T})
\end{aligned}
\tag{207}
$$

where

$$
p_{\mathbf{q}i}(\mathbf{s},\mathbf{G}) = \frac{4\pi}{\Omega}\frac{e^{-(\mathbf{q}+\mathbf{G})^2/4\alpha^2}}{(\mathbf{q}+\mathbf{G})^2} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{s}} \int_{\Omega} d^3r\; e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} b_{\mathbf{q}i}(\mathbf{r}) \tag{208}
$$

and

$$
p_{\mathbf{q}i}(\mathbf{s},T) = \alpha e^{i\mathbf{q}\cdot\mathbf{T}} \int_{\Omega} d^3r\; \frac{erfc(\alpha|\mathbf{s}-\mathbf{r}-\mathbf{T}|)}{\alpha|\mathbf{s}-\mathbf{r}-\mathbf{T}|} b_{\mathbf{q}i}(\mathbf{r}) \tag{209}
$$

It is straightforward to calculate $p_{\mathbf{q}j}(\mathbf{s},\mathbf{G})$, since it is a Fourier transform of $b_{\mathbf{q}j}(\mathbf{r})$. To calculate $p_{\mathbf{q}i}(\mathbf{s},\mathbf{T})$, we use the following expansion formulas

$$
\frac{1}{|\mathbf{s}-\mathbf{r}|} = \sum_L \frac{4\pi}{2l+1}\frac{r_<^l}{r_>^{l+1}} y_L(\hat{\mathbf{s}}) y_L(\hat{\mathbf{r}}) \tag{210}
$$

and

$$
\frac{erf(\alpha|\mathbf{s}-\mathbf{r}|)}{\alpha|\mathbf{s}-\mathbf{r}|} = \sum_L \frac{4\pi}{2l+1} g_l(r,s) y_L(\hat{\mathbf{s}}) y_L(\hat{\mathbf{r}}) \tag{211}
$$

The coefficients $g_l(r,s)$ are determined by numerical integrations using special directions in Eq. (199).

At the central sphere, $b_{\mathbf{q}i}(\mathbf{r})$ has no $\mathbf{q}$ dependence and is given by Eq. (191), so that

$$
p_{\mathbf{q}i}(\mathbf{s},\mathbf{T}) = \alpha e^{i\mathbf{q}\cdot\mathbf{T}} \sum_L w_{li}(s_{\mathbf{T}}) y_L(\hat{\mathbf{s}_{\mathbf{T}}}) \int d\Omega y_{L_1} y_{L_2} y_L \tag{212}
$$

where $\mathbf{s_T} = \mathbf{s} - \mathbf{T}$ and

$$w_{li}(s_{\mathbf{T}}) = \frac{4\pi}{2l+1}\int_0^R dr r^2 \left\{ \frac{r_<^l}{r_>^{l+1}} - g_l(r, s_{\mathbf{T}}) \right\} b_i(r) \tag{213}$$

We note that the sum over l is cut off by l_1 and l_2 in the product function b. Finally,

$$v_{ij}(\mathbf{q}) = \int_\Omega d^3s B^*_{\mathbf{q}i}(\mathbf{s}) \Phi_{\mathbf{q}j}(\mathbf{s}) \tag{214}$$

which is easily done with special directions in Eq. (199).

7.4 Evaluation of the polarization P

Calculations of P are the most time consuming due to a large number of matrix elements, a summation over the Brillouin zone, which is not restricted to the irreducible zone, and a sum over occupied and unoccupied states as may be seen from the following expression:

$$\begin{aligned} P_{ij}(\mathbf{q},\omega) &= \sum_{\mathbf{k}} \sum_{n\le\mu} \sum_{n'>\mu} \langle B_{\mathbf{q}i}\psi_{\mathbf{k}n} | \psi_{\mathbf{k}+\mathbf{q},n'} \rangle \langle \psi_{\mathbf{k}+\mathbf{q},n'} | \psi_{\mathbf{k}n} B_{\mathbf{q}j} \rangle \\ &\quad \times \left\{ \frac{1}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q},n'} + \varepsilon_{\mathbf{k}n} + i\delta} - \frac{1}{\omega + \varepsilon_{\mathbf{k}+\mathbf{q},n'} - \varepsilon_{\mathbf{k}n} - i\delta} \right\} \end{aligned} \tag{215}$$

For real frequencies, we calculate S^0 in Eq. (76). The δ function is replaced by a Gaussian

$$\delta(x) \Longrightarrow \frac{\exp -(x/\sigma)^2}{\sigma\sqrt{\pi}} \tag{216}$$

The self-energy is not sensitive to the choice of σ.

The matrix elements reduce into integrals of four orbitals

$$\begin{aligned} \int d^3r\ \phi_{RL_1\nu_1}\phi_{RL_2\nu_2}\phi_{RL_3\nu_3}\phi_{RL_4\nu_4} &= \int dr\ r^2 \varphi_{Rl_1\nu_1}\varphi_{Rl_2\nu_2}\varphi_{Rl_3\nu_3}\varphi_{Rl_4\nu_4} \\ &\quad \times \int d\Omega\ y_{L_1} y_{L_2} y_{L_3} y_{L_4} \end{aligned} \tag{217}$$

The angular integral is calculated by using special directions in Eq. (199).

To obtain the real part of P we calculate the Hilbert transform in Eq. (61) using the anti-symmetry of S:

$$Re\ P(\omega) = \int_0^\infty d\omega'\ S^0(\omega') \left\{ \frac{1}{\omega - \omega'} - \frac{1}{\omega + \omega'} \right\} \tag{218}$$

Although the integrand diverges when $\omega' = \omega$, the integral is well-defined because it is a principal value integral. In practice, S^0 is expanded in Taylor series around ω within an interval $\omega - h$ and $\omega + h$.

For imaginary frequency, we calculate P directly from Eq. (215) by setting $\omega \to i\omega$:

$$\begin{aligned} P_{ij}(\mathbf{q},i\omega) &= \sum_{\mathbf{k}} \sum_{n\le\mu} \sum_{n'>\mu} \langle B_{\mathbf{q}i}\psi_{\mathbf{k}n} | \psi_{\mathbf{k}+\mathbf{q},n'} \rangle \langle \psi_{\mathbf{k}+\mathbf{q},n'} | \psi_{\mathbf{k}n} B_{\mathbf{q}j} \rangle \\ &\quad \times \frac{-2(\varepsilon_{\mathbf{k}+\mathbf{q},n'} - \varepsilon_{\mathbf{k}n})}{\omega^2 + (\varepsilon_{\mathbf{k}+\mathbf{q},n'} - \varepsilon_{\mathbf{k}n})^2} \end{aligned} \tag{219}$$

Thus $P(\mathbf{r},\mathbf{r}',i\omega)$ is real along the imaginary axis although the matrix representation P_{ij} may be complex due to the matrix elements.

The Brillouin zone integration is performed using a simple sampling method. It is also possible to use a more accurate tetrahedron method but the replacement of the δ function by a Gaussian is not possible anymore, resulting in a significantly more complicated programming.

Once we have obtain P, the response function R and the screened Coulomb interaction W can be calculated straightforwardly.

7.5 Evaluation of the self-energy Σ

Taking the matrix element of the bare exchange in a Bloch state $\psi_{\mathbf{q}n}$ we get from Eq. (78)

$$\Sigma^x_{\mathbf{q}n} = \sum_{\mathbf{k}} \sum_{n' \le \mu} \sum_{ij} \langle \psi_{\mathbf{q}n}\psi_{\mathbf{k}-\mathbf{q}n'}|B_{\mathbf{k}i}\rangle \; v_{ij}(\mathbf{k}) \; \langle B_{\mathbf{k}j}|\psi_{\mathbf{k}-\mathbf{q}n'}\psi_{\mathbf{q}n}\rangle \tag{220}$$

obtained by expanding the Coulomb potential like in Eq. (182). ¿From Eq. (77) the correlated part of $Im\ \Sigma$ is given by

$$\Gamma_{\mathbf{q}n}(\omega) = \begin{cases} \sum_{\mathbf{k}} \sum_{n' \le \mu} \sum_{ij} \langle \psi_{\mathbf{q}n}\psi_{\mathbf{k}-\mathbf{q},n'}|B_{\mathbf{k}i}\rangle \; D_{ij}(\mathbf{k}, \omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'}) \\ \qquad \times \langle B_{\mathbf{k}j}|\psi_{\mathbf{k}-\mathbf{q},n'}\psi_{\mathbf{q}n}\rangle \; \theta(\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'}) \qquad for\ \ \omega \le \mu \\ \sum_{\mathbf{k}} \sum_{n' > \mu} \sum_{ij} \langle \psi_{\mathbf{q}n}\psi_{\mathbf{k}-\mathbf{q},n'}|B_{\mathbf{k}i}\rangle \; D_{ij}(\mathbf{k}, \varepsilon_{\mathbf{k}-\mathbf{q},n'} - \omega) \\ \qquad \times \langle B_{\mathbf{k}j}|\psi_{\mathbf{k}-\mathbf{q},n'}\psi_{\mathbf{q}n}\rangle \; \theta(\varepsilon_{\mathbf{k}-\mathbf{q},n'} - \omega) \qquad for\ \ \omega > \mu \end{cases} \tag{221}$$

The real part of Σ^c is obtained from the Hilbert transform (principal value integral) in Eq. (66). As in the case of the polarization, care must be taken when $\omega' = \omega$ by expanding Γ in Taylor series around ω.

The quasiparticle energy can now be calculated as follows:

$$\begin{aligned} E_{\mathbf{q}n} &= \varepsilon_{\mathbf{q}n} + \Delta\Sigma_{\mathbf{q}n}(E_{\mathbf{q}n}) \\ &= \varepsilon_{\mathbf{q}n} + \Delta\Sigma_{\mathbf{q}n}(\varepsilon_{\mathbf{q}n}) + (E_{\mathbf{q}n} - \varepsilon_{\mathbf{q}n})\frac{\partial\Delta\Sigma_{\mathbf{q}n}(\varepsilon_{\mathbf{q}n})}{\partial\omega} \end{aligned} \tag{222}$$

where

$$\Delta\Sigma_{\mathbf{q}n}(\omega) = \langle \psi_{\mathbf{q}n}|Re\ \Sigma(\omega) - V_{xc}|\psi_{\mathbf{q}n}\rangle \tag{223}$$

The self-energy correction $\Delta\Sigma$ is obtained from first order perturbation theory from Eq. (39) and the Kohn-Sham equation:

$$(H_0 + \Sigma)\Psi = E\Psi \tag{224}$$

$$(H_0 + V^{xc})\psi = \varepsilon\psi \tag{225}$$

where H_0 is the kinetic energy plus the Hartree potential. The self-energy correction to $\varepsilon_{\mathbf{q}n}$ is given by

$$\begin{aligned} \Delta\varepsilon_{\mathbf{q}n} &= E_{\mathbf{q}n} - \varepsilon_{\mathbf{q}n} \\ &= Z_{\mathbf{q}n}\Delta\Sigma_{\mathbf{q}n}(\varepsilon_{\mathbf{q}n}) \end{aligned} \tag{226}$$

where

$$Z_{\mathbf{q}n} = \left[1 - \frac{\partial \Delta\Sigma_{\mathbf{q}n}(\varepsilon_{\mathbf{q}n})}{\partial\omega}\right]^{-1} < 1 \tag{227}$$

is the quasiparticle weight.

The frequency integration of the self-energy may also be performed along the imaginary axis [41] plus contributions from the poles of the Green function (Fig. 4). ¿From Eq. (82) we have

$$\Sigma^c_{\mathbf{q}n}(\omega) = \sum_{\mathbf{k}} \sum_{n' \le \mu} \sum_{ij} \langle \psi_{\mathbf{q}n} \psi_{\mathbf{k}-\mathbf{q},n'} | B_{\mathbf{k}i} \rangle \ \langle B_{\mathbf{k}j} | \psi_{\mathbf{k}-\mathbf{q},n'} \psi_{\mathbf{q}n} \rangle \times \frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' \ \frac{W^c_{ij}(\mathbf{k},\omega')}{\omega + \omega' - \varepsilon_{\mathbf{k}-\mathbf{q},n'} - i\delta} \tag{228}$$

$$+ \sum_{\mathbf{k}} \sum_{n' > \mu} \sum_{ij} \langle \psi_{\mathbf{q}n} \psi_{\mathbf{k}-\mathbf{q},n'} | B_{\mathbf{k}i} \rangle \ \langle B_{\mathbf{k}j} | \psi_{\mathbf{k}-\mathbf{q},n'} \psi_{\mathbf{q}n} \rangle \times \frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' \ \frac{W^c_{ij}(\mathbf{k},\omega')}{\omega + \omega' - \varepsilon_{\mathbf{k}-\mathbf{q},n'} + i\delta} \tag{229}$$

We consider the integration along the imaginary axis with $\omega' \to i\omega''$, ω'' real, and along the path C (Fig. 1):

$$\begin{aligned}
&\frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' \ \frac{W^c_{ij}(\mathbf{k},\omega')}{\omega + \omega' - \varepsilon_{\mathbf{k}-\mathbf{q},n'} \pm i\delta} \\
&= -\int_0^{\infty} \frac{d\omega''}{2\pi} W^c_{ij}(\mathbf{k}, i\omega'') \left\{ \frac{1}{\omega + i\omega'' - \varepsilon_{\mathbf{k}-\mathbf{q},n'}} + \frac{1}{\omega - i\omega'' - \varepsilon_{\mathbf{k}-\mathbf{q},n'}} \right\} \\
&\quad + \frac{i}{2\pi} \int_C d\omega' \ \frac{W^c_{ij}(\mathbf{k},\omega')}{\omega + \omega' - \varepsilon_{\mathbf{k}-\mathbf{q},n'} \pm i\delta} \\
&= -\int_0^{\infty} d\omega'' \ W^c_{ij}(\mathbf{k}, i\omega'') \frac{1}{\pi} \frac{\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'}}{(\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'})^2 + \omega''^2} \\
&\quad \pm W^c_{ij}[\mathbf{k}, \pm(\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'})] \ \theta[\pm(\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'})] \ \theta[\pm(\omega - \mu)] \\
&\quad \times \theta[\pm(\varepsilon_{\mathbf{k}-\mathbf{q},n'} - \mu)]
\end{aligned} \tag{230}$$

The first term is the contribution along the imaginary axis and the second from the poles of G. The integrand in the first term is very peaked around $\omega' = 0$ when $\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'}$ is small. To handel this problem, we add and substract the following term

$$\begin{aligned}
&\int_0^{\infty} d\omega' \ W^c_{ij}(\mathbf{k},0) e^{-\alpha^2\omega'^2} \frac{1}{\pi} \frac{\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'}}{(\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'})^2 + \omega''^2} \\
&= W^c_{ij}(\mathbf{k},0) \frac{\pi}{2(\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'})} e^{\alpha^2(\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'})^2} erfc[\alpha(\omega - \varepsilon_{\mathbf{k}-\mathbf{q},n'})]
\end{aligned} \tag{231}$$

When this term is substracted from the integrand in the first term, the resulting integrand is smooth and a Gaussian quadrature may be used.

The GWA has been applied to many systems. Here we show some results for Ni (Fig. 5) and NiO (Fig. 7). The self-energy of Ni at the Γ-point is shown in Fig. 6. The imporatnce of the starting Hamiltonian in GW calculations is illustrated in Fig. 9.

APPENDIX A

¿From the spectral representation of W_c in Eq. (67) we have

$$\begin{aligned} W_c(0) &= \int_{-\infty}^{0} d\omega' \frac{D(\omega')}{-\omega' - i\delta} + \int_{0}^{\infty} d\omega' \frac{D(\omega')}{-\omega' + i\delta} \\ &= -2\int_{0}^{\infty} d\omega' \frac{D(\omega')}{\omega' - i\delta} \end{aligned} \tag{232}$$

using the fact that $D(\omega)$ is odd.

$$\begin{aligned} W_c^-(0) &= \frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega' \frac{W_c(\omega')}{\omega' - i\delta} \\ &= \frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega' \frac{1}{\omega' - i\delta} \\ &\times \left\{ \int_{-\infty}^{0} d\omega'' \frac{D(\omega'')}{\omega' - \omega'' - i\delta} + \int_{0}^{\infty} d\omega'' \frac{D(\omega'')}{\omega' - \omega'' + i\delta} \right\} \\ &= \int_{0}^{\infty} d\omega'' \frac{D(\omega'')}{\omega'' - i\delta} \\ &= -\frac{1}{2} W_c(0) \end{aligned} \tag{233}$$

APPENDIX B

$$\Sigma^c(E) = \frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega\, G(E+\omega) W^c(\omega) \tag{234}$$

The space variables have been dropped out for clarity. Using the spectral representations of G and W^c in Eqs. (24) and (67) we get

$$\begin{aligned} \Sigma^c(E) &= \frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega \left\{ \int_{-\infty}^{\mu} d\omega_1 \frac{A(\omega_1)}{E+\omega-\omega_1-i\delta} \right. \\ &\qquad \left. + \int_{\mu}^{\infty} d\omega_1 \frac{A(\omega_1)}{E+\omega-\omega_1+i\delta} \right\} \\ &\times \int_{0}^{\infty} d\omega_2\, D(\omega_2) \left\{ \frac{1}{\omega-\omega_2+i\eta} - \frac{1}{\omega+\omega_2-i\eta} \right\} \end{aligned} \tag{235}$$

Performing the contour integration in ω yields

$$\begin{aligned} \Sigma^c(E) &= \int_{-\infty}^{\mu} d\omega_1 \int_{0}^{\infty} d\omega_2 \frac{A(\omega_1)D(\omega_2)}{E+\omega_2-\omega_1-i\delta} \\ &+ \int_{\mu}^{\infty} d\omega_1 \int_{0}^{\infty} d\omega_2 \frac{A(\omega_1)D(\omega_2)}{E-\omega_2-\omega_1+i\delta} \end{aligned} \tag{236}$$

The spectral function of Σ^c in Eq. (66) is then

$$\begin{aligned} \Gamma(E) &= -\mathrm{sgn}(E-\mu)\int_{-\infty}^{\mu} d\omega_1 \int_{0}^{\infty} d\omega_2 A(\omega_1)D(\omega_2)\delta(E+\omega_2-\omega_1) \\ &+\mathrm{sgn}(E-\mu)\int_{\mu}^{\infty} d\omega_1 \int_{0}^{\infty} d\omega_2 A(\omega_1)D(\omega_2)\delta(-E+\omega_2+\omega_1) \\ &-\mathrm{sgn}(E-\mu)\int_{0}^{\infty} d\omega_2 \theta(\mu-E-\omega_2)A(E+\omega_2)D(\omega_2) \\ &+\mathrm{sgn}(E-\mu)\int_{0}^{\infty} d\omega_2 \theta(-\mu+E-\omega_2)A(E-\omega_2)D(\omega_2) \end{aligned} \tag{237}$$

Putting back the space variables and using A in Eq. (75) we obtain

$$\begin{aligned}\Gamma(\mathbf{r},\mathbf{r}',E) &= -\mathrm{sgn}(E-\mu)\sum_i \theta(\mu-\varepsilon_i)\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')D(\mathbf{r},\mathbf{r}';\varepsilon_i-E)\theta(\varepsilon_i-E) \\ &\quad +\mathrm{sgn}(E-\mu)\sum_i \theta(\varepsilon_i-\mu)\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')D(\mathbf{r},\mathbf{r}';E-\varepsilon_i)\theta(E-\varepsilon_i)\end{aligned} \tag{238}$$

Finally,

$$\Gamma(\mathbf{r},\mathbf{r}',E) = \begin{cases} \sum_{i\le\mu}\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')D(\mathbf{r},\mathbf{r}',\varepsilon_i-E)\theta(\varepsilon_i-E) & \text{for } E\le\mu \\ \sum_{i>\mu}\psi_i^*(\mathbf{r})\psi_i(\mathbf{r}')D(\mathbf{r},\mathbf{r}',E-\varepsilon_i)\theta(E-\varepsilon_i) & \text{for } E>\mu \end{cases} \tag{239}$$

References

[1] Almbladh C-O and Hedin L 1983 *Handbook on Synchrotron Radiation* **1** 686 ed. E. E. Koch (North-Holland)

[2] Almbladh C-O, von Barth U, and van Leeuwen R 1999, *Int. J. Mod. Phys. B* **13** 535

[3] Andersen O K 1975 *Phys. Rev. B* **12** 3060

[4] Anisimov V I, Zaanen J, and Andersen O K 1991 *Phys. Rev. B* **44** 943

[5] Anisimov V I, Solovyev I V, Korotin M A, Czyzyk M T, and Sawatzky G A 1993 *Phys. Rev. B* **48** 16929

[6] Anisimov V I, Aryasetiawan F, and Lichtenstein A I 1997 *J. Phys.: Condens. Matter* **9** 767-808

[7] Arai M and Fujiwara T 1995 *Phys. Rev. B* **51** 1477-89

[8] Aryasetiawan F 1992a *Phys. Rev. B* **46** 13051-64

[9] Aryasetiawan F and Gunnarsson O 1994a *Phys. Rev.B* **49** 7219

[10] Aryasetiawan F and Gunnarsson O 1994b *Phys. Rev. B* **49** 16214-22

[11] Aryasetiawan F and Gunnarsson O 1995 *Phys. Rev. Lett.* **74** 3221-24

[12] Aryasetiawan F and Karlsson K 1996 *Phys. Rev. B* **54** 5353-7

[13] Aryasetiawan F, Hedin L, and Karlsson K 1996 *Phys. Rev. Lett.* **77** 2268-71

[14] Baym G and Kadanoff L P 1961 *Phys. Rev.* **124** 287

[15] Baym G 1962 *Phys. Rev.* **127** 1391

[16] Becke A D 1988 *Phys. Rev. A* **38** 3098

[17] Becke A D 1992 *J. Chem. Phys.* **96** 2155

[18] Becke A D 1996 *J. Chem. Phys.* **104** 1040-6

[19] Bergersen B, Kus F W, and Blomberg C 1973 *Can. J. Phys.* **51** 102-110

[20] Berglund C N and Spicer W E 1964 *Phys. Rev* **136** A 1030

[21] Bobbert P A and van Haeringen W 1994 *Phys. Rev. B* **49** 10326-31

[22] Bylander D M and Kleinman L 1995a *Phys. Rev. Lett.* **74** 3660

[23] Bylander D M and Kleinman L 1995b *Phys. Rev. B* **52** 14566

[24] Calandra C and Manghi F 1992 *Phys. Rev. B* **45** 5819

[25] Ceperley D M and Alder B J 1980 *Phys. Rev. Lett.* **45** 566

[26] Cowan R D 1967 *Phys. Rev.* **163** 54

[27] Daling R and van Haeringen W 1989 *Phys. Rev. B* **40** 11659-65

[28] Daling R, Unger P, Fulde P, and van Haeringen W 1991 *Phys. Rev. B* **43** 1851-4

[29] de Groot H J, Bobbert P A, and van Haeringen W 1995 *Phys. Rev. B* **52** 11000

[30] Deisz J J, Eguiluz A, and Hanke W 1993 *Phys. Rev. Lett.* **71** 2793-96

[31] Deisz J J and Eguiluz A 1997 *Phys. Rev. B* **55** 9195-9

[32] Del Sole R, Reining L, and Godby R W 1994 *Phys. Rev. B* **49** 8024-8

[33] Faddeev L D 1963 *Sov. Phys. JETP* **12** 275

[34] Fan H Y 1945 *Phys. Rev* **68** 43

[35] Farid B, Godby R W, and Needs R J 1990 *20th International Conference on the Physics of Semiconductors* editors Anastassakis E M and Joannopoulos J D, Vol. 3 1759-62 (World Scientific, Singapore)

[36] Fetter A L and Walecka J D 1971 *Quantum Theory of Many-Particle Systems* (McGraw-Hill)

[37] Frota H O and Mahan G D 1992 *Phys. Rev. B* **45** 6243

[38] Galitskii V M and Migdal A B 1958 *Sov. Phys. JETP* **7** 96

[39] Galitskii V M 1958 *Sov. Phys. JETP* **7** 104

[40] Georges A, Kotliar G, Krauth W and Rozenberg M J 1996 *Rev. Mod. Phys.* **68** 13-125

[41] Godby R W, Schlüter M, and Sham L J 1988 *Phys. Rev. B* **37** 10159-75

[42] Gunnarsson O, Andersen O K, Jepsen O, and Zaanen J 1989 *Phys. Rev. B* **39** 1708

[43] Gunnarsson O, Meden V, and Schönhammer K 1994 *Phys. Rev. B* **50** 10462

[44] Hanke W and Sham L J 1975 *Phys. Rev. B* **12** 4501-11

[45] Hanke W and Sham L J 1988 *Phys. Rev. B* **38** 13361-70

[46] Hedin L 1965a *Phys. Rev.* **139** A796

[47] Hedin L and Lundqvist S 1969 *Solid State Physics vol. 23*, eds. H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York).

[48] Hedin L, Lundqvist B I, and Lundqvist S 1970 *J. Res. Natl. Bur. Stand. Sect. A* **74A** 417

[49] Hedin L 1980 *Physica Scripta* **21** 477-80

[50] Hedin L 1999 *J. Phys.: Condens. Matter* **11** R489-R528

[51] Hindgren M 1997 (PhD thesis, University of Lund)

[52] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864

[53] Holm B and Aryasetiawan F 1997 *Phys. Rev. B* **56**, 12825

[54] Holm B and von Barth U 1998 *Phys. Rev. B* **57**, 2108

[55] Hybertsen M S and Louie S G 1986 *Phys. Rev. B* **34** 5390-413

[56] Igarashi J 1983 *J. Phys. Soc. Jpn.* **52** 2827

[57] Igarashi J 1985 *J. Phys. Soc. Jpn.* **54** 260

[58] Igarashi J, Unger P, Hirai K, and Fulde P 1994 *Phys. Rev. B* **49** 16181

[59] Inkson J C 1984 *Many-body Theory of Solids* (Plenum Press, New York, 1984)

[60] Kanamori J 1963 *Prog. Theor. Phys.* **30** 275

[61] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133

[62] Kotani T 1995 *Phys. Rev. Lett.* **74** 2989

[63] Kotani T and Akai H 1996 *Phys. Rev. B* **54** 16502

[64] Landau L D 1956 *Zh. Eksperimen. i Teor. Fiz.* **30** 1058

[65] Landau L D 1957 *Zh. Eksperimen. i Teor. Fiz.* **32** 59

[66] Langreth D C 1970 *Phys. Rev. B***1** 471

[67] Langreth D C and Mehl M J 1983 *Phys. Rev. B* **28** 1809

[68] Lichtenstein A I, Zaanen J, and Anisimov V I 1995 *Phys. Rev. B* **52** R5467

[69] Liebsch A 1979 *Phys. Rev. Lett.* **43** 1431-4

[70] Liebsch A 1981 *Phys. Rev. B* **23** 5203-12

[71] Lindgren I 1971 *Int. J. Quantum Chem.* **5** 411

[72] Lundqvist B I 1967a *Phys. Kondens. Mater.* **6** 193

[73] Lundqvist B I 1967b *Phys. Kondens. Mater.* **6** 206

[74] Luttinger J M and Ward J C 1960 *Phys. Rev* **118** 1417-27

[75] Mahan G D 1967a *Phys. Rev.***153** 882

[76] Mahan G D 1967b *Phys. Rev.* **163** 612

[77] Mahan G D and Sernelius B E 1989 *Phys. Rev. Lett.* **62** 2718

[78] Mahan G D 1990 *Many-particle Physics* (Plenum Press, New York)

[79] Mahan G D 1994 *Comments Cond. Mat. Phys.* **16** 333

[80] Manghi F, Calandra C and Ossicini S 1994 *Phys. Rev. Lett.* **73** 3129-32

[81] Mayer H and Thomas H 1957 *Z. Physik* **147** 419

[82] Northrup J E, Hybertsen M S, and Louie S G 1987 *Phys. Rev. Lett.* **59** 819

[83] Northrup J E, Hybertsen M S, and Louie S G 1989 *Phys. Rev. B* **39** 8198

[84] Nozières P and de Dominicis C J 1969 *Phys. Rev.* **178** 1097

[85] Onida G, Reining L, Godby R W, Del Sole R and Andreoni W 1995 *Phys. Rev. Lett.* **75** 818-21

[86] Penn D R 1979 *Phys. Rev. Lett.* **42** 921-4

[87] Perdew J P and Zunger A 1981 *Phys. Rev. B* **23** 5048

[88] Perdew J P, Burke K, and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865-8

[89] Puff H 1961 *Phys. Stat. Sol.* **1** 704

[90] Rice T M 1965 *Ann. Phys.* **31** 100

[91] Shirley E L 1996 *Phys. Rev. B* **54** 7758-64

[92] Smith N V 1978 in *Photoemission in Solids I*, eds. Cardona M and Ley L (Springer-Verlag, Berlin, Heidelberg)

[93] Solovyev I V, Dederichs P H, and Anisimov V I (1994) *Phys. Rev. B* **50** 16861-71

[94] Solovyev I, Hamada N, and Terakura K 1996 *Phys. Rev. B* **53** 7158

[95] Springer M, Svendsen P S, and von Barth U 1996 *Phys. Rev. B* **54** 17392-401

[96] Springer M, Thesis 1997 (Lund University)

[97] Springer M, Aryasetiawan F, and Karlsson K 1998 *Phys. Rev. Lett.* **80**, 2389

[98] Steiner M M, Albers R C, and Sham L J 1992 *Phys. Rev. B* **45** 13272

[99] Surh M P, Northrup J E, and Louie S G 1988 *Phys. Rev. B* **38** 5976-80

[100] Svane A and Gunnarsson O 1990 *Phys. Rev. Lett.* **65** 1148-51

[101] Svendsen P S and von Barth U 1996 *Phys. Rev. B* **54** 17402-13

[102] Szotek Z, Temmerman W M, and Winter H 1993 *Phys. Rev. B* **47** 4029

[103] Talman J D and Shadwick W F 1976 *Phys. Rev. A* **14** 36

[104] Treglia G, Ducastelle F, and Spanjaard D 1980 *Phys. Rev. B* **21** 3729
Treglia G, Ducastelle F, and Spanjaard D 1980 *Phys. Rev. B* **21** 3729

[105] Verdozzi C, Godby R W, and Holloway S 1995 *Phys. Rev. Lett.* **74** 2327-30

[106] von Barth U and Holm B 1996 *Phys. Rev. B* **54** 8411-9

[107] Ward J C 1950 *Phys. Rev.* **78** 182

[108] Zunger A, Perdew J P, and Oliver G L 1980 *Solid State Commun.* **34** 933

$Core$	$\lvert < \psi_{\mathbf{k}n} \lvert \psi_{core} > \rvert^2$	$\sum_i \lvert c_i \rvert^2$	error
3s	.158114	.158113	.000001
3p	.066833	.066832	.000001
3p	.209174	.209172	.000002
3p	.168184	.168183	.000001

$\epsilon_{\mathbf{k}'n'}(eV)$	$\lvert < \psi_{\mathbf{k}n} \lvert \psi_{\mathbf{k}'n'} > \rvert^2$	$\sum_i \lvert c_i \rvert^2$	error
-9.09	.052222	.052220	.000002
-2.23	.181724	.181722	.000002
-2.23	.101124	.101124	.000000
-2.23	.167753	.167743	.000010
-1.22	.119586	.119583	.000003
-1.22	.014565	.014537	.000028
24.39	.060638	.060634	.000004
28.19	.013188	.013074	.000114
28.19	.008852	.008736	.000116
28.19	.001883	.001738	.000145
42.29	.015563	.015400	.000163
42.29	.018353	.018167	.000186
42.29	.006292	.006040	.000252
73.89	.017300	.016660	.000640
73.89	.011847	.011409	.000438
73.89	.013877	.012911	.000966

Table 1: The completeness test of the optimal product basis for Nickel. A product of two wavefunctions is expanded in the basis: $\psi^*_{\mathbf{k}n}\psi_{\mathbf{k}'n'} = \sum_i B_i c_i$ with k = (0 0 0), $\epsilon_{\mathbf{k}n} = -1.22\ eV$ (the higest valence state at the Γ point) and $\mathbf{k}' = (.5\ .5\ .5)$. The basis is complete if $\lvert \sum_i \lvert c_i \rvert^2 = \lvert < \psi_{\mathbf{k}n} \lvert \psi_{\mathbf{k}'n'} > \rvert^2$. The number of optimal product basis functions is 101 and 82 with and without $3s, 3p$ core states respectively.

FIGURE CAPTIONS

Figure 1: Bandgaps of a semiconductors and insulators calculated within the LDA and the GWA, compared with experiment. The data are provided by Dr. Eric Shirley [50].

Figure 2: The loss spectra of Ni with (solid line) and without (dotted line) local field compared with the experimental spectrum (full circles). Both theoretical spectra are calculated with 4s, 4p, 3d, 4f, and 5g LMTO orbitals, including an empty sphere at (0.5 0.5 0.5)a and core excitations

Figure 3: The loss spectra of Si with (solid line) and without (dotted line) local field compared with the experimental spectrum (full circles). Both theoretical spectra are calculated with 3s, 3p, 3d, and 4f LMTO orbitals including core excitations [10].

Figure 4: The analytic structure of $\Sigma^c = iGW^c$ for $\omega > E_F$ (a) and $\omega \leq E_F$ (b). Frequency integration of the self-energy along the real axis from $-\infty$ to ∞ is equivalent to the integration along the imaginary axis including the path C.

Figure 5: The bandstructure of Ni along ΓX and ΓL averaged over the majority and minority channels. The solid curves are the experiment and the dotted curves are the LDA (Mårtensson and Nilsson 1984). The filled circles are the quasiparticle energies in the GWA [8].

Figure 6: Ni self-energy in the GWA. (a) The real and imaginary parts of the correlation part of the self-energy for the minority spin state Γ'_{25}. (b) The real and imaginary parts of the correlation part of the self-energy for the majority spin state Γ'_{25} [8].

Figure 7: NiO bandstructure. (a) Comparison between the LDA (solid line) and the experimental bandstructure (Shen *et al* 1990, 1991a,b). (b) Comparison between the LDA (solid line) and the quasiparticle bandstructure in the GWA [11].

Figure 8: The energy-loss spectra of NiO. The smooth solid curve corresponds to the calculated spectrum with virtually no gap in the LDA Hamiltonian and the dashed one with $\approx$5 eV gap. The other curve is the experiment [12]

Figure 9: The spectral function of NiO at the Γ point for a Ni 3d state. The solid line corresponds to the case where the starting LDA Hamiltonian has virtually no gap and the dashed line to the case with ≈5 eV gap. (b) The same as in (a) but magnified 10 times [12]

Figure 10: Diagrammatic expansion for the Green function to second order in the GWA and the cumulant expansion respectively. The solid line represents the noninteracting Green function G_0 and the wiggly line represents the screened interaction W.

Figure 11: (a) The experimental spectral function for Na (dots). The solid line is a synthetic spectrum obtained by convoluting the density of states from a bandstructure calculation and the the experimental core level spectrum. BG is the estimated background contribution. The data are taken from Steiner, Höchst, and Hüfner (1979). (b) The total spectral function of Na for the occupied states. The solid and dashed line correspond to the cumulant expansion and GWA respectively [13].

Figure 12: Feynman diagrams for the T-matrix (square): The wiggly and the solid line with arrow represent the screened interaction U and the Green function G respectively.

Figure 13: Ni spectral functions at the X point for the second band [97].

Figure 14: The quasiparticle dispersions for $r_s = 4$ corresponding to full self-consistency (solid line), partial self-consistency (dashed line), first iteration (dotted line) and the free electron (dashed-dotted line) [54].

Figure 15: The partially self-consistent spectral function $A(k = 0.5k_F, \omega)$ compared to that of the first iteration for $r_s = 4$ [106].

Figure 16: The fully self-consistent spectral function $A(k = 0.5k_F, \omega)$ compared to that of the first iteration for $r_s = 4$ [53].

Figure 17: The spectral function of the partially self-consistent self-energy, $\Gamma = |\mathrm{Im}\Sigma|/\pi$, for $k = k_F$ compared to that of the first iteration [106].

Figure 18: The real part of the partially self-consistent self-energy for $k = k_F$ compared to that of the first iteration [106].

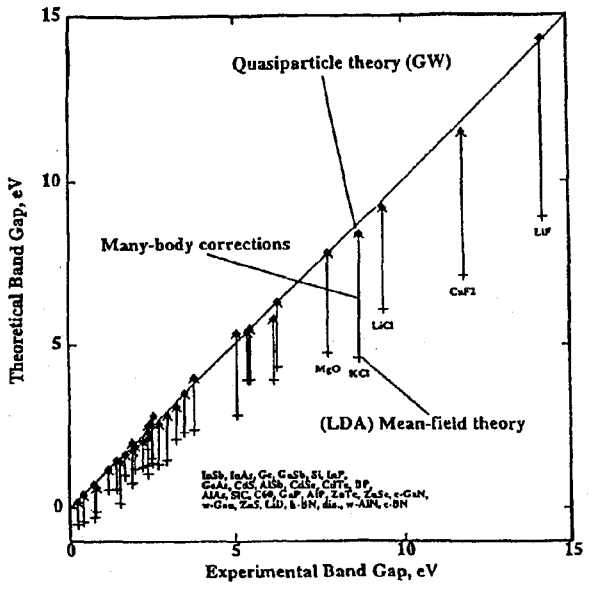

Figure 1

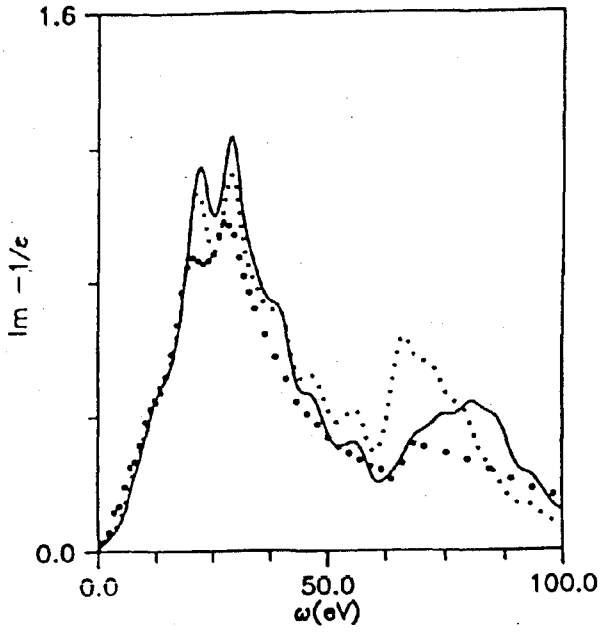

Figure 2

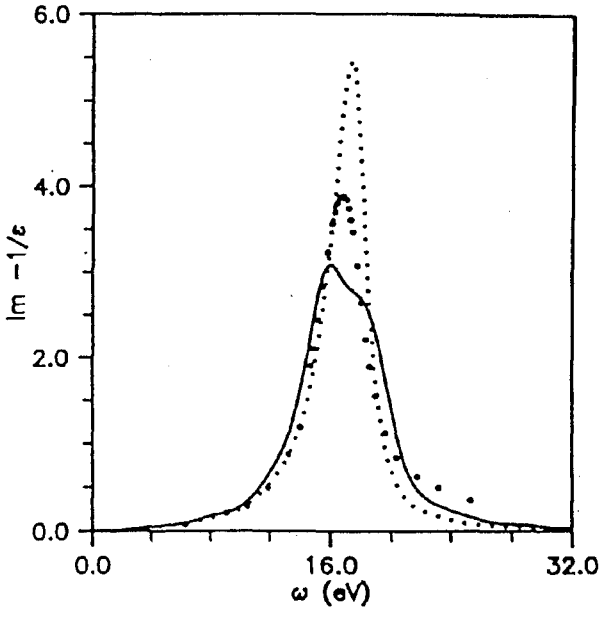

Figure 3

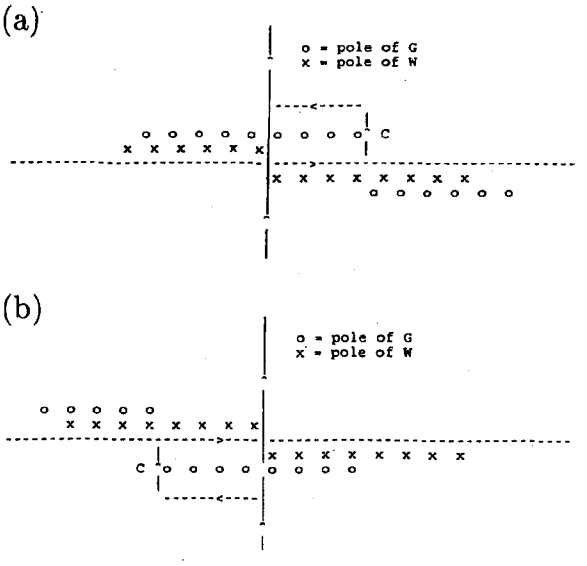

Figure 4

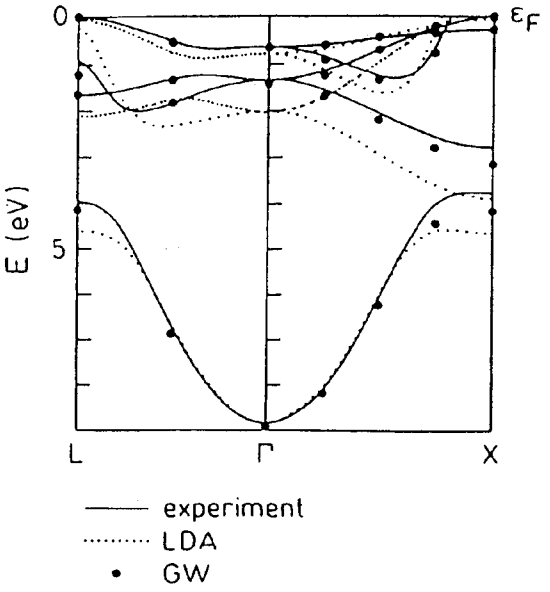

Figure 5

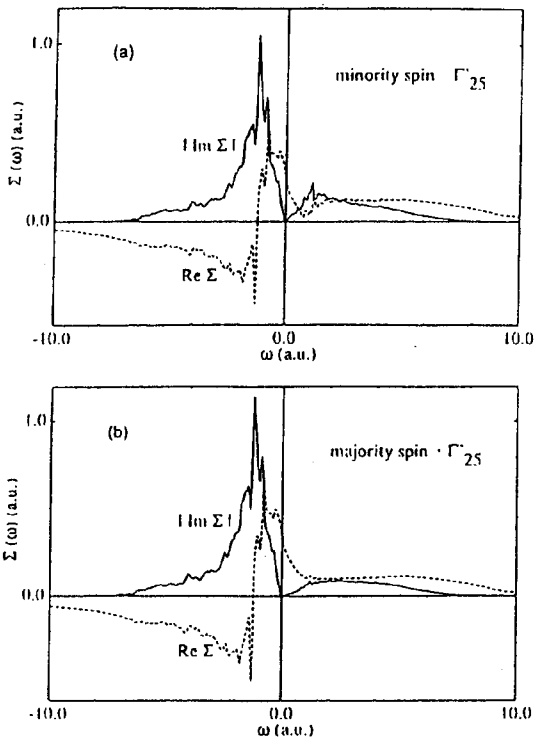

Figure 6

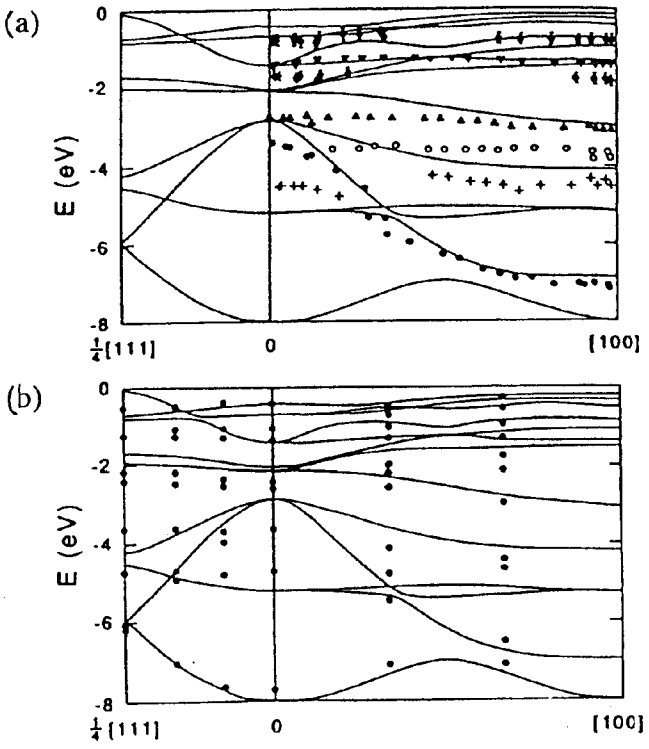

Figure 7

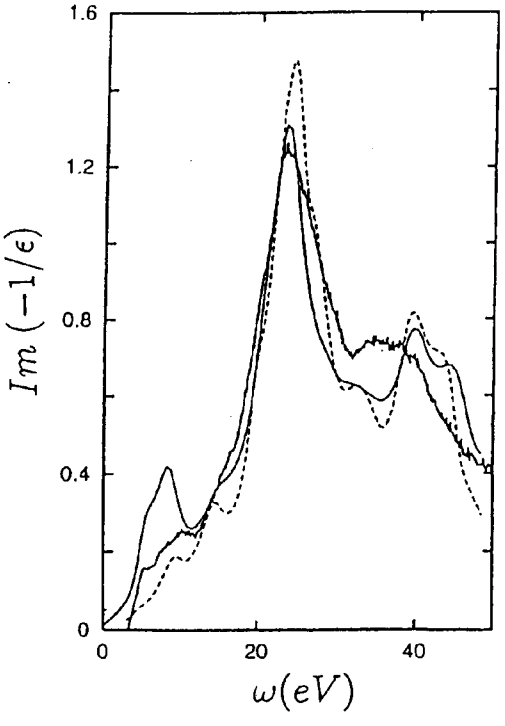

Figure 8

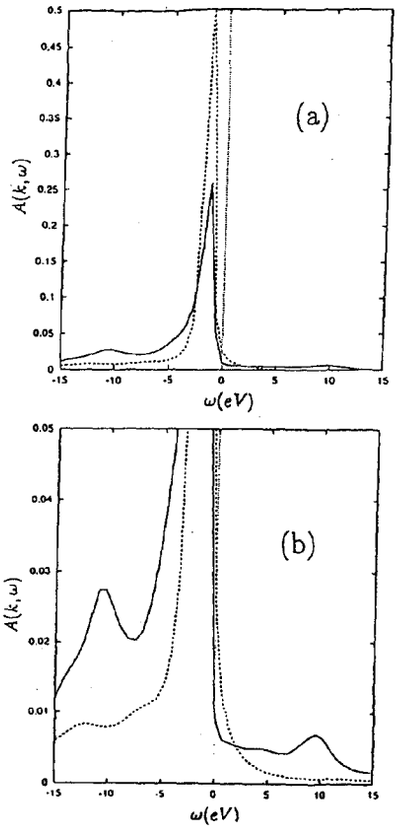

Figure 9

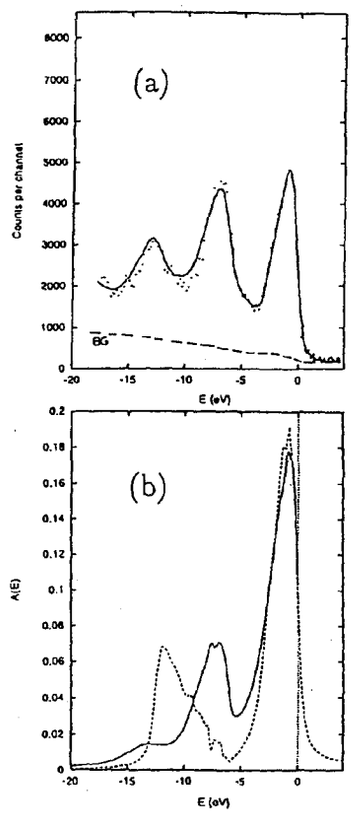

Figure 11

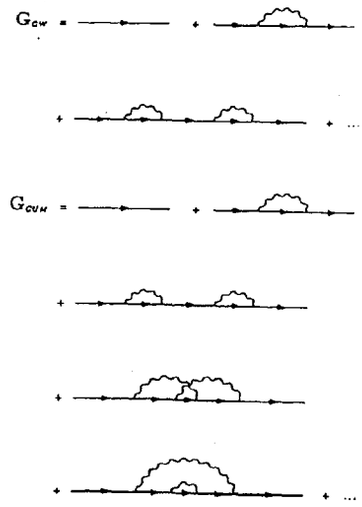

Figure 10

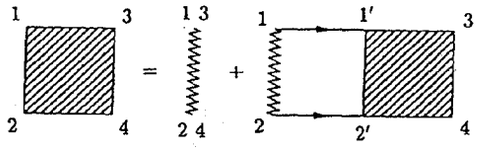

Figure 12

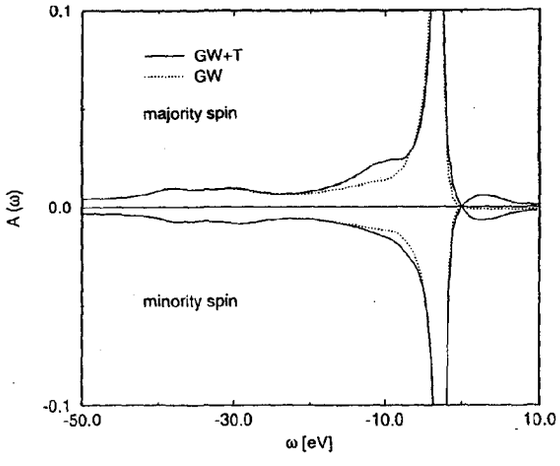

Figure 13

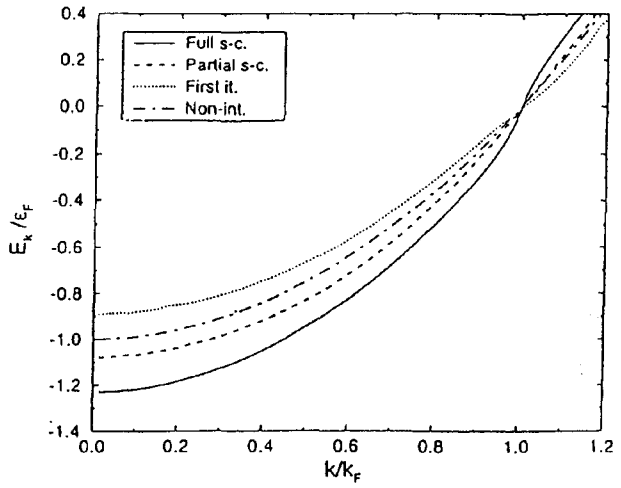

Figure 14

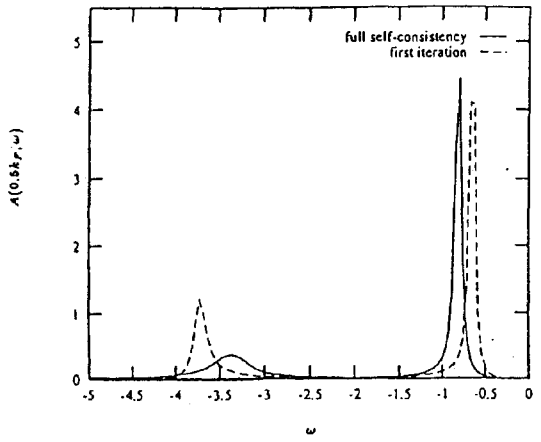

Figure 15

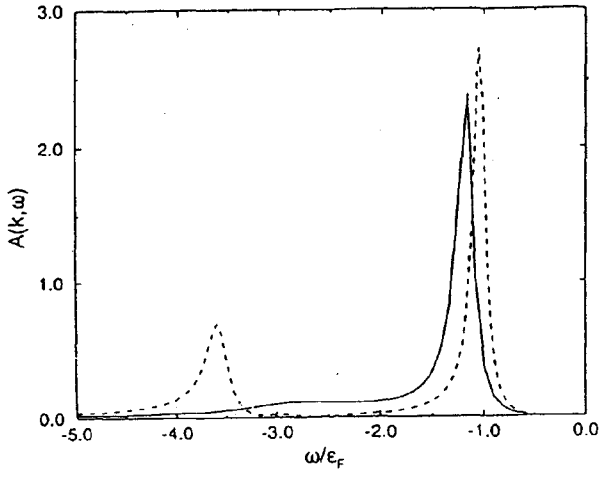

Figure 16

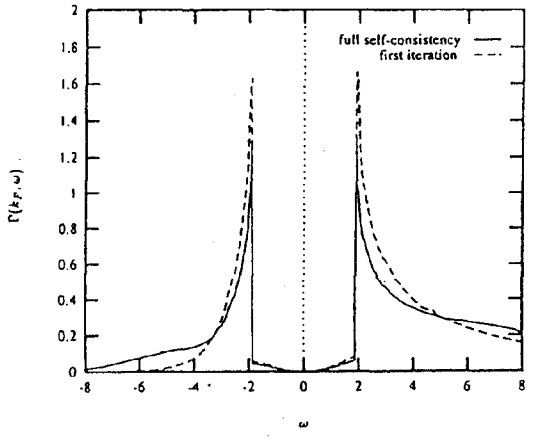

Figure 17

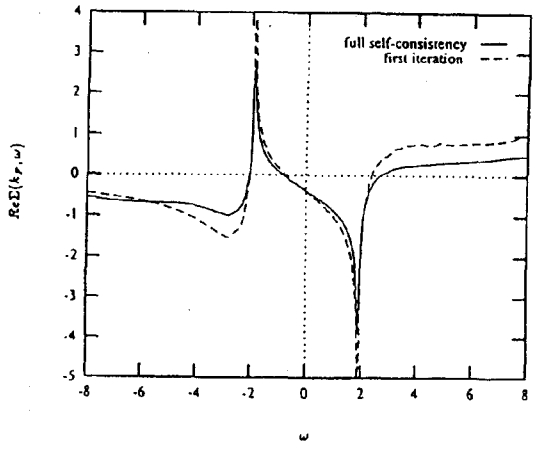

Figure 18