Self-energy of ferromagnetic nickel in the GW approximation

F. Aryasetiawan*

Institute for Theoretical Physics, University of Lund, Sölvegatan 14A, S-223 62 Lund, Sweden

(Received 29 April 1992)

The one-electron excitation spectra of ferromagnetic nickel have been obtained from a first-principles calculation of the self-energy operator within the so-called GW approximation. The dielectric matrix, needed to form the screened potential W, is computed within the random-phase approximation. The quasiparticle energies are in very good agreement with angle-resolved photoemission data. The bottom of the d band is raised by about 1 eV resulting in band narrowing as observed experimentally and the quasiparticle widths are also in favorable agreement with experiment. The exchange splittings, however, are the same for most cases as those given by the local-density approximation in density-functional theory. The satellite at 6 eV is not reproduced. Instead, we found a significant contribution to the spectral weight from quasiparticle peaks around that energy. We discuss the success and the shortcomings of the GW approximation in the light of our results.

I. INTRODUCTION

Measurements of one-electron excitation spectra by high-energy photoemission and inverse-photoemission experiments on crystalline solids show in most cases the presence of well-defined although broadened peaks at certain energies. These peaks are called quasiparticle peaks. In some cases, we may get additional peaks arising from many-body interactions, but they are usually weaker and we refer to them as satellite structures. When the positions of the main peaks are plotted as a function of the momenta of the emitted electrons, we obtain what are termed the band structures of the solids.

A theoretical tool for calculating the positions and widths of the quasiparticles is provided by a many-body perturbation theory in which the Green function is the basic quantity.[1] To obtain the Green function, we require knowledge of the electronic self-energy or mass operator which is in general nonlocal and energy dependent. The self-energy operator acts like a potential in a Schrödinger-like equation, known as the Dyson equation, whose solution gives the Green function. Thus the self-energy operator is a basic quantity needed in the calculations of one-electron excitation spectra.

In practice, the exact self-energy operators for real systems are impossible to compute and we always have to resort to approximate forms. Perhaps the simplest approximation would be to replace the self-energy with that of the homogeneous electron gas and neglect its nonlocality and energy dependence.[2] This approximation leads to the so-called local-density approximation (LDA) of the exchange and correlation potential, v^{xc}, which is identical to the local-density approximation in the density functional theory (DFT).[3,4] The exact v^{xc} in DFT may also be thought of as an approximation to the self-energy operator; it is a local, energy independent self-energy operator that yields the exact ground-state density.

The LDA is very appealing because the local nature of the v^{xc} yields a set of single-particle equations which are numerically much easier to solve than, e.g., the integro-differential equations in the Hartree-Fock approximation.

Due to its simplicity, it has been applied to a wide class of systems and in many cases with surprisingly good results, particularly in simple metals but even in systems with relatively localized electrons, systems which show no resemblance to the homogeneous electron gas. This suggests that for these systems, the self-energy operator is short ranged, but the fact that it may be approximated by an energy independent v^{xc} is due to subtle cancellations between the effects of the strong energy dependence and of the nonlocality.

There are, however, cases where the LDA gives results which differ considerably from experiments. Notable among these is the so-called "band-gap problem" in semiconductors where the LDA yields band gaps which are 30–100 % too small[5,6] compared to the experimental values. Even the exact DFT v^{xc}'s are unlikely to significantly reduce the discrepancies, as previous studies on the exact v^{xc} for small atoms have shown.[7,8] The work of Godby, Schlüter, and Sham[9] on semiconductors suggests that the large parts of the discrepancies are more likely to originate from the discontinuity in the v^{xc} as a functional of the particle number[10] than from the difference between the LDA and exact v^{xc}. If we do want to use the LDA, a more sound procedure would be to compute total energy differences between the N- and the $(N\pm1)$-particle systems.

The failure of the LDA indicates the importance of nonlocality and dynamic correlations in describing quasiparticles. A more realistic but relatively simple approximation to the self-energy, which takes into account both nonlocality and dynamic correlations, was developed in the early 1960s by Hedin,[1,11] known as the GW approximation (GWA). This approximation was originally derived from a many-body perturbation theory as a first term in the expansion of the self-energy in the screened Coulomb potential W, rather than the bare Coulomb potential v. The expansion is in general divergent, in the sense that the second-order term is not necessarily smaller than the first-order term. However, the use of the approximation is physically motivated if we view it as a Hartree-Fock approximation (Gv) with the bare

Coulomb potential v replaced by the dynamically screened one.

Despite the theoretical simplicity of the GWA, its applications to real systems have been hampered by the large size of the computations. Due to the progress in computer technology, it is now feasible to apply the GWA to some real systems. Recently, Hybertsen and Louie,[12,13] Godby, Schlüter, and Sham,[9] and von der Linden and Horsch[14] performed self-energy calculations for semiconductors within the GWA with encouraging results: the large discrepancies in the band gaps are removed.

The success of the GWA in semiconductors and several other systems leads us to test it in systems with more complicated electronic structures. An interesting class of systems is one where a narrow band, resulting from localized electrons, is imbedded in a free-electron-like band. Transition metals form such a class and among the $3d$ series, nickel provides the most anomalous case. More extreme cases of this class are found in heavy-Fermion systems where the f electrons form very narrow bands.

As in the case of semiconductors, the LDA band structure of nickel[15] deviates significantly from angle-resolved photoemission data.

The width of the occupied $3d$ band is 30% smaller than that of the LDA (3.3 eV vs 4.5 eV).[16,17]

The exchange splitting for states at the top of the occupied band is half the LDA value (0.25–0.35 eV vs 0.6 eV).[17,18]

A satellite at ~ 6 eV below the Fermi level which resonates at the $3p$ threshold and shows no dispersion is observed experimentally but entirely missing in LDA.[16,19]

On the other hand, good agreement with experiment is found for the ground-state properties such as equilibrium lattice constant, bulk modulus, and magnetic moment, with the exception of the cohesive energy where the LDA value is about 1 eV too small.[20]

The above discrepancies clearly show the breakdown of the LDA and necessitate the inclusion of nonlocality and dynamic correlations in the self-energy operator. Another indication of the importance of many-body effects in nickel is demonstrated by the relatively large widths of the quasiparticles—up to 2 eV at the bottom of the band[18]—which imply short lifetimes. Generally speaking, a long lifetime justifies the use of single-particle theory where the many-body state is approximated by a single-Slater determinant whereas a short lifetime implies a strong interaction between the quasiparticle and the rest of the system which changes the quasiparticle state and results in its decay.

There are several features of nickel which may give a qualitative explanation for the inadequacy of the LDA in describing quasiparticle properties.

(1) Nickel is ferromagnetic with a fully occupied majority spin channel and a partially occupied minority spin channel giving a magnetic moment of $0.59\mu_B$/atom.[15]

(2) The density of states around the Fermi level is very large, about 1.7 states/eV.[20]

(3) The d states are very localized, more than 95% of the states are confined to the muffin-tin sphere (Table I) resulting in a narrow-band width (~ 3.3 eV).

(4) The correlation strength measured in terms of the ratio between the Hubbard U and the width of the band is intermediate.

(5) The two lowest atomic configurations $3d^9 4s$ and $3d^8 4s^2$ are almost degenerate, differing by only 0.025 eV.[21]

(6) The LDA eigenvalues for the d states are lower than the experimental values, whereas in most cases the LDA eigenvalues are higher than the experiment.[17,18]

The localized nature of the d states suggests that we may use an atomic picture in explaining the characteristic properties of nickel, but on the other hand, de Haas–van Alphen measurements[22,23] clearly show the existence of a Fermi surface. The itinerant character of the d states should therefore be taken into account. This is important when we consider screening of a hole left behind during a photoemission experiment. In the itinerant picture, electrons from neighboring atoms can participate in the screening whereas there is no such possibility in the atomic picture. But in many cases, the atomic picture should be sufficient and this is supported by experiment on NiO,[24] which is an insulator with localized $3d$ electrons and which indeed shows similar features as those of nickel.

The existence of the satellite at 6 eV below the Fermi level and the reduction in the exchange splittings are attributed in a Hubbard model[25,26] to a two-hole bound state which has a long lifetime due to the localized character of the d electrons and the presence of unoccupied states just above the Fermi level.

Previous works on the self-energy of nickel have been based mainly on Hubbard models, treated either with the t matrix formulation[25,26] or second-order perturbation theory.[27] These models are useful as a means of identifying important physical processes responsible for band narrowing, satellite structure, etc., but the existence of the Hubbard parameters does not fully justify direct comparisons with experiments. On the other hand, the GWA provides a more proper way for the calculation of the exchange-correlation operator. The inclusion of dynamic screening in the GWA should result in band narrowing, whereas its ability to account for the satellite and the reduction in the exchange splittings is doubtful if the above explanation is correct. The purpose of the present paper is to investigate the ability of the GWA to describe

TABLE I. The self-energies at the X point in eV for majority and minority spin (alternately). The charge of each state inside the muffin-tin sphere is denoted by "charge" and Z is given by Eq. (43).

kn	Charge	Σ^{xc}	v^{xc}	Z
X_1	0.862	-24.7	-27.4	0.73
	0.853	-23.7	-26.4	0.72
X_3	0.936	-27.2	-30.6	0.64
	0.932	-26.1	-29.7	0.63
X_2	0.996	-34.8	-37.7	0.50
	0.998	-34.1	-36.9	0.51

systems with relatively strong correlations with nickel as a prototype. In view of the above discussions, a first-principle calculation of nickel self-energy within the GWA has been performed. It also forms a preliminary study for more complicated nickel compounds such as NiO which has vanishingly small gap in the LDA.

To do a GW calculation, we start from a zeroth-order Hamiltonian. Since we may view the GWA as a first-order perturbation theory, it is natural to choose a starting Hamiltonian which is as close as possible to the real system. The LDA Hamiltonian is a reasonable choice and this has also been used in previous works on semiconductors with good results.

Self-energy calculation of nickel is a major computational challenge. One of the main problems is choosing a suitable set of single-particle basis functions for the many-body perturbation calculation in the GWA. In the case of semiconductors, plane waves provide a good basis since the valence electrons can be described very well by a pseudopotential. They are also a natural basis for the Coulomb potential and the evaluation of matrix elements becomes very simple. The large and important bare exchange part of the self-energy can then be evaluated with high accuracy. In contrast, nickel has a strong potential which rules out the possibility of using a pseudopotential. Clearly, plane waves are not a good choice since a large number of them is required to describe core oscillations. Instead we have chosen to use a modified linearized augmented plane-wave (LAPW) basis[28] to do our perturbation calculation. The advantage of this basis is that a relatively small number of basis functions is needed because they are constructed from radial wave functions which are solutions to the Schrödinger equation inside the muffin-tin sphere. The disadvantage, however, is that the evaluation of matrix elements becomes complicated. We have developed a simple but efficient method for evaluating these which is described in the previous paper.[28]

The second computational problem is the k-space integration. In semiconductors only a few k points are needed due to the absence of Fermi surface but in nickel more points are required to take into account the Fermi surface. Surprisingly we found that a relatively small number of k points (20 in the irreducible wedge) is sufficient to get an accuracy of 0.1–0.2 eV in the quasiparticle energies. Since we are interested in the relative quasiparticle energies rather than their absolute values, this appears to be the result of a systematic cancellation of errors.

The paper is organized as follows: theory and numerical procedure are described in Secs. II and III, respectively. Results and discussions are presented in Sec. IV. Section V is reserved for conclusions.

II. THEORY

The quasiparticles

The one-particle Green function is defined by[29]

$$G(xt,x't')= -i\langle N|[T\{\hat{\psi}(xt)\hat{\psi}^\dagger(x't')\}]|N\rangle ,$$

$$x \equiv \mathbf{r},\sigma . \quad (1)$$

The field operators are defined in the Heisenberg picture. For $t>t'$ the Green function describes the motion of an added particle from $\mathbf{r}',t'$ to $\mathbf{r},t$ with a possible spin flip σ' to σ whereas for $t<t'$ it describes the motion of an added hole from $\mathbf{r},t$ to $\mathbf{r}',t'$, also with a possible spin flip σ to σ'. The time ordering operator is simply a convenient way of treating both particle and hole at the same time.

From the Heisenberg equation of motion for operators it can be shown that the Green function satisfies the Dyson equation

$$[\omega - H^0(x;\omega)]G(x,x';\omega) - \int dx''\Sigma(x,x'';\omega)G(x'',x';\omega)$$

$$=\delta(x-x') . \quad (2)$$

A solution to the above equation may be obtained from the classical theory of Green function.[30] We define $g_s(x;\omega)$ to be the solution to

$$[E_s(\omega)-H^0(x;\omega)]g_s(x;\omega) - \int dx'\Sigma(x,x';\omega)g_s(x';\omega)$$

$$=0 . \quad (3)$$

The eigenvalues $E_s(\omega)$ are in general complex and ω may be treated as a parameter. The Green function is given by

$$G(x,x';\omega)= \sum_s \frac{g_s(x;\omega)g_s^\dagger(x';\omega)}{\omega - E_s(\omega)} , \quad (4)$$

where $g_s^\dagger$ are the corresponding "left-hand" eigenfunctions. When the self-energy operator is Hermitian, g_s and $g_s^\dagger$ are complex conjugate of one another and the g_s's are orthogonal, but in general they are not.

For crystalline solids, the Green function is normally represented in some Bloch basis $\phi_{\mathbf{k}n}$ which we assume to be orthonormal. We form the spectral function

$$A(\mathbf{k},\omega)= -\frac{1}{\pi} \sum_n \mathrm{Im}G_{nn}(\mathbf{k},\omega) , \quad (5)$$

where $G_{nn'}(\mathbf{k},\omega)$ is the matrix element of G in the Bloch basis. We note that since k is a good quantum number, $A(\mathbf{k},\omega)$ is independent of the Bloch basis. We now consider Eq. (3) and suppose that at some $\omega=\omega_s$, we find that

$$\omega_s = \mathrm{Re}E_s(\omega_s) . \quad (6)$$

It follows from Eqs. (5) and (4) that the spectral function $A(\mathbf{k},\omega)$ has a peak at $\omega=\omega_s$. We identify this peak as a quasiparticle peak and define the wave function $g_s(x;\omega_s)$ as the quasiparticle wave function.

It may happen that Eq. (6) has no solution which implies that Eq. (3) has no solution either. In this case, we find $\omega=\omega_s$ that minimizes $|\omega_s - \mathrm{Re}E_s(\omega_s)|$ and we should still find a peak in the spectral function around ω_s. We identify ω_s as the quasiparticle energy and define the corresponding quasiparticle wave function to be that g_s which satisfies Eq. (3) with ω_s as the eigenvalue. For real systems, however, such a case is unlikely to occur. In the nickel case, we did not encounter any difficulties in finding solutions to Eq. (6). A practical way of finding the quasiparticle energy, however, is provided by Eq. (11) below.

The quasiparticle concept may be understood in a more physical way by considering an exact expression for the one-electron spectral function (Ref. 1):

$$A(x,x';\omega) = \sum_s f_s(x) f_s^\dagger(x') \delta(\omega - \varepsilon_s) , \qquad (7)$$

where

$$f_s(x) = \langle N | \psi(x) | N+1, s \rangle ,$$
$$\varepsilon_s = E_{N+1,s} - E_N \text{ for } \varepsilon_s \geq \mu , \qquad (8)$$

$$f_s(x) = \langle N-1, s | \psi(x) | N \rangle ,$$
$$\varepsilon_s = E_N - E_{n-1,s} \text{ for } \varepsilon_s < \mu . \qquad (9)$$

ε_s are excitations energies of the $(N-1)$- and $(N+1)$-particle systems. We consider a noninteracting Hamiltonian H^0 which may be chosen to be the LDA Hamiltonian and form the following quantity:

$$A_n(\mathbf{k},\omega) = \sum_s |\langle f_s | \psi_{\mathbf{k}n} \rangle|^2 \delta(\omega - \varepsilon_s) . \qquad (10)$$

Without the self-energy correction, one of the functions f_s's say f_p, is identical to the Bloch state $\psi_{\mathbf{k}n}$ and the above expression reduces to a δ function centered at the LDA eigenvalue $\varepsilon_p = \varepsilon_{\mathbf{k}n}^{\text{LDA}}$. If we now take into account the self-energy correction, the function f_p is no longer identical to $\psi_{\mathbf{k}n}$. Consequently, the matrix element $\langle f_p | \psi_{\mathbf{k}n} \rangle$ is reduced from unity since other f_s's will give nonzero matrix elements, resulting in the broadening of the δ function and we might even get a satellite structure. Provided $\psi_{\mathbf{k}n}$ is a good starting wave function, we expect that the broadened δ function is still centered around ω_p and we identify the peak as the quasiparticle peak. If we sum over the index n in Eq. (10), we get the spectral function, identical to that in Eq. (5).

In terms of the noninteracting Green function and the self-energy, the spectral function is formally given by

$$\mathbf{A}(\omega) = \frac{1}{\pi} |\text{Im}\mathbf{G}(\omega)|$$

$$= \frac{1}{\pi} \frac{|\text{Im}\Sigma(\omega)|}{|\text{Re}[(\mathbf{G}^0)^{-1}(\omega) - \Sigma(\omega)]|^2 + |\text{Im}\Sigma(\omega)|^2} . \qquad (11)$$

Thus the imaginary part of the self-energy is proportional to the lifetime of the quasiparticle. For materials with Fermi surface (metals) the number of excited states can be shown from phase-space argument to vary like $\sim(\varepsilon_s - \mu)^2$ which in turn is proportional to $\text{Im}\Sigma$. Thus quasiparticles close to the Fermi surface, if they exist, have a long lifetime unless when there is a van Hove singularity at the Fermi level which can make the imaginary part of the self-energy behave linearly.

The spectral function $A(\omega)$ in metals can be measured by means of photoemission experiments. Provided the final states of the photoelectrons are plane waves, i.e., the photoelectrons are completely decoupled from the solid, the spectral distributions of the holes left behind do correspond to the theoretical spectral functions if we neglect electron-phonon coupling and extrinsic effects which

cause energy loss of the electron on its way to and through the surface. The presence of well-defined peaks in the spectral function indicates the existence of long lifetime excitations or quasiparticles.

The screened interaction

The screened interaction takes the form

$$W(\mathbf{r},\mathbf{r}';\omega) = \int d^3 r'' \epsilon^{-1}(\mathbf{r},\mathbf{r}'';\omega) v(|\mathbf{r}'' - \mathbf{r}'|) . \qquad (12)$$

The dielectric matrix needed in forming the screened potential W is given by

$$\epsilon^{-1}(\mathbf{r},\mathbf{r}';\omega) = \delta(\mathbf{r}-\mathbf{r}') + \int d^3 r'' v(|\mathbf{r}-\mathbf{r}''|) \chi(\mathbf{r}'',\mathbf{r}';\omega) , \qquad (13)$$

where $\chi(\mathbf{r},\mathbf{r}';\omega)$ is the Fourier transform of the time-ordered density-density response function

$$\chi(\mathbf{r}t,\mathbf{r}'t') = i\{ \langle 0 | T[\hat\rho(\mathbf{r}t)\hat\rho(\mathbf{r}'t')] | 0 \rangle - \langle \rho(\mathbf{r}) \rangle \langle \rho(\mathbf{r}') \rangle \} , \qquad (14)$$

which is computed in the random-phase approximation (RPA):[31]

$$\chi = \chi^0 + \chi^0 v \chi . \qquad (15)$$

In RPA particle number is conserved and therefore the f-sum rule is satisfied. χ^0 is the time-ordered density-density response function for noninteracting electrons which may be written in its spectral representation

$$\chi^0(\mathbf{r},\mathbf{r}';\omega) = \int_{-\infty}^{\infty} d\omega' \frac{S^0(\mathbf{r},\mathbf{r}';\omega')}{\omega - \omega' + i\omega'\delta} , \qquad (16)$$

where the spectral function S^0 is proportional to the imaginary part of χ^0

$$S^0(\mathbf{r},\mathbf{r}';\omega) = -\frac{1}{\pi} \text{Im}\chi^0(\mathbf{r},\mathbf{r}';\omega)\text{sgn}(\omega) \qquad (17)$$

and is given explicitly by

$$S^0(\mathbf{r},\mathbf{r}';\omega) = \sum_{\mathbf{k}n}^{\text{occ}} \sum_{\mathbf{k}'n'}^{\text{unocc}} \psi_{\mathbf{k}n}^*(\mathbf{r}) \psi_{\mathbf{k}n}(\mathbf{r}') \psi_{\mathbf{k}'n'}(\mathbf{r})$$
$$\times \psi_{\mathbf{k}'n'}^*(\mathbf{r}') \delta(\omega - \varepsilon_{\mathbf{k}'n'} + \varepsilon_{\mathbf{k}n}) . \qquad (18)$$

The physical meaning of the RPA is that the electrons respond to the external and induced field as if they were noninteracting. The Bloch states and eigenvalues $\{\psi_{\mathbf{k}n}, \varepsilon_{\mathbf{k}n}\}$ correspond to a zeroth-order noninteracting Hamiltonian chosen to be the LDA:

$$H^0 = T + V^H + v_{\text{LDA}}^{\text{xc}} , \qquad (19)$$

where T is the kinetic energy operator, V^H is the Hartree potential, and $v_{\text{LDA}}^{\text{xc}}$ is the local-density approximation for exchange and correlation potential.

The self-energy

The self-energy is approximated by the first term in its expansion in power of the screened potential W,[1]

$$\Sigma(\mathbf{r},\mathbf{r}';\omega)=i\int\frac{d\omega'}{2\pi}G(\mathbf{r},\mathbf{r}';\omega+\omega')W(\mathbf{r},\mathbf{r}';\omega')e^{i\delta\omega'} \ .$$

$$(20)$$

It is convenient and physically appealing to divide the self-energy into the frequency independent bare exchange (Fock) term and the frequency-dependent correlated part which contains the dynamical effects of correlation:

$$\Sigma=\Sigma^x+\Sigma^c$$

$$=Gv+GW^c \ , \qquad (21)$$

where $W^c=W-v$. The GWA is similar in form to the Hartree-Fock approximation. One can think of the GWA as a Hartree-Fock approximation with a dynamically screened interaction rather than as a perturbation theory.

Like χ, the quantities W^c and Σ^c obey the Kramers-Kronig relations

$$W^c(\omega)=\int_{-\infty}^{\infty}d\omega'\frac{B(\omega')}{\omega-\omega'+i\omega'\delta} \ , \qquad (22)$$

$$\Sigma^c(\omega)=\int_{-\infty}^{\infty}d\omega'\frac{\Gamma(\omega')}{\omega-\omega'+i(\omega'-\mu)\delta} \ , \qquad (23)$$

where

$$B(\omega)=-\frac{1}{\pi}\mathrm{Im}W^c(\omega)\mathrm{sgn}(\omega) \ , \qquad (24)$$

$$\Gamma(\omega)=-\frac{1}{\pi}\mathrm{Im}\Sigma^c(\omega)\mathrm{sgn}(\omega-\mu) \ . \qquad (25)$$

In the above equations the space variables $\mathbf{r},\mathbf{r}'$ have been suppressed for simplicity.

The Green function G in W and Σ is in principle a self-consistent G obtained from the Dyson equation

$$G=G^0+G^0\Sigma G \ . \qquad (26)$$

To calculate the self-consistent G is beyond our present computational capability so that in practice we use in Σ the zeroth-order Green function G^0 with its spectral function given by

$$A^0(\mathbf{r},\mathbf{r}';\omega)=\sum_{\mathbf{k}n}\psi_{\mathbf{k}n}(\mathbf{r})\psi_{\mathbf{k}n}^*(\mathbf{r}')\delta(\omega-\varepsilon_{\mathbf{k}n}) \ . \qquad (27)$$

With $G=G^0$ the spectral function of the self-energy is given by $\Gamma=\Gamma^h+\Gamma^e$ where

$$\Gamma^h(\mathbf{r},\mathbf{r}';\omega)=\sum_{\mathbf{k}n}^{\mathrm{occ}}\psi_{\mathbf{k}n}(\mathbf{r})\psi_{\mathbf{k}n}^*(\mathbf{r}')B(\mathbf{r},\mathbf{r}';\varepsilon_{\mathbf{k}n}-\omega)\theta(\varepsilon_{\mathbf{k}n}-\omega) \ ,$$

$$\omega\leq\mu \ , \qquad (28)$$

$$\Gamma^e(\mathbf{r},\mathbf{r}',\omega)=\sum_{\mathbf{k}n}^{\mathrm{unocc}}\psi_{\mathbf{k}n}(\mathbf{r})\psi_{\mathbf{k}n}^*(\mathbf{r}')B(\mathbf{r},\mathbf{r}';\omega-\varepsilon_{\mathbf{k}n})\theta(\omega-\varepsilon_{\mathbf{k}n}) \ ,$$

$$\omega>\mu \ . \qquad (29)$$

Γ^h and Γ^e denote the hole and electron parts, respectively. We note that the GWA is a conserving approximation provided we have a self-consistent G.[32,33] With the approximation $G=G^0$ particle number is still conserved but energy and momentum ar not necessarily conserved anymore.

III. NUMERICAL PROCEDURE

We start with a self-consistent LAPW band structure calculation of spin polarized nickel within the local-density approximation.[34−36] The resulting self-consistent potential is then used to construct the basis functions which are described in detail in a previous paper.[28] These basis functions are subsequently used to solve the secular equation to generate a set of orthonormal Bloch states with a corresponding set of eigenvalues: $\{\psi_{\mathbf{k}n},\varepsilon_{\mathbf{k}n}\}$. An energy cutoff of 9 Ry is used giving on average 40 Bloch states. These together with the nine core states $(1s^12s^12p^33s^13p^3)$ represent the zeroth-order approximation in our many-body perturbation calculation and form the basis for the representation of ϵ^{-1}, W, and Σ.

The next step is to calculate the inverse of the dielectric matrix ϵ^{-1}. For details the reader is referred to an earlier work.[28] The main quantity to be computed is the spectral function S^0 which in the Bloch-state basis is given by

$$S_{ll'}^0(\mathbf{q};\omega)=\sum_{\mathbf{k}n}^{\mathrm{occ}}\sum_{n'}^{\mathrm{unocc}}\langle\mathbf{q}l,\mathbf{k}n|\mathbf{k}+\mathbf{q}n'\rangle$$

$$\times\langle\mathbf{k}+\mathbf{q}n'|\mathbf{k}n,\mathbf{q}l'\rangle$$

$$\times\delta(\omega-\varepsilon_{\mathbf{k}+\mathbf{q}n'}+\varepsilon_{\mathbf{k}n}) \ . \qquad (30)$$

The computation of the matrix elements

$$\langle\mathbf{q},l,\mathbf{k}n|\mathbf{k}+\mathbf{q}n'\rangle=\int d^3r\,\psi_{\mathbf{q}l}^*\psi_{\mathbf{k}n}^*\psi_{\mathbf{k}+\mathbf{q}n'} \qquad (31)$$

forms a major computational effort. A simple scheme has been developed for a fast and accurate evaluation of these matrix elements.[28] Brillouin zone integration is done with trapezoidal rule. The δ function is replaced by a Gaussian with a width of 0.1 Ry and the numbers of points in the irreducible wedge is 20. Convergence test shows very little change in the energy loss function $\epsilon^{-1}(\mathbf{q};\omega)$ when the number of $\mathbf{k}$ points is increased to 89 and the width of the Gaussian is changed up to 0.3 Ry. For the purpose of obtaining the self-energy this is more than adequate since the screened potential $W=\epsilon^{-1}v$ is convoluted with the Green function G. Experience with semiconductors shows that even plasmon pole approximation gives very good self-energy.[13] The Hilbert transform in Eq. (16) is found to converge with a frequency mesh of 0.04 Ry.

Once we have obtained the screened potential W, we are in a position to calculate the self-energy. The self-energy is divided into the bare exchange Σ^x and the correlated part Σ^c. We first describe the evaluation of Σ^c. In the Bloch-state basis, the spectral function for the self-energy in Eqs. (28) and (29) is given by

$$\Gamma_{ll'}^h(\mathbf{q};\omega)=\sum_{\mathbf{k}n}^{\mathrm{occ}}\sum_{m,m'}\langle\mathbf{q}l,\mathbf{k}-\mathbf{q}n|\mathbf{k}m\rangle B_{mm'}(\mathbf{k};\omega-\varepsilon_{\mathbf{k}-\mathbf{q}n})$$

$$\times\langle\mathbf{k}m'|\mathbf{k}-\mathbf{q}n,\mathbf{q}l'\rangle$$

$$\times\theta(\varepsilon_{\mathbf{k}-\mathbf{q}n}-\omega) \ , \qquad (32)$$

$$\Gamma^e_{ll'}(\mathbf{q};\omega) \sum_{kn}^{\text{unocc}} \sum_{m,m'} \langle \mathbf{q}l, \mathbf{k}-\mathbf{q}n | \mathbf{k}m \rangle B_{mm'}(\mathbf{k};\omega - \varepsilon_{\mathbf{k}-\mathbf{q}n})$$

$$\times \langle \mathbf{k}m' | \mathbf{k}-\mathbf{q}n, \mathbf{q}l' \rangle$$

$$\times \theta(\omega - \varepsilon_{\mathbf{k}-\mathbf{q}n}) . \tag{33}$$

Comparison with Eq. (31) shows that the evaluation of Γ involves an extra summation over the band and the frequency ω extends to twice the maximum energy cutoff.

$$\Sigma^x_{ll'}(\mathbf{q}) = -\sum_{kn}^{\text{occ}} \sum_{\mathbf{R}} \int_\Omega d^3r \int_\Omega d^3r' \frac{\psi^*_{\mathbf{q}l}(\mathbf{r})\psi_{\mathbf{k}n}(\mathbf{r})\psi^*_{\mathbf{k}n}(\mathbf{r}')\psi_{\mathbf{q}l'}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'+\mathbf{R}|} e^{-i(\mathbf{q}-\mathbf{k})\cdot\mathbf{R}} , \tag{34}$$

where $\mathbf{R}$ is a lattice vector and Ω is the unit cell volume. The summation over $\mathbf{R}$ may be carried out with the Ewald method but the summation over $\mathbf{k}$ and n makes the computation disproportionately large.

A simple way of avoiding the summation over $\mathbf{R}$ would be to express v in the Bloch state basis

$$v(|\mathbf{r}-\mathbf{r}'|) = \sum_{knn'} \psi_{\mathbf{k}n}(\mathbf{r}) v_{nn'}(\mathbf{k}) \psi^*_{\mathbf{k}n'}(\mathbf{r}') , \tag{35}$$

so that we obtain

$$\Sigma^x_{ll'}(\mathbf{q}) = -\sum_{kn}^{\text{occ}} \langle \mathbf{q}l, \mathbf{k}-\mathbf{q}n | \mathbf{k}m \rangle v_{mm'}(\mathbf{k})$$

$$\times \mathbf{k}m' | \mathbf{k}-\mathbf{q}n, \mathbf{q}l' \rangle . \tag{36}$$

However, a significant error (as large as 1 eV) in the sp band is introduced due to the finite number of basis functions. This arises from the different nature of the cancellations of errors between the localized d states and the extended sp states. We can of course increase the number of basis functions but this is computationally disadvantageous. To improve on the bare exchange, we use the following procedure: we write

$$v(|\mathbf{r}-\mathbf{r}'|) = \bar{v}(|\mathbf{r}-\mathbf{r}'|) + \Delta v(|\mathbf{r}-\mathbf{r}'|), \quad \Delta v \equiv v - \bar{v} , \tag{37}$$

where $\bar{v}$ is given by Eq. (35). We then expand Δv in a set of plane waves and use it to calculate the correction to the bare exchange which results in

$$\Delta\Sigma^x_{ll'}(\mathbf{q}) = \sum_{kn}^{\text{occ}} \sum_{GG'} \langle \mathbf{q}l, \mathbf{k}-\mathbf{q}n | \mathbf{k}+\mathbf{G} \rangle \Delta v_{GG'}(\mathbf{k})$$

$$\times \langle \mathbf{k}+\mathbf{G}' | \mathbf{k}-\mathbf{q}n, \mathbf{q}l' \rangle , \tag{38}$$

where

$$\Delta v_{GG'}(\mathbf{k}) = \frac{4\pi\delta_{GG'}}{(\mathbf{k}+\mathbf{G})^2}$$

$$-\sum_{nn'} \langle \mathbf{k}+\mathbf{G} | \mathbf{k}n \rangle v_{nn'}(\mathbf{k}) \langle \mathbf{k}n' | \mathbf{k}+\mathbf{G}' \rangle . \tag{39}$$

$v_{nn'}(\mathbf{k})$ are the matrix elements of v in the Bloch-state basis. With this exchange correction, the sp band is in much better agreement with experiment.

The real part is obtained by Hilbert transform as in Eq. (23). Brillouin zone integration is done with 20 points in the irreducible wedge and a change of less than 0.1 eV is observed when the number of points is increased to 89.

Numerical evaluation of Σ^x presents some difficulties due to its magnitude which is typically 5–10 times larger than Σ^c. This means Σ^x must be computed with one significant figure more than Σ^c. Direct evaluation of Σ^x in the Bloch-state basis is numerically prohibitive as may be seen below:

Having obtained the self-energy, the quasiparticle energies are easily found:

$$E_{\mathbf{k}n} = \varepsilon_{\mathbf{k}n} + \Delta\Sigma_{nn}(\mathbf{k};E_{\mathbf{k}n})$$

$$= \varepsilon_{\mathbf{k}n} + \Delta\Sigma_{nn}(\mathbf{k};\varepsilon_{\mathbf{k}n}) + (E_{\mathbf{k}n} - \varepsilon_{\mathbf{k}n})\frac{\partial\Delta\Sigma_{nn}(\varepsilon_{\mathbf{k}n})}{\partial\omega} , \tag{40}$$

where

$$\Delta\Sigma_{nn}(\mathbf{k};\omega) = \langle \mathbf{k}n | \Sigma^x(\omega) + \Sigma^c(\omega) - v^{xc} | \mathbf{k}n \rangle . \tag{41}$$

The self-energy correction is given by

$$\Delta\varepsilon_{\mathbf{k}n} = E_{\mathbf{k}n} - \varepsilon_{\mathbf{k}n}$$

$$= Z_{\mathbf{k}n}\Delta\Sigma_{nn}(\mathbf{k};\varepsilon_{\mathbf{k}n}) , \tag{42}$$

where

$$Z_{\mathbf{k}n} = \left[1 - \frac{\partial\Delta\Sigma_{nn}(\mathbf{k};\varepsilon_{\mathbf{k}n})}{\partial\omega} \right]^{-1} < 1 . \tag{43}$$

IV. RESULTS AND DISCUSSIONS

We first present the results and describe their main features. Later we will discuss and attempt to give a unified explanation for the results

The real and imaginary parts of the correlated part of the self-energies for the Γ'_{25} state, X'_4 state, and L'_2 state are shown in Figs. 1–3. The imaginary parts show a parabolic behavior around the Fermi level which is in accordance with theory and typical of a Fermi liquid. The two-peak structure about 20 and 30 eV below the Fermi level originates from a similar structure in the imaginary part of the inverse dielectric function.[28] These two peaks are most likely to be the plasmon peaks. Estimate of the plasmon frequency from the electron gas formula $\omega^2_{\text{plasmon}} = 4\pi n$ with the density n taken to be the average valence density gives $\omega_{\text{plasmon}} = 30.8$ eV, in close agreement with the position of the large peak in the screened potential. The little peak just below the Fermi level originates from the first peak in the dielectric function around
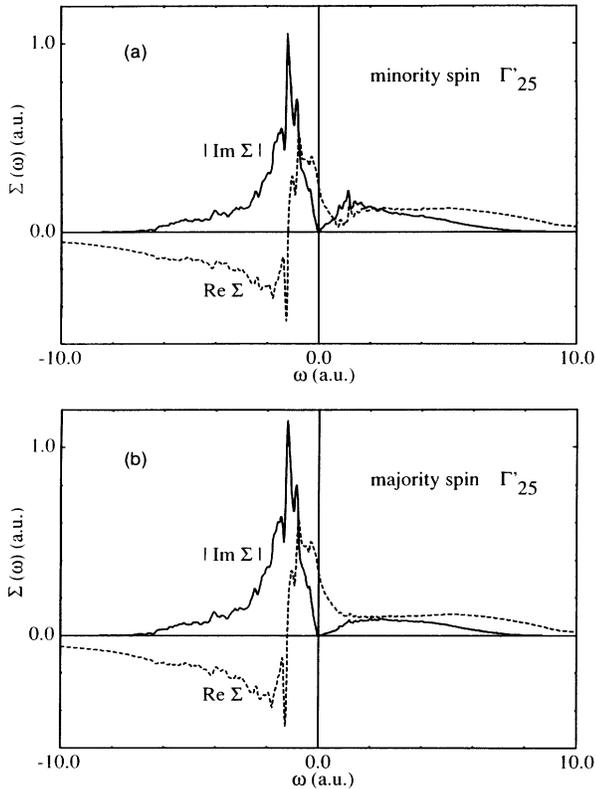
FIG. 1. (a) The real and imaginary parts of the correlated part of the self-energy for the minority spin state Γ'_{25}. (b) The real and imaginary parts of the correlated part of the self-energy for the majority spin state Γ'_{25}.

5–6 eV and may be associated with transitions from a relatively large number of s states about 5–6 eV below the Fermi level to states just above the Fermi level. But as we will discuss later, the peak is not strong enough to produce a satellite structure.

The large frequency behaviors of the imaginary parts appear to be rather similar for both the hole and particle parts. As a consequence, most of the contribution to the real parts, which are the Hilbert transform of the imaginary parts, comes mainly from frequency regions 50 eV below and above the Fermi level, since the large frequency parts of the hole and particle parts cancel one another. This is in agreement with the physical picture, where the most important contribution to the self-energy comes from the energy region below the plasmon frequency. This can be viewed as a justification for the use of energy cutoff of 9 Ry in the basis functions.

The real parts of the self-energies show large derivatives around the Fermi level. Self-consistency is therefore important and we should take the derivatives into account when computing the quasiparticle energies. A typical Z value is between 0.5 and 0.7, which is of similar magnitude as in simple metals, semiconductors, and the electron gas at metallic densities. This range of values appears to be material independent.

In Table I we present the self-energies at the X point which are representative for other points in the Brillouin zone (BZ). The self-energies were calculated without the

exchange corrections as described in Sec. III. The charge inside the muffin-tin sphere is a measure of the degree of localization of the state. We see a clear relationship between the Σ and the degree of localization: the more localized the state the higher the effect of exchange and correlation. This may be understood in terms of the local-density picture where the exchange-correlation potential varies like $\sim n^{1/3}$. A similar relationship holds between Z and localization. The quantity $1/Z$ gives the effective-mass ratio: the more localized the state, the flatter the band and therefore the larger the effective mass.

The average band structure of the up and down spin along the ΓX and ΓL is plotted in Fig. 4. The bottom of the d band along the ΓX has a deviation from experiment ranging from nothing at the Γ point to the 0.4 eV at the zone edge. One might attribute the discrepancy to the Bloch state being inaccurate but, on the other hand, the starting eigenvalue is very accurate and therefore the corresponding Bloch state should be accurate too. We postpone for the moment a possible explanation for this discrepancy.

The band structure of the d states is quite complicated and as may be seen in Fig. 4, the self-energy correction is strongly state dependent. For example, the self-energy correction to the Γ'_{25} state is positive whereas at the L'_2 state it is negative. This point has been brought up be-
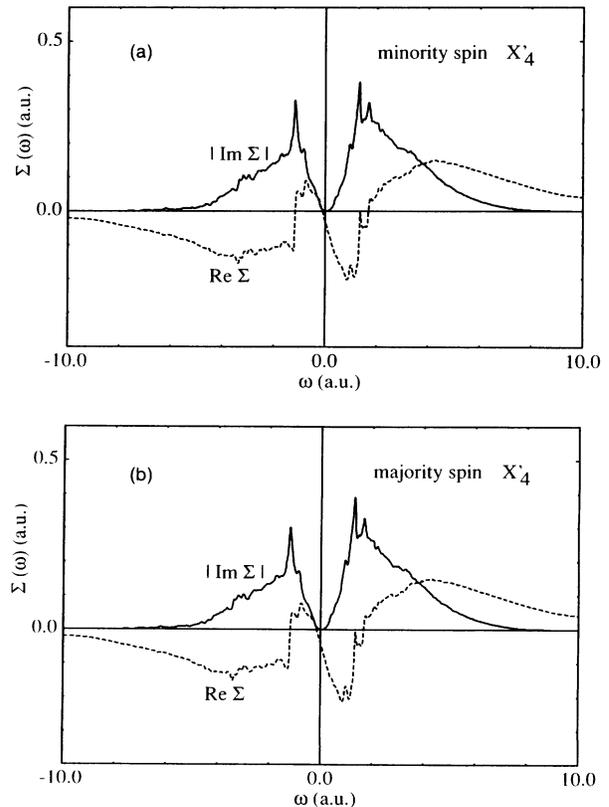
FIG. 2. (a) The real and imaginary parts of the correlated part of the self-energy for the minority spin state X'_4. (b) The real and imaginary parts of the correlated part of the self-energy for the majority spin state X'_4.
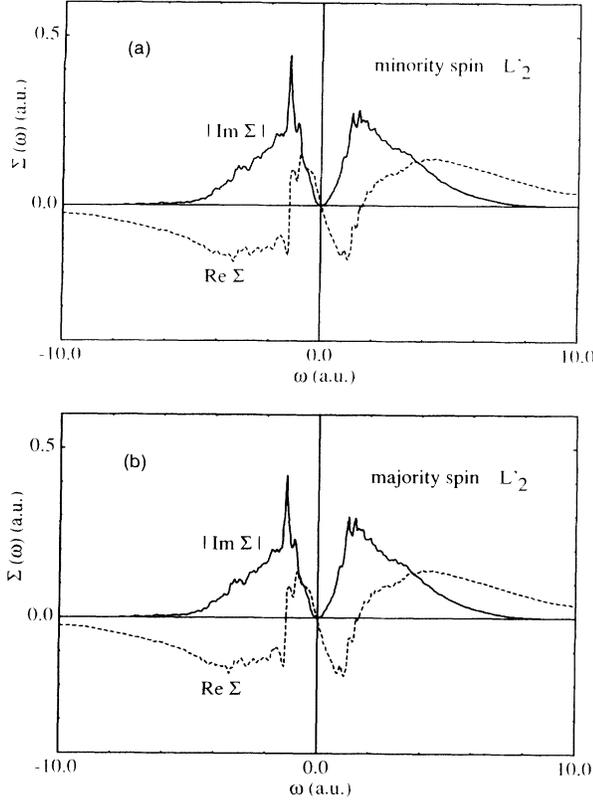
FIG. 3. (a) The real and imaginary parts of the correlated part of the self-energy for the minority spin state L_2'. (b) The real and imaginary parts of the correlated part of the self-energy for the majority spin state L_2'.

fore by von der Linden and Horsch[14] against the statement by some authors that the self-energy may be approximated by a scissor operator, i.e., a uniform shift across the Brillouin zone. Although the statement may be true in semiconductors except at a few states, it is cer-
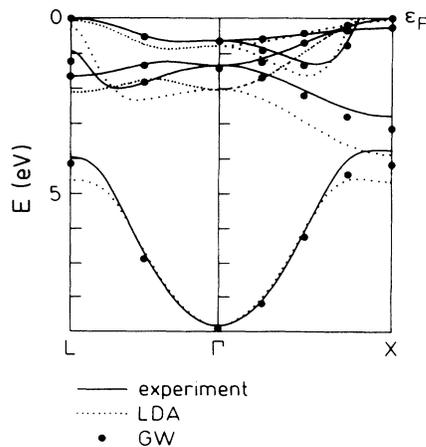
——— experiment
········· LDA
  • GW

FIG. 4. The band structure along the ΓX and ΓL directions. The solid curves are the experiment and the dotted curves are the LDA, both taken from Ref. 37. The filled circles are the quasiparticle energies in the GWA. The unit is in eV and the energy is measured with respect to the Fermi level.

tainly not true in the case of nickel.

In Table II we present the quasiparticle energies at high symmetry points for the occupied states and compare them with experimental values when available. The numerical error is estimated from convergence test to lie between 0.1 and 0.2 eV. Since the splitting is small, the experimental quasiparticle energies are the average of the up- and down-spin channels. Only states near the Fermi level are resolved, but states which lie further away from the Fermi level have larger quasiparticle widths and therefore the up- and down-spin channel are unresolved. Due to experimental difficulties in determining the Fermi level, absolute comparison of the quasiparticle energies is not very meaningful. We have found, however, that the data obtained by Mårtensson and Nilsson[37] fit well with our calculations without adjustment, whereas a downward shift of 0.3 eV of the data of Eberhardt and Plummer[18] appears necessary for comparison. Indeed, it has been pointed out that the latter data are too high by $0.2-0.4$ eV.[37]

With the fore-mentioned shift, agreement with Eberhardt and Plummer data is in most cases very good except for the following states.

(1) X_2 state: the experimental point appears too low. That this is the case is confirmed by comparison with two other sets of data.[17,37]

(2) X_1 state: the experimental point appears too high. Again comparison with two other sets of data confirms this. Here, the calculated quasiparticle energy is not in very good agreement with the other two sets of data but still it is unlikely that the self-energy correction should be as large as 1.5 eV.

(3) W_2' and W_3: the discrepancies are not large and no other experimental data are available. It is difficult to judge what the correct values should be.

In Table II we also give the full width at half maximum (FWHM) of the quasiparticle energies obtained from the imaginary parts of the self-energies and compare them with experiment. The agreement is very favorable. The widths vary from zero at the Fermi level to about 2 eV at the bottom of the band which corresponds to a lifetime of $\sim 7 \times 10^{-16}$ sec and a scattering length of $l = v_g \tau \sim 4$ Å if we take v_g to be the group velocity of the free-electron band.

Inclusion of transitions from the core states in the response function has little effect on the quasiparticle energies, apart from an almost uniform upward shift of ~ 0.1 eV. Since we are interested in relative eigenvalues, this shift is not of much importance and the core electrons may therefore be neglected in the computation of the response function. On the other hand, the core electrons should be included when we are interested in the high energy spectrum.

The exchange splittings are not improved from their LDA values to the accuracy of the computation. In some cases they are improved by 0.1 eV or at most 0.2 eV. In fact, when the LDA splitting is small, the GWA tends to increase it marginally. The decrease in the exchange splittings mainly occur at the top of the band. It has been suggested that the splittings may indeed increase with the distance from the Fermi level.[18] Since the exper-

imental exchange splittings are only about 0.3 eV, we might argue that they lie outside the computational accuracy which is between 0.1 and 0.2 eV. But judging from the uniformity of the results across the Brillouin zone and the good result for the quasiparticle energies, it seems that the discrepancy is not due to numerical inaccuracy but rather to the inadequacy of the GWA itself. The exchange splitting seems to be closely related to the existence of the 6-eV satellite which we turn to in the next paragraph.

The spectral function $A(\omega) = -(1/\pi)\mathrm{Im}\,\mathrm{tr}G(\omega)$, where the Green function is obtained from the Dyson equation, is plotted in Figs. 5 and 6. The main spectral width is about 3.5 eV with a shoulder structure extending to 4.5 eV. This shoulder structure comes mainly from the X_1 and L_1 states, both at the zone edges with high density of states. Between 5 and 6 eV we see a satellite-like structure. This structure, however, originates from quasiparticle peaks corresponding to the lowest valence states at $\mathbf{k}$ around (0.5, 0.25, 0.0) and (0.5, 0.25, 0.25). When contributions from these two states are taken out, the peak disappears as shown in the figure. A similar observation has been made before[38] and the conclusion is that the weight of the satellite coming from many-body effects may not be as large as it is assumed to be. We would like to emphasize that this does not imply that there is no satellite structure resulting from many-body correlations. On the contrary, the resonance at the $3p$ threshold clearly indicates the presence of many-body effects.

At lower energies (Fig. 6) we find two peaks at -24 and -34 eV which we believe are due to plasmonlike excitations. These two peaks may be traced back to the two large peaks in the dielectric function at about the same

TABLE II. Quasiparticle energies in eV at high symmetry points for majority and minority spin (alternately). $\epsilon^{\mathrm{LDA}}_{\mathbf{k}n}$ is the LDA eigenvalue calculated with our modified LAPW method (Ref. 28). $\langle E_{\mathbf{k}n} \rangle$ is the average quasiparticle energy. The experimental data are taken from Ref. 18 and those in the parentheses from Ref. 37. The experimental data from Ref. 18 have been shifted down by 0.3 eV for the purpose of comparison.

$\mathbf{k}n$	$\epsilon^{\mathrm{LDA}}_{\mathbf{k}n}$	$E_{\mathbf{k}n}$	$\langle E_{\mathbf{k}n} \rangle$	Expt.	FWHM GW	FWHM Expt.
Γ_1	-9.26	-9.0			2.1	
	-9.24	-9.0	-9.0	-9.1 ± 0.2	2.1	1.8
Γ'_{25}	-2.35	-1.7			0.8	
	-1.76	-1.1	-1.4	-1.4 ± 0.2	0.4	
Γ_{12}	-1.23	-0.9			0.3	
	-0.59	-0.3	-0.6	-0.7 ± 0.1	0.2	
X_1	-4.93	-4.3			2.9	
	-4.54	-3.9	-4.1	$-3.6\pm0.2(-3.8)$	2.5	1.25
X_3	-4.28	-3.5			2.2	
	-3.80	-2.9	-3.2	$-3.1\pm0.2(-2.8)$	1.5	1.4
X_2	-0.60	-0.5			0.1	
	$+0.07$	$+0.2$	-0.3	$-1.15\pm0.1(-0.2)$	0.1	
L_1	-4.95	-4.3			2.5	
	-4.60	-4.0	-4.1	-3.9 ± 0.2	2.0	
L_3	-2.44	-1.9			1.0	
	-1.86	-1.3	-1.6	-1.6 ± 0.1	0.6	0.9
L'_2	-0.74	-1.3			0.02	
	-0.73	-1.3	-1.3	$-1.3\pm(-1.0)0.1$	0.02	
W'_2	-4.00	-3.4			2.2	
	-3.58	-3.1	-3.3	-2.9 ± 0.2	1.9	
W_3	-3.09	-2.6			1.4	
	-2.61	-2.0	-2.3	-2.0 ± 0.2	0.8	1.3
W_1	-1.40	-1.1			0.8	
	-0.79	-0.5	-0.8	-0.95 ± 0.1	0.2	0.8
K_1	-4.18	-3.6			2.4	
	-3.76	-3.2	-3.4	-3.4 ± 0.2	2.0	1.3
K_1	-3.80	-3.2			1.9	
	-3.36	-2.6	-2.9	-2.85 ± 0.1	1.2	1.0
K_3	-2.13	-1.8			0.4	
	-1.65	-1.2	-1.5	-1.2 ± 0.2	0.8	0.8
K_3	-1.14	-1.0			0.3	
	-0.51	-0.4	-0.7	-0.75 ± 0.1	0.2	

F. ARYASETIAWAN

positions.[28] They are not as large as in simple metals because the d electrons in nickel are very localized and tightly bound as indicated by the small bandwidth and the large charges inside the muffin-tin sphere (Table I). Moreover, the plasmons in Ni merge with single-particle excitations so that their widths are broadened and their strengths are consequently reduced.

We have also calculated the spectral function from the following expression:

$$A(\omega) = -\frac{1}{\pi} \sum_{\mathbf{k}n} \frac{\mathrm{Im}\Sigma_{nn}(\mathbf{k},\omega)}{[\omega - \varepsilon_{\mathbf{k}n} - \mathrm{Re}\Sigma_{nn}(\mathbf{k},\omega)]^2 + [\mathrm{Im}\Sigma_{nn}(\mathbf{k},\omega)]^2}$$ (44)

and found little difference between this and the spectral function obtained from the full Green function, as may be seen in an example shown in Fig. 7. This is because the off-diagonal elements of the full Green function are small, which means that the quasiparticle wave functions do not differ much from the LDA wave functions. A similar result has been obtained by Hybertsen and Louie[13] in their GW calculations of semiconductors. The numerical implication is quite significant since, to a very good approximation, only the matrix elements of the self-energy taken between the valence states are required instead of the whole 50×50 self-energy matrix.

The quasiparticle energies may also be determined from the peaks in the imaginary parts of the Green function. The results agree to within the accuracy of the computation with the results obtained from the formula in (42).

We summarize the main results: the quasiparticle band structure as well as the quasiparticle widths are given very well by the GWA but the exchange splittings remain significantly large and the satellite structure at 6 eV is not reproduced in our calculation. We will try to give a unified explanation for the discrepancies between the GWA and experiment. In short, the satellite and the correction to the exchange splittings are not contained in the screening but rather in the so-called vertex correc-

tions, i.e., beyond the GWA.

We first mention some facts about the 6-eV satellite. The satellite does not appear to arise from any feature in the single-particle band structure and shows no dispersion in the angle-resolved photoemission spectra.[18,24] At 67 eV incident photon energy corresponding to the binding energy of the $3p$ level, it shows a Fano-type (asymmetric) resonant enhancement in intensity and at the same time the main $3d$ emissions show a strong antiresonance.[39] It has also been found that the satellite is spin-polarized near resonance.[40]

The commonly accepted explanation for the existence of the satellite is the following:[25,26] during a photoemission process, a d electron is emitted out of the solid and another at the same atomic site is excited to an empty d state just above the Fermi level. In the atomic picture, this corresponds to the configuration $3d^7 4s^2$ which is separated from the main band $(3d^8 4s)$ by more than 6 eV, but metallic screening should reduce this value, making it closer to the observed value. The created two holes multiple scatter by Coulomb interaction and form a virtual bound state at 6 eV. Or in a simple single-particle picture, the photon energy is used to take one d electron out and to excite another into an empty d state so that the emitted electron appears to have a lower binding energy corresponding to the satellite energy. Due to selection rules, electrons at the bottom of the d band, which hybridize with the s-p band, have the largest probability of being excited to the empty d states. As a consequence,
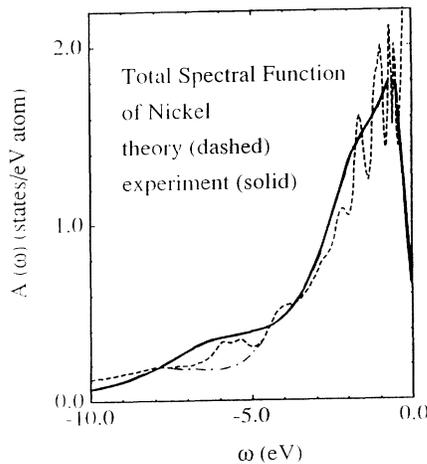
FIG. 5. The total spectral function of nickel. The solid curve is the experimental result taken from Ref. 17 and the dashed curves is the theoretical result within the GWA. The dashed-dotted curve corresponds to the case when two quasiparticle peaks at around 5–6 eV are taken out.
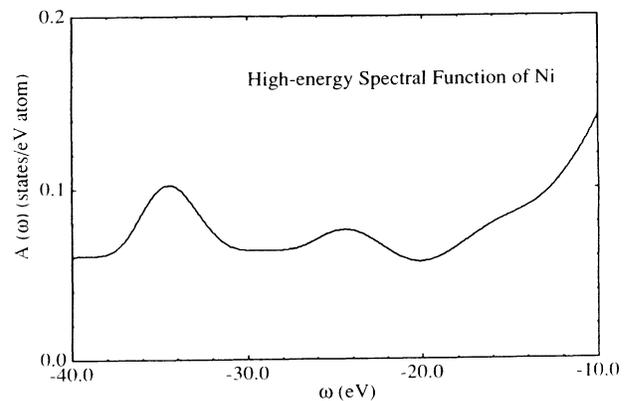
FIG. 6. The spectral function of nickel at high energy with two plasmon peaks at -24 and -34 eV and a shoulder structure at -16 eV.
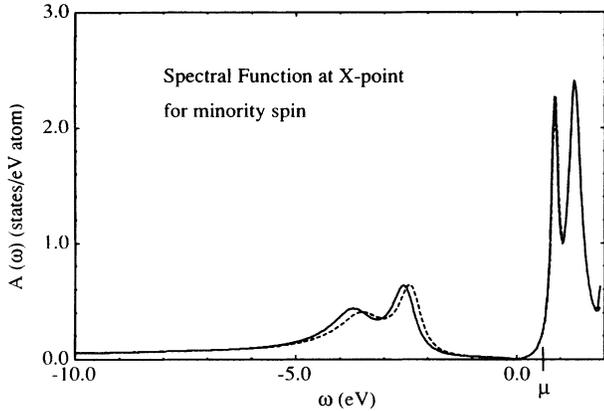
FIG. 7. The spectral function at the X point for the minority spin. The solid curve is obtained from the full Green function, calculated from the Dyson equation, and the dashed curve is obtained without including the off-diagonal elements of the self-energy matrix [see Eq. (44)].

the spectral weight that goes to the satellite mainly comes from the bottom of the band which then results in band narrowing.

The main source of band narrowing, however, comes from screening, whereas contribution from the hole-hole interactions is secondary. The band narrowing seems to show that the screening mechanism in nickel is quite different from that in iron, cobalt, or copper. Due to the localized character of the d states, a d hole is also localized. The hole is screened by the surrounding d electrons in the same atomic site as well as electrons from neighboring atoms, mainly s electrons but also d electrons to a lesser degree. The hole then travels with the screening cloud from site to site, forming a quasihole which turns out to have a heavier effective mass than the corresponding screened LDA state, resulting in band narrowing. In terms of perturbation theory, the screening is associated with the excitations of electrons from below to above the Fermi level. The largest contribution comes from excitations of $3d$ electrons to the empty states just above the Fermi level. This is evident when we look at the imaginary part of the noninteracting response function where the main peak lies below 5 eV.[28] These excitation processes are very much energy dependent and their effects on the quasihole energies are embodied in the energy-dependent, nonlocal self-energy operator through the density-density response function.

The resonance at the $3p$ threshold is attributed to an Auger process where a $3p$ electron is excited to fill the empty d states followed by a super-Coster-Kronig decay which in atomic configuration interaction picture is given by

$$3p^6 3d^9 4s + h\nu \rightarrow 3p^5 3d^{10} 4s \rightarrow 3p^6 3d^8 4s + \varepsilon l \ . \quad (45)$$
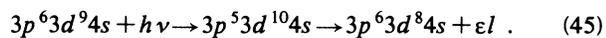
The $3p$ resonance indicates that the presence of unoccupied d states is indispensible in explaining the narrowing of the band. Indeed, in the case of Cu with a filled d band, LDA is known to give a reasonable bandwidth.

The reduction in the exchange splittings may also be

explained by the same model. For simplicity we consider a single-band model with a fully occupied majority channel and a partially occupied minority channel. A hole in the majority channel can excite an electron in the minority channel to the unoccupied states, resulting in two holes. On the other hand, a hole in the minority channel cannot excite an electron from the majority channel since there is no empty state to go into. This means that there is no contribution to the satellite from the minority channel. This implies in the real case that the majority channel has a larger band narrowing, resulting in a decreasing in the exchange splittings.

The above physical mechanism is modeled by Penn[25] and Liebsch[26] with a Hubbard Hamiltonian. The self-energy arising from the hole-hole interaction is calculated in the t-matrix formulation,[41] assuming that there is at most one electron in an unfilled d state because the number of such states is small. The self-energy diagrams for this model are drawn in Fig. 8. The t matrix is given by $t = U/(1 + UG_2)$, where U is the Hubbard parameter and G_2 is the two-hole Green function evaluated with the Hartree-Fock eigenvalues. The t matrix gives the effective hole-hole interaction. The denominator of t represents the multiple scattering shown in Fig. 8 and a pole in t corresponds to the satellite binding energy. In the Penn model, the d orbitals are assumed to be degenerate whereas in the Liebsch model the multiplets are taken into account. The latter has the advantage that it goes over to the atomic case in the limit of no overlap between neighboring atoms.

To investigate the discrepancies between the GWA and experiment, we look for physical processes which are left out in the GWA. We may do this by making a comparison between the Hubbard model and the GWA. Direct comparison between the two is not straightforward for several reasons. The GWA is a first-principle theory whereas the Hubbard model is, as the name implies, a model of a physical system with an adjustable Hubbard parameter U which contains the effect of screening to all orders in perturbation theory. U is a hard-core, contact interaction whereas W is a screened Coulomb potential which is softer and nonlocal as well as frequency dependent. But for the purpose of identifying physical processes responsible for the presence of the satellite and the reduction in the exchange splittings, it is reasonable to make a diagrammatic comparison between the GWA and the Hubbard model.
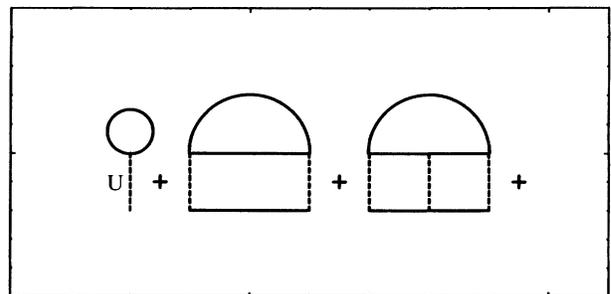
FIG. 8. The self-energy diagrams in the Hubbard models used by Penn and Liebsch.

From the point of view of the GWA, the correlation between the hole created during the photoemission process and the rest of the system is described by the diagrams in Fig. 9. The hole interacts with the rest of the system through virtual creation and annihilation of electron-hole pairs (polarization) which amounts to screening the Coulomb interaction. Polarization processes whereby another d electron is excited to an unoccupied state just above the Fermi level are included in the diagrams. However, ladder diagrams shown in Fig. 8 representing repeated hole-hole scatterings are absent in the GWA approximation. These processes which lead to the virtual two-hole bound state are important for the description of the satellite and the exchange splittings. Therefore the GWA is not expected to be able to describe the satellite and give the correct exchange splittings as our calculation has shown.

The discrepancy in the band structure for the bottom of the d band may also be explained by the absence of hole-hole interactions in the GWA. As discussed earlier, the weight of the satellite mainly comes from the bottom of the d band and this satellite arises from the hole-hole interactions. Since these interactions are missing in the GWA, we expect discrepancies in the quasiparticle energies at the bottom of the d band. In Fig. 4 we see the discrepancy for the lowest d band increases as we move from the Γ point to the X point.

The GWA, on the other hand, takes into account interatomic interactions which, as shown by Kanamori,[42] can increase the effective intra-atomic Coulomb interactions by as much as 30% in the Hartree-Fock approximation. The interatomic interactions in any case should be of much less importance than the intra-atomic ones due to the very localized nature of the d states.

We might still argue that the virtual excitations of electrons from the d band to unoccupied states just above the Fermi level would give a large contribution to the response function due to the small energy denominator and consequently a satellite peak in the spectral function. But inspection of the inverse of the dielectric function[28] shows no such strong peak at around 6 eV both experimentally and theoretically. In RPA the screening is affected by excitations of particle-hole pairs which through inversion of Eq. (15) may conspire to form collective excitations (plasmons). The full response function χ is in general very different from the noninteracting response function χ^0 since the latter contains no

plasmons. The 6-eV satellite does not correspond to plasmon excitations of the $3d$ electrons since such plasmon peaks are absent in the imaginary part of χ. It is clear that the existence of the satellite is not a direct consequence of screening effects but due rather to vertex corrections in the self-energy which in this case takes the form of the ladder diagrams shown in Fig. 8. The effects of the self-energy corrections on the exchange splittings are probably not contained in the screening either since W is spin independent.

Adding vertex corrections to the GWA is unfortunately no simple task, both from numerical and theoretical points of view. Theoretically, there is a difficult problem associated with conservation laws. The GWA with self-consistent Green function is conserving in particle number, energy, and momentum, but with the zeroth-order Green function it is only particle-number conserving as has been stated before. Energy and momentum conservations are important when we are interested in transport properties which involve rates of change of energy and momentum, but for ground-state properties such as total energy, it should be a good approximation to replace G with G^0. A general prescription for constructing a conserving approximation has been given by Baym.[33] Provided the self-energy is expressed as a functional derivative with respect to G of some scalar $\Phi[G]$, the approximation is conserving. There are two practical problems: the first is the choice of $\Phi[G]$ and the second is self-consistency. $\Phi[G]$ is a functional of the self-consistent Green function which in practice is very difficult if not impossible to obtain. Very little progress has been made in this field so far.

The reason for the success of the GWA in predicting the quasiparticle energies is not entirely clear. For the core-electron case, the following model Hamiltonian[43,44]

$$H = H_v + \varepsilon_c^0 b^\dagger b + V b b^\dagger \qquad (46)$$

yields the same core-electrons energy as in the GWA. H_v is the Hamiltonian of the valence electrons which include valence-valence and valence-core interactions, V is the potential experienced by the valence electrons when a hole is created, and $b, b^\dagger$ are the annihilation and creation operators for the core electrons. For the valence and conduction electrons, the situation is less clear. In the case of nickel, the band narrowing (from the corresponding LDA band) has a main contribution from dynamic screening and to a lesser extent from hole-hole interactions which may be thought of as vertex corrections (going beyond the GWA).

We finish this section with discussions on the numerical aspects of the calculation. We first consider the basis functions. A question may be raised whether the number of basis functions used in the calculation is sufficient to describe the screened potential and the self-energy. We have used basis functions which are eigenfunctions of the zeroth-order Hamiltonian (LDA) and which should be rather close to the quasiparticle wave functions. Although the basis functions are not complete, they form a suitable space for describing the physics at low energy. The zeroth-order Green function G^0, for instance, is ex-
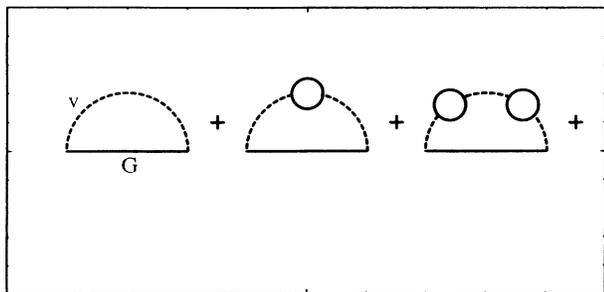
FIG. 9. The self-energy diagrams in the GWA.

actly represented in the basis as long as we do not go beyond the energy cutoff. In contrast, a large number of plane waves would be required to represent G^0. Although absolute convergence may not be obtained with a finite basis, we are interested in relative quasiparticle energies and their variation in k space so that systematic errors that may arise tend to cancel one another.

To substantiate our argument, we list the following numerical facts. (1) The f-sum rule

$$\int_0^\infty d\omega\,\omega\,\mathrm{Im}\epsilon(\mathbf{q};\omega) = -\int_0^\infty d\omega\,\omega\,\mathrm{Im}\epsilon^{-1}(\mathbf{q};\omega) = \frac{\pi}{2}\omega_p^2 \quad (47)$$

is satisfied to 90% for a number of q's.

(2) The Coulomb potential

$$\langle \mathbf{q}|v|\mathbf{q}\rangle = \sum_{nn'} \langle \mathbf{q}|\mathbf{q}n\rangle v_{nn'}(\mathbf{q})\langle \mathbf{q}n'|\mathbf{q}\rangle\ , \quad (48)$$

where $v_{nn'}(\mathbf{q}) = \langle \mathbf{q}n|v|\mathbf{q}n'\rangle$, differs from the exact value of $4\pi/q^2$ by less than 3%.

(3) The high-frequency regions of the imaginary parts of the self-energies shown in Fig. 3 are of similar magnitude for the hole and electron parts. When performing the Hilbert transform to obtain the real parts of the self-energies, the hole and electron parts tend to cancel one another so that the high-frequency regions are of little significance. Most of the contribution to the real parts come from frequency regions ± 5 Ry. The second numerical aspect concerns the integration over the BZ. Considering the complicated structure of the Fermi surface in nickel, one would expect that a large number of k points is required to get convergence. Surprisingly, the real parts of the self-energies are found to differ by only ~ 0.1 eV when the number of k points is increased from 20 to 89, suggesting that the integrands in k space are a slowly varying function of k. In Eqs. (31), (32), and (33), it is the matrix elements 31 that determine the structure of the integrands in k space. Provided we have a correct band labeling, these matrix elements should be a slowly varying functions of k since the wave functions should not change character dramatically except in some exceptional cases.

V. CONCLUSIONS

We have performed a first-principle calculation of the self-energy of nickel in the GWA starting from a self-consistent LDA Hamiltonian with LAPW basis. The quasiparticle energies are in very good agreement with experiment except at the bottom of the d band where a discrepancy of 0.3–0.4 eV remains. Consequently, the d band is narrowed by ~ 1 eV. The widths of the quasiparticles are also in favorable agreement with the observed values, in particular the unusually large widths at the bottom of the d band. The exchange splittings, however, remain essentially unchanged to the accuracy of the computation. In some cases, mostly at the top of the band, a reduction of 0.10–0.20 eV is found.

The core electrons have little effect on the quasiparticle energies. They merely give an almost uniform upward shift of ~ 0.1 eV. Only when we are interested in the high energy spectrum is it necessary to include core electrons in the response function.

A satellitelike structure is found between 5 and 6 eV but this comes entirely from quasiparticle peaks, not as a result of many-body shake-up. Experimental evidence strongly suggests many-body effects as the origin of the satellite; the resonance at the $3p$ threshold is particularly convincing. The theoretical result shows, however, that the single-particle contribution to the satellite weight is quite significant as has been suggested before.[38] It seems clear that the satellite peak and the reduction in exchange splittings are due to the two-hole bound state as model calculations have shown. The two-hole interactions are absent in the GWA and therefore it is not surprising that the GWA does not give the satellite and the reduction in exchange splittings. In the case of nickel we may identify two contributions to the band narrowing: the largest comes from screening or polarization processes and a further contribution arises from the two-hole interactions mainly for states at the bottom of the d band.

It is also found that the spectral function obtained from the full G differs very little from that obtained from G^0. This suggests that the quasiparticle wave functions do not deviate much from the LDA wave functions. For most purposes, it is then sufficient to compute the diagonal components of the self-energy corresponding to the states of interest instead of computing the whole self-energy matrix.

Since the quasiparticle energies are in good agreement with experiment, one would expect that the charge density and magnetic moment obtained from the Green function should improve those of the LDA or at least maintain the good results of the LDA in these two quantities. To calculate a quantity like a magnetic moment, however, requires a self-consistent calculation and a larger number of k points, which are beyond our computational capability.

The domain of applicability of the GWA appears to be larger than one might anticipate from a relatively simple many-body theory. The strength of the theory is that it contains the essential physical processes, in determining quasiparticle properties, which are common to a large class of materials. Qualitatively, it is the dynamical screening of a hole or particle with the formation of a localized polarization cloud which determines the effective mass of the quasihole (particle) as it moves from site to site. Quantitatively, we do not know the exact reasons for the success of the GWA in predicting quasiparticle energies. Only in the case of core electrons do we have a clear explanation but the situation is still obscure for the valence and conduction electrons. It is our hope that the work we have presented in this paper serves as a further guide in understanding the fundamental mechanism behind the theory.

The good results for nickel encourage further applications to even more strongly correlated systems such as NiO. It would be useful, however, to do a preliminary study in the form of model calculations in order to identify what physical processes are responsible for the widening of the gap. Controversy around the validity of perturbation theory like the GWA to systems like NiO is difficult to settle until numerical calculations have been performed.

*Present address: Max-Planck-Institut für Festkörperforschung, Heisenbergstrasse 1, D-7000 Stuttgart 80, Federal Republic of Germany.

[1] L. Hedin and S. Lundqvist, in *Solid State Physics,* edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1969), Vol. 23, p. 1.

[2] L. J. Sham and W. Kohn, Phys. Rev. **145**, 561 (1966); see also J. C. Slater, *Self-Consistent Field for Molecules and Solids* (McGraw-Hill, New York, 1974), Vol. IV.

[3] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[4] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[5] D. R. Hamann, Phys. Rev. Lett. **42**, 662 (1979).

[6] G. B. Bachelet and N. E. Christensen, Phys. Rev. B **31**, 879 (1985).

[7] C.-O. Almbladh and A. C. Pedroza, Phys. Rev. A **29**, 2322 (1984).

[8] F. Aryasetiawan and M. J. Stott, Phys. Rev. B **38**, 2974 (1988).

[9] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10 159 (1988).

[10] C.-O. Almbladh and U. von Barth, in *Density Functional Methods in Physics,* Vol. 123 of *NATO Advanced Study Institute, Series B: Physics,* edited by R. M. Dreizler and J. da Providencia (Plenum, New York, 1985).

[11] L. Hedin, Phys. Rev. **139**, A796 (1965).

[12] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[13] M. S. Hybersten and S. G. Louie, Comments Condens. Matter Phys. **13**, 223 (1987).

[14] W. von der Linden and P. Horsch, Phys. Rev. B **37**, 8351 (1988).

[15] C. S. Wang and J. Callaway, Phys. Rev. B **15**, 298 (1977).

[16] S. Hüfner, G. K. Wertheim, N. V. Smith, and M. M. Traum, Solid State Commun. **11**, 323 (1972).

[17] F. J. Himpsel, J. A. Knapp, and D. E. Eastman, Phys. Rev. B **19**, 2919 (1979).

[18] W. Eberhardt and E. W. Plummer, Phys. Rev. B **21**, 3245 (1980).

[19] P. C. Kemeny and N. J. Shevchik, Solid State Commun. **17**, 255 (1975).

[20] V. L. Moruzzi, J. F. Janak, and A. R. Williams, *Calculated Electronic Properties of Metals* (Pergamon, New York, 1978).

[21] C. E. Moore, *Atomic Energy Levels,* Natl. Bur. Stand. Ref. Data Ser., Natl. Bur. Stand (U.S.) Circ. No. 467 (U.S. GPO, Washington, D.C., 1949, 1952, 1958).

[22] D. C. Tsui, Phys. Rev. **164**, 669 (1967).

[23] D. C. Tsui and R. W. Stark, J. Appl. Phys. **39**, 1056 (1968).

[24] S. Hüfner and G. K. Wertheim, Phys. Rev. B **8**, 4857 (1973).

[25] D. R. Penn, Phys. Rev. Lett. **42**, 921 (1979).

[26] A. Liebsch, Phys. Rev. B **23**, 5203 (1981).

[27] G. Treglia and F. Ducastelle, Phys. Rev. B **21**, 3729 (1980).

[28] F. Aryasetiawan, U. von Barth, P. Blaha, and K. Schwarz (unpublished).

[29] See, e.g., A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).

[30] P. M. Moore and M. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).

[31] D. Pines, *Elementary Excitations in Solids* (Benjamin, New York, 1963).

[32] G. Baym and L. P. Kadanoff, Phys. Rev. **124**, 287 (1961).

[33] G. Baym, Phys. Rev. **127**, 1391 (1962).

[34] O. K. Andersen, Phys. Rev. B **12**, 3060 (1975).

[35] D. D. Koelling and G. O. Arbman, J. Phys. F **5**, 2041 (1983).

[36] P. Blaha and K. Schwarz, J. Phys. F **17**, 899 (1987).

[37] H. Mårtensson and P. O. Nilsson, Phys. Rev. B **30**, 3047 (1984).

[38] J. Kanski, P. O. Nilsson, and C. G. Larsson, Solid State Commun. **35**, 397 (1980).

[39] C. Guillot, Y. Ballu, J. Paigné, J. Lecante, K. P. Jain, P. Thiry, R. Pinchaux, Y. Pétroff, and L. M. Falicov, Phys. Rev. Lett. **39**, 1632 (1977).

[40] R. Clauberg, W. Gudat, E. Kisher, E. Kuhlmann, and G. M. Rothberg, Phys. Rev. Lett. **47**, 1314 (1981).

[41] J. Kanamori, Prog. Theor. Phys. **30**, 275 (1963).

[42] J. Kanamori, in *Electron Correlation and Magnetism in Narrow-Band Systems,* edited by T. Moriya, Springer Series in Solid-State Sciences Vol. 29 (Springer-Verlag, Berlin, 1981).

[43] D. C. Langreth, Phys. Rev. B **1**, 471 (1970).

[44] C.-O. Almbladh and L. Hedin, *Beyond the One-Electron Model,* Handbook on Synchroton Radiation, Vol. 1, edited by E. E. Koch (North-Holland, Amsterdam, 1983).