

Excitation spectra of semiconductors and insulators: a density-functional approach to many-body theory

E G Maksimov, I I Mazin, S Yu Savrasov and Yu A Uspenski

Academy of Sciences of the USSR, P N Lebedev Physical Institute, 117924 Moscow, Leninsky Prospect 53, USSR

Received 20 January 1988, in final form 7 June 1988

Abstract. We analyse the self-energy operator of the Dyson equation in solids in comparison with the Kohn–Sham local potential of the density-functional theory (DFT). This allows us to explain why there is little difference between the DFT one-particle spectra and the spectra of actual one-electron excitations (OEE) in metals and why this difference is considerable in insulators. We can identify the terms responsible for this difference. This allows us to propose a very simple method for deriving the OEE spectra from the DFT calculation. In particular, the calculations of the energy gaps in semiconductors (C, Si), large-gap dielectrics (Ne, Ar, Kr) and ionic solids (MgO, LiF, NaCl) provide an agreement with experiment well within 5%, as well as the pressure dependence of the gap in Xe up to metallisation. The calculational effort involved is little different from that required in the usual self-consistent DFT calculation.

1. Introduction

Recently great progress has been achieved in the computational electronic theory of solids. The theoretical basis of most calculations is provided by the density-functional theory (DFT) (Hohenberg and Kohn 1964, Kohn and Sham 1965). This formalism is exact for the ground-state properties (total energy, pressure, cohesive energy, etc.). However, the one-particle spectrum of the DFT non-interacting reference system may differ from the spectrum of the actual one-electron excitations (OEE) in a solid, defined by the poles of the one-particle Green function. Nevertheless, the DFT spectrum in metals turns out to be very close to the OEE one†.

The situation is different in insulating crystals. A well known problem is due to the fact that DFT heavily underestimates the dielectric gap, sometimes by more than 50%, and this underestimation remains even in very accurate DFT calculations (e.g. Godby *et al* 1987). At the same time, many-body perturbation theory gives quite reasonable results for the gaps (Pickett and Wang 1984, Strinati *et al* 1982, Hybertsen and Louie 1986). It has been noted by Perdew and Levy (1983) and by Sham and Schlüter (1983) that the effective potential of the DFT is defined only with an additive constant which may depend on the total number of electrons in the system. The variation of this constant when the electron number changes from N to $N + 1$ gives an additional contribution

† This is confirmed both by many-particle calculations (e.g. Hedin 1965, MacDonald 1980, Schreiber and Bross 1983, and others) and by comparison with experiments. The discrepancy is usually less than 10%.

(the so-called ‘energy derivative discontinuity’) to the calculated value of the gap. However, the calculation of this contribution is as complicated as the full many-body calculation of the OEE spectra.

The main practical question in this connection is the following one: Is it possible to retain the simplicity of the DFT while extending it to describe accurately enough the dielectric gap and the whole spectrum of OEE in insulating crystals? In order to answer this question we have performed a comparative analysis of the DFT equations and the corresponding equations for the Green function. This allowed us to extract the contributions which are responsible for the qualitative difference between metals and insulators. It turns out that the corrections to the DFT quasi-particle spectrum may be calculated by use of a simple formula. The accuracy is about 5% in a wide range of ionic and covalent dielectrics and semiconductors. Our microscopic analysis also helps to understand better the causes of successes and failures of some other approaches.

2. Description of one-electron spectra in density-functional theory and Green function method

The one-particle equations of the density-functional theory (Kohn and Sham 1965) are

$$[-(\hbar^2\nabla^2/2m) + V_{\text{eff}}(r)]\tilde{\psi}_i(r) = \tilde{E}_i\tilde{\psi}_i(r) \quad (1)$$

where

$$V_{\text{eff}}(r) = V_{\text{ext}}(r) + V_{\text{H}}(r) + V_{\text{xc}}(r)$$

V_{H} is the Hartree potential, $V_{\text{xc}}(r) = \delta E_{\text{xc}}\{n(r)\}/\delta n(r)$, and $E_{\text{xc}}\{n(r)\}$ is the usual exchange–correlation energy. The OEE spectrum, defined by the poles of the one-particle Green function, is given by the Dyson-type equation

$$[-(\hbar^2\nabla^2/2m) + V_{\text{ext}}(r) + V_{\text{H}}(r)]\psi_i(r) + \int dr' \Sigma_{\text{xc}}(r, r', E_i)\psi_i(r') = E_i\psi_i(r). \quad (2)$$

In contrast with $V_{\text{xc}}(r)$, the self-energy operator $\Sigma_{\text{xc}}(r, r', E)$ is non-local, i.e. depends on both r and r' , and is also energy-dependent. Such a difference leads to some important consequences. For instance, in the homogeneous electron gas $V_{\text{eff}}(r) = \text{const}$, because of the translational invariance. Therefore the DFT spectrum coincides with the spectrum of the non-interacting electrons. However, the OEE spectrum is always different from the non-interacting one. In particular, in the Hartree–Fock approximation, where the exchange is non-screened and consequently long-range, the non-locality of Σ_{xc} leads to the vanishing density of states at $E = \mu$ (μ is the chemical potential). The screening of the exchange interaction makes Σ_{xc} more local (more short-range) and the renormalisation of the spectrum in the vicinity of μ is strongly suppressed, being only a few per cent (Hedin and Lundqvist 1969). The origin of the above-mentioned similarity between the DFT and the OEE spectra in metals also lies in the full screening, as we shall see.

We shall use for our analysis the lowest-order approximation to equation (2), i.e. the so-called GW approximation. In this approximation $\Sigma_{\text{xc}}(r, r', E)$ is given by the following equation:

$$\Sigma_{\text{xc}}(r, r', E) = -(i/2\pi) \int dE' G(r, r', E - E')W(r, r', E'). \quad (3)$$

Here $G(r, r', E)$ is the one-particle Green function and W is the dynamically screened Coulomb interaction. $G(r, r', E)$ is expressed in terms of the eigenfunctions ψ_i and eigenvalues E_i of the Dyson equation (2). However, as a first approximation, in the calculation of Σ_{xc} one can use ψ_i and \tilde{E}_i instead of ψ_i and E_i .

Numerous investigations of the homogeneous electron gas (Hedin and Lundqvist 1969), metals (MacDonald 1980, Schreiber and Bross 1983) and semiconductors (Strinati *et al* 1982, Hybertsen and Louie 1986) have demonstrated that the GW approximation is accurate enough for the calculation of the OEE spectrum. Further we shall concentrate on the non-locality and energy dependence of Σ_{xc} in the framework of this approach.

Let us use the spectral representation for the dielectric function $\epsilon^{-1}(r, r', E)$

$$\epsilon^{-1}(r, r', E) = \delta(r - r') - \frac{2}{\pi} \int_0^\infty \frac{E' \text{Im} \epsilon^{-1}(r, r', E')}{E^2 - E'^2 + i\delta} dE'. \tag{4}$$

Then (3) may be written as

$$\begin{aligned} \Sigma_{xc}(r, r', \tilde{E}_i) &= \Sigma_{xc}^{\text{sex}}(r, r', \tilde{E}_i) + \Sigma_{xc}^{\text{coh}}(r, r', \tilde{E}_i) \\ &= \sum_j \theta(\mu - \tilde{E}_j) \tilde{\psi}_j(r) \tilde{\psi}_j^*(r') W(r, r', \tilde{E}_i - \tilde{E}_j) \\ &\quad + \frac{1}{\pi} \sum_j \int d r'' \frac{e^2}{|r - r''|} \int_0^\infty \frac{dE' \text{Im} \epsilon^{-1}(r, r', E')}{\tilde{E}_i - \tilde{E}_j + E'} \tilde{\psi}_j(r) \tilde{\psi}_j^*(r'). \end{aligned} \tag{5}$$

The first term in (5) is called the screened exchange and is due to the pole of the one-particle Green function. The second one, known as the Coulomb hole, is due to the plasmon pole in the dielectric function. When can equation (5) be reasonably approximated by a local potential like $V_{xc}(r)\delta(r - r')$? Apparently, $\Sigma_{xc}(r, r', E)$ should decrease sufficiently fast with $|r - r'|$ and should not depend considerably on the energy. Let us see if these conditions are satisfied in real systems.

$\text{Im} \epsilon^{-1}(r, r', E')$ has a sharp maximum at an energy $E' \sim \hbar\omega_p$ where ω_p is the plasma frequency (in the well known plasmon pole approximation this function has a pole at $E' = \hbar\omega_p$). So the main contribution to the Coulomb hole term in (5) comes from the region $E' \sim \hbar\omega_p \gg \tilde{E}_i$. In this region the sum $\tilde{E}_i + E'$ only depends slightly on \tilde{E}_i . If we neglect $\tilde{E}_i - \tilde{E}_j$ in comparison with E' in the denominator, then, because of the completeness of the eigenfunction set ψ_i , we shall have

$$\Sigma_{xc}^{\text{coh}}(r, r', E_i) \sim \sum_j \tilde{\psi}_j(r) \tilde{\psi}_j^*(r') = \delta(r - r').$$

Such a reasoning has been used by Hedin and Lundqvist (1969) in connection with the homogeneous electron gas, but we see that it is true also for real metals as well as for insulators. Surely this part of Σ_{xc} is well approximated by the corresponding part of V_{xc} and here there is no substantial difference between DFT and Green function method.

The situation with the dynamically screened exchange is different. Here there is a great difference between metals and insulators owing to the different kinds of screening in these systems. As we shall see, it is the full screening that makes Σ_{xc}^{sex} in metals quasi-local, as well as in the homogeneous gas. In insulators Σ_{xc}^{sex} has a strongly non-local part which cannot be described by the local DFT potential and should be included separately in the calculations of OEE spectra.

The contribution of the dynamically screened exchange to the OEE energy $E_{k\lambda}$ in a band λ with a wavevector k may be written as follows:

$$\langle \tilde{k}\lambda | \Sigma_{xc}^{\text{sex}}(r, r', E_{k\lambda}) | \tilde{k}\lambda \rangle = - \frac{2}{(2\pi)^3} \sum_{\lambda'} \int d q \theta(\mu - E_{k+q,\lambda'})$$

$$\begin{aligned} & \times \sum_{GG'} \bar{\rho}_{k+q, \lambda', k, \lambda}^*(q+G) \bar{\rho}_{k+q, \lambda', k, \lambda}(q+G') \\ & \times W(q+G, q+G', E_{k\lambda} - E_{k+q, \lambda'}) \end{aligned} \quad (6)$$

where G and G' are reciprocal-lattice vectors, and $\bar{\rho}_{k+q, \lambda', k, \lambda}(q+G)$ is a form factor defined as

$$\bar{\rho}_{k+q, \lambda', k, \lambda}(q+G) = \int_{\Omega} dr \tilde{\psi}_{k+q, \lambda'}^*(r) e^{i(q+G)r} \tilde{\psi}_{k\lambda}(r). \quad (7)$$

The non-locality of $\Sigma_{xc}(r, r', \vec{E}_{k\lambda})$ is due to the fact that it decreases too slowly at long distances $|r - r'|$. The long-range asymptotics of Σ_{xc} are defined by the terms with $G = G' = 0$ and small q in (6). Specifically, the long-range contribution in Σ_{xc}^{sex} (if any) arises when the effective interaction $W(q+0, q+0, 0)$ diverges at $q \rightarrow 0$ like the bare Coulomb interaction $4\pi e^2/q^2$.

It follows from the orthonormality of $\tilde{\psi}_{k\lambda}$ that at $q \rightarrow 0$

$$\bar{\rho}_{k+q, \lambda', k, \lambda}(q+0) = \delta_{\lambda\lambda'} + (1 - \delta_{\lambda\lambda'}) iq P_{\lambda\lambda'}^k + O(q^2) \quad (8)$$

where $P_{\lambda\lambda'}^k$ is the matrix element of the dipole transition between the states λ and λ' :

$$P_{\lambda\lambda'}^k = \int_{\Omega} \tilde{\psi}_{k\lambda}^*(r) r \tilde{\psi}_{k\lambda'}(r) dr. \quad (9)$$

It is easily seen from (6)–(8) that in components with $\lambda \neq \lambda'$ of (6) the Coulomb divergence $1/q^2$ cancels out due to the factor q in (8). The resulting interaction is short-range in comparison with the bare Coulomb interaction. In the tight-binding limit equation (6) may be considered at $|r - r'| \rightarrow \infty$ as a dipole–dipole interaction contribution to Σ_{xc}^{sex} , taking into account (8) and (9)†. The components with $\lambda \neq \lambda'$ are qualitatively the same both in metals and in insulators. In both cases at $q \rightarrow 0$ the effective interaction corresponds to the energy $|E_{k\lambda'} - E_{k\lambda}|$ which is the energy of the inter-band transitions. At such energies the full screening does not exist, so that at large distance $|r - r'|$

$$W(r, r', E_{k\lambda} - E_{k\lambda'}) \sim V_C(r - r') \varepsilon^{-1}(\omega_{\lambda\lambda'})$$

where $\omega_{\lambda\lambda'} = |E_{k\lambda} - E_{k\lambda'}|/\hbar$. Owing to the fact that the interacting objects here are ‘dipoles’, the effective interaction decreases with distance comparatively fast, as $1/|r - r'|^3$. In many crystals the dipole–dipole interaction vanishes due to symmetry so that only ‘quadrupoles’ contribute to the effective interaction, which decreases in this case even faster.

Thus we see that the non-locality in terms $\lambda \neq \lambda'$ is weak. Let us turn to the diagonal terms $\lambda = \lambda'$ which exist when the band λ is fully or partly occupied. Form factors (7) in this case do not depend on q at $q \rightarrow 0$ and hence these terms correspond to the interactions of unit charges $\int dr |\psi_{k\lambda}(r)|^2 = 1$. The energy argument of the effective interaction $W(r, r', E_{k\lambda} - E_{k+q, \lambda})$ at $q \rightarrow 0$ tends to zero as $qV_{k\lambda}/\hbar$, where $V_{k\lambda}$ is the electron velocity. In insulators $\lim_{q \rightarrow 0} \varepsilon^{-1}(q+0, q+0, qV/\hbar) = \varepsilon_0$, ε_0 being the usual static macroscopic dielectric constant. Thus at large distance

$$W(r, r', qV/\hbar) \xrightarrow{|r-r'| \rightarrow \infty} V_C(r - r') \varepsilon_0^{-1} = e^2 \varepsilon_0^{-1} / |r - r'|.$$

The corresponding contribution in Σ_{xc}^{sex} is therefore long-range. It should be noted that

† Sterne and Inkson (1984) have discussed the problem of the dielectric gap in the tight-binding limit and have come to a similar conclusion.

Table 1. Energy gap (eV) in the Penn model. E_g is the initial gap, DFT is the local DFT result, HF is the Hartree–Fock approximation, equation (4) is the ‘exact’ value and equation (11) is our formula. ϵ_0 is the static macroscopic dielectric constant.

E_g	ϵ_0	DFT	HF	Equation (11)	Equation (4)
1.36	47.3	1.9	6.8	2.01	2.45
2.72	15.6	3.4	10.2	3.83	4.35
4.08	8.4	4.9	12.65	5.82	6.12
5.44	5.8	6.8	14.7	8.16	7.62
6.8	4.06	8.43	16.32	10.06	9.52
8.16	3.2	9.93	19.04	12.78	11.15

this contribution may be written in terms of the Hartree–Fock exchange self-energy Σ_{xc}^{HF} as

$$\Sigma_{xc}^{\text{non-local}} = \Sigma_{xc}^{\text{HF}} / \epsilon_0. \quad (10)$$

In other words, the non-local contribution in $\Sigma_{xc}(r, r', E_{k\lambda})$ stems from the dynamically screened exchange part but can be represented by the statically screened bare (Fock) exchange interaction. In fact only the diagonal part of Σ_{xc}^{sex} is substantially non-local, but the non-diagonal part is quasi-local both in Σ_{xc}^{sex} and in Σ_{xc}^{HF} , so we may hope to correct the DFT spectra relying on equation (10).

Indeed, it is well known that the Hartree–Fock (HF) approximation overestimates the dielectric gap, sometimes by more than 100%. Supposing that the non-local part of Σ_{xc} may be described by use of HF approximation with (10), and that all other parts are reasonably described by the DFT potential V_{xc} , we get the following simple formula:

$$\Delta E_{k\lambda} = \langle \tilde{k}\lambda | \Sigma_{xc}(r, r', E_{k\lambda}) - V_{xc}(r) \delta(r - r') | \tilde{k}\lambda \rangle = (E_{k\lambda}^{\text{HF}} - E_{k\lambda}^{\text{HF-LDA}}) / \epsilon_0. \quad (11)$$

This formula must give reasonable results both in the strong screening limit and in the opposite one. Indeed, in materials $\epsilon_0^{-1} \rightarrow 0$ so that the OEE spectrum calculated from (11) coincides with the DFT spectrum. As has been mentioned above, it is a good approximation in metals. In the tight-binding limit ($\epsilon_0^{-1} \rightarrow 1$) (11) gives the HF spectrum which is a good approximation for atomic systems. Moreover in the intermediate regime, e.g. in semiconductors, the gap overestimation of HF approximation is of the same order as the DFT underestimation. In this case reasonable results are very likely too. The simplest semiconductor model (Penn 1962) has been investigated recently by Guinea and Tejedor (1980). We use their results to convince ourselves that our formula (11) indeed gives good results (table 1). Then we check it for real solids.

3. One-electron spectra in real solids

Using the data of different authors we have calculated using equation (11) gaps and valence band widths of various semiconductors and dielectrics (table 2)†. For gaps equation (11) gives excellent results, in spite of its simplicity. The maximal error is 10%, the average one is less than 5%,‡ i.e. an order better than DFT or HF method separately.

† We do not distinguish between the conventional and exchange-only LDA band structures because the difference is usually not very large; however, in case of small-gap semiconductors (Ge) one should use the exchange-only (Gaspar–Kohn–Sham) potential in equation (11).

‡ The only exception is Ge where equation (11) underestimates the direct gap by 25%. It is known, however, that the band structure of Ge is especially sensitive to calculational details. For instance, in the many-body calculations of Hybertsen and Louie (1986), the error for Ge was greater than for other materials. Further investigation should show if our approach is applicable to other small-gap semiconductors with the same accuracy as to Si, C, etc.

Table 2. The direct gap E_g (eV) and the valence band width W (eV) in different solids.

	ϵ_0	E_g			
		DFT	HF	Equation (11)	Expt
Ge	15.8	0.5 ^a	4.3 ^b	0	1.0 ^b
C	5.84	5.6 ^c	15.0 ^c	7.2	7.3 ^c
Si	11.90	2.6 ^c	9.4 ^c	3.2	3.4 ^c
Ne	1.23	21.2 ^c	25.1 ^c	22.5	21.4 ^c
Ar	1.62	8.3 ^c	18.5 ^c	14.6	14.3 ^c
Kr	1.82	6.8 ^c	16.4 ^c	12.1	11.6 ^c
LiF	1.94	10.0 ^d	22.4 ^e	16.3	14.9 ^d
NaCl	2.38	5.05 ^d	12.1 ^f	8.0	8.97 ^d
MgO	3.01	4.2 ^d	18.0 ^g	8.8	7.6 ^d

	ϵ_0	W			
		DFT	HF	Equation (11)	Expt
Ne	1.23	0.61 ^h	0.5 ⁱ	0.52	1.31 ^h
Ar	1.62	0.85 ^h	1.21 ⁱ	1.07	1.66 ^h
C	5.84	20.4 ^j	29.0 ^j	21.9	24.2 ^j
Si	11.90	11.9 ^k	16.9 ⁱ	12.3	12.5 ^k

^a Our calculations.^b Svane (1987).^c Horsch (1985).^d Carlsson (1985).^e Perrot (1972).^f Lipari and Kunz (1971).^g Pantelides *et al* (1974).^h Trickey *et al* (1981).ⁱ Dagens and Perrot (1972).^j Strinati *et al* (1982).^k Hybertsen and Louie (1986).^l Ohkosi (1985).

Let us note that our test included semiconductors (Si) as well as wide-gap insulators (Ne, Ar, Kr) and ionic crystals (LiF, NaCl). The valence band width of classical semiconductors—Si and C—also agrees well with the experiment. The narrow valence bands in Ne and Ar cannot be described by this method, probably because the simple GW approximation, which was the starting point of our analysis, is not valid here. An indication of this is the fact that in rare-gas crystals both the HF method and DFT underestimate this quantity.

It is interesting to test our approach on the band structure in several points of the Brillouin zone. The necessary HF data for diamond exist (Strinati *et al* 1982). Unfortunately, the experiments are not sufficiently accurate, so we compare our results with the many-particle calculations of Strinati *et al* (1982) (table 3).

Such a high accuracy of the very simple formula (11) may seem a little strange. We think that it is due to the fact that in (11) we have saved the high accuracy of the DFT in describing the short-range effects and have also taken into account the main factor responsible for the renormalisation of the gap, namely the long-range screened exchange interaction.

Another interesting consequence of our analysis is due to the so-called ‘self-interaction correction’ (SIC) method. This method was proposed by Perdew and Zunger (1981), who noted that the usual LDA approximation in DFT is wrong for a system of one electron, namely it includes the non-physical self-interaction to describe also any one-electron system, so Perdew and Zunger have proposed the scheme that had been devised as a ‘correction’ to the LDA (but not to the DFT itself). In this scheme, however, the effective potential is orbital-dependent, in contradiction with the DFT:

$$V_i^{\text{SIC}}(r) = V^{\text{DFT}}(r)\{n(r)\}V_{\text{H}}(r)\{n_i(r)\} - V_{\text{xc}}^{\text{DFT}}(r)\{n_i(r)\} \quad (12)$$

Table 3. Inter-band transition energies (eV) in diamond. The time-dependent Hartree-Fock approximation (TDHF) is taken as 'exact' for those transitions where the experimental values are not available.

Transition	DFT	HF	Equation (11)	TDHF	Expt
$\Gamma^{v1}-\Gamma^{v25'}$	21.7	29.0	22.9	25.2	24.0
$\Gamma^{c15}-\Gamma^{v25'}$	5.6	15.0	7.2	7.4	7.3
$\Gamma^{c2'}-\Gamma^{c15}$	7.9	11.0	8.44	10.2	8.0
$X^{v1}-X^{v4}$	6.5	8.0	6.73	6.85	
$X^{c4}-X^{v4}$	23.3	40.0	26.3	29.25	
$L^{v2'}-L^{v1}$	2.06	6.1	2.75	4.7	
$L^{v1}-L^{c1}$	22.6	38.0	25.3	37.6	
$L^{c1}-L^{c3}$	0.43	2.0	0.7	1.6	
$L^{c1}-L^{v2'}$	24.7	44.1	28.0	32.3	

where $n_i(r)$ is the density of the i th orbital. This approach has given some improvements in the total energy of atoms and considerable improvements of the OEE spectra. Later it has been generalised for solids (Heaton *et al* 1983), however somewhat artificially, with similar results. Later new approximations in the DFT have been devised that are self-interaction-free in the Perdew-Zunger sense[†]. Recent calculations (e.g. Manghi *et al* 1985, Godby *et al* 1987) gave nearly the same results for gaps as the LDA. We shall not go into details but it is almost clear now that in most cases even the exact DFT cannot give the results of the same accuracy for the excitation spectra as SIC does. The answer is that SIC is not an improvement of the LDA in the framework of the DFT, but a very indirect approximation for the self-energy in the Dyson equations[‡]. Our analysis confirms this: indeed, the non-local part of Σ_{xc} is due to the intra-band interaction between occupied states with nearly equal wavevectors. These are just the same terms that are affected by the SIC method. The difference is, however, in how these terms are handled in both cases.

The last problem due to the use of equation (11) lies in the HF eigenenergies that enter this equation. The full HF calculation for a solid is an extremely tedious procedure. Fortunately, there is a very simple procedure, based on the quasi-classical WKB method, that yields accurate enough HF spectra.

4. Quasi-classical approximation in the Hartree-Fock method

Sham and Kohn (1966) have proposed a quasi-classical ('local') approximation for the self-energy Σ_{xc} . In this approximation the non-local operator Σ_{xc} is approximated by a local energy-dependent operator:

$$\Sigma_{xc}(r, r', E) \rightarrow \Sigma_{xc}^{\text{loc}}(p(r, E), E)\delta(r - r'). \quad (13)$$

In order to find the local momentum $p(r, E)$ one should write the wavefunction in the quasi-classical form $e^{ip(r, E)r}$ where p is slowly varying with r . Then

$$\frac{\hbar^2 p^2(r)}{2m} - \frac{\hbar^2 k_F^2(r)}{2m} = E_{k\lambda} - E_F - [\Sigma_{xc}^{\text{loc}}(p(r, E_{k\lambda}), E_{k\lambda}) - V_{xc}(r)] \quad (14)$$

[†] Formally in the DFT there is no such notion as self-interaction, so for many-electron systems it is not possible, strictly speaking, to discuss self-interactions. It has been mentioned, for instance, by Pickett (1985).

[‡] The question if SIC is a correction to the LDA or a kind of approximation to $\Sigma_{xc}(r, r', E)$ has been recently discussed by Perdew (1986).

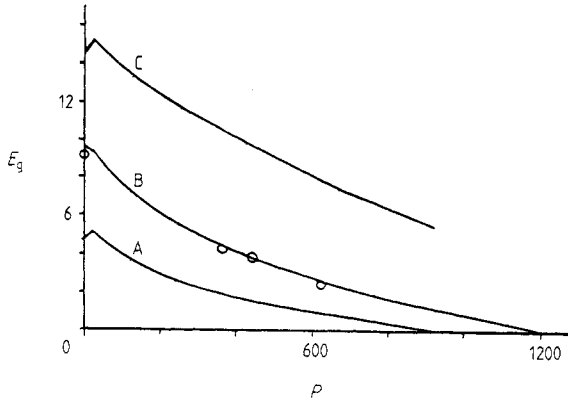


Figure 1. Pressure dependence of the minimum energy gap in Xe. Curve A represents full relativistic DFT calculations. Curve B is WKB-HF approach described in § 4. Curve C is obtained from the first and second ones after equation (11). The fitting for ϵ_0 uses experiments of Itie and Le Toullec (1984) and Makarenko and Weil (1982). Circles denote available experimental data (Makarenko and Weil 1982, Syassen 1982).

where the local Fermi momentum $k_F(r)$ is related to the local density $n(r)$ and $\Sigma_{xc}(p, E)$ is the well known self-energy operator in the homogeneous electron gas. Equation (14) has been used successfully in metals (e.g. Liberman 1971). However, it has been mentioned (Levine and Louie 1982, Horsch 1985) that for an insulator the homogeneous electron gas is a poor reference system. It is much better to use as a reference system, for instance, the Penn (1962) model. In this model the electron spectrum possesses spherical symmetry and has a gap over the Fermi sphere. This gap is a parameter of the model. The wavefunction is given by the linear combination of two plane waves:

$$\psi_p(r) = a_p e^{ipr} + b_p e^{i(2k_F - p)r}. \quad (15)$$

The quasi-classical approach to this model demands the substitution of p in equation (15) by the local momentum $p(r, E)$. The equation for p has the following form:

$$\begin{aligned} \frac{\hbar^2 p^2(r)}{2m} - \frac{\hbar^2 k_F^2(r)}{2m} + \frac{4\hbar^2}{2m} k_F(r) b^2 [p(r)] \{k_F(r) - p(r)\} \\ = E_{\vec{k}\vec{\lambda}}^{\text{HF}} - E_F - \{\Sigma_x^{\text{loc}}[p(r)] - V_{xc}(r)\} \end{aligned} \quad (16)$$

where $\Sigma_x^{\text{loc}}[p(r)]$ is calculated in the Penn model. A very similar approach has been used by Horsch (1985), but he has simply used instead of $p(r)$ the value of p averaged over the Wigner-Seitz cell.

The results of the quasi-classical calculation (15) for Ne and Ar (table 4) demonstrate the sufficient accuracy of this method. Usage of the quasi-classical HF approach combined with equation (11) demands negligible computer time in comparison to the DFT band structure calculations themselves.

It is worth noting that in some cases, namely in rare-gas crystals, the WKB-HF band structure depends strongly on the Penn gap E_g^{Penn} used in construction of Σ_x^{loc} (in Si this dependence is weak enough). However, this parameter may be fixed in the following manner.

Table 4. Inter-band transition energies (Ryd) in Ar and Ne calculated by the HF method and in the quasi-classical HF approximation (WKB-HF).

Ar	Transition	HF ^a	WKB-HF ^b
	$\Gamma^{v15}-\Gamma c1$	1.36	1.44
	$\Gamma c1-\Gamma c25'$	0.61	0.65
	$X^{v4'}-Xc1$	1.63	1.73
	$L^{v2'}-Lc1$	1.68	1.70
	$L^{c1}-Lc3$	0.33	0.42
Ne	Transition	HF ^a	WKB-HF ^c
	$\Gamma^{v15}-\Gamma c1$	1.87	1.86
	$\Gamma c1-\Gamma c25'$	1.33	1.29
	$X^{v41}-Xc1$	2.38	2.40
	$L^{v2'}-Lc1$	2.35	2.27
	$L^{c1}-Lc3$	0.77	0.87

^a Dagens and Perrot (1972).

^b $E_{\text{Penn}}^{\text{g}} = 1$ Ryd.

^c $E_{\text{Penn}}^{\text{g}} = 1.6$ Ryd.

Let us neglect the difference between the Hartree-Fock approximation (HFA) and exchange LDA wavefunctions. The energy spectrum in the HFA we shall approximate by the WKB-HF method as described above. The total energy in the HFA must be equal to the total energy in the exchange-only DFT which is well represented by the exchange-LDA total energy. Then, for these energies we have

$$E_{\text{tot}}^{\text{HF}} = \sum_i^{\text{occ}} E_i^{\text{HF}} - \frac{1}{2} \int V_{\text{H}}(r)n(r) \, dr - \int V_{\text{ext}}(r)n(r) \, dr - \frac{1}{2} \sum_i^{\text{occ}} \int V_{\text{F}}(r, r')\psi_i^*(r)\psi_i(r') \, dr \, dr' \quad (17)$$

$$E_{\text{tot}}^{\text{DFT}} = \sum_i^{\text{occ}} E_i^{\text{DFT}} - \frac{1}{2} \int V_{\text{H}}(r)n(r) \, dr - \int V_{\text{ext}}(r)n(r) \, dr + \int [E_{\text{x}}^{\text{DFT}}\{n(r)\} - V_{\text{x}}^{\text{DFT}}\{n(r)\}]n(r) \, dr. \quad (18)$$

In the WKB-HF approach we assume that

$$\int V_{\text{F}}(r, r')\psi_i(r') \, dr' = V_i(r)\psi_i(r) \quad (19)$$

(as far as we neglect the difference between the wavefunctions). Then, equality of (17) and (18) gives:

$$\sum_i^{\text{occ}} E_i^{\text{HF}} - \frac{1}{2} \sum_i^{\text{occ}} \int V_i(r)|\psi_i(r)|^2 \, dr = \sum_i^{\text{occ}} E_i^{\text{DFT}} + \int [E_{\text{x}}^{\text{DFT}}\{n(r)\} - V_{\text{x}}^{\text{DFT}}\{n(r)\}]n(r) \, dr. \quad (20)$$

Table 5. Determining the gap parameter ($E^{\text{Penn}}_{\text{g}}$) in WKB-HF approach by use of equation (24). Brackets denote averaging over occupied bands (three k -points used). Condition (24) demands that $\langle E^{\alpha=1} - E^{\text{HF}} \rangle$ be equal to zero. All energies are in Ryd.

Argon $E^{\text{Penn}}_{\text{g}}$	$\langle E^{\alpha=1} - E^{\alpha=2/3} \rangle = -0.53$		
	$\langle E^{\text{WKB-HF}} - E^{\alpha=1} \rangle$	$E^{\text{WKB-HF}}_{\text{g}}$	E^{HF}_{g}
0.2	0.30	1.04	
0.4	0.12	1.20	
0.6	0.09	1.30	
0.8	0.03	1.36	
1.0	0.00	1.42	1.36
1.2	-0.04	1.46	

Moreover, in the WKB-HFA:

$$E_i^{\text{HF}} - E_i^{\text{DFT}} = \int [V_i(r) - V_x^{\text{DFT}}(r)] |\psi_i(r)|^2 dr \quad (21)$$

and hence

$$\frac{1}{2} \sum_i^{\text{occ}} \int V_i(r) |\psi_i(r)|^2 dr = \int \mathcal{E}_x^{\text{DFT}} \{n(r)\} n(r) dr. \quad (22)$$

In LDA (exchange-only) $\mathcal{E}_x(n) = \frac{3}{4} V_x(n)$, then

$$\sum_i^{\text{occ}} \int V_i(r) |\psi_i(r)|^2 dr = \frac{3}{2} \int V_x(r) n(r) dr. \quad (23)$$

In other words

$$\sum_i^{\text{occ}} E_i^{\text{HF}} - \sum_i^{\text{occ}} E_i^{\text{DFT}} = \frac{1}{2} \int dr V_x^{\text{DFT}}(r) n(r). \quad (24)$$

This means that the well known Slater result that the average Fock potential is equal to $\frac{3}{2}$ of the KS exchange-only potential is true not only for a homogeneous electron gas but (in WKB approximation) for any density distribution. In fact, for occupied bands the main effect of changing α in the $X\alpha$ approach is the rigid displacement of bands; shape variations are much more subtle, so only a few energies are sufficient in sums in (24). Table 5 illustrates how equation (24) may be used to fix the Penn gap parameter in the WKB-HF procedure.

5. Application to the pressure dependence of dielectric gaps

In recent years, calculations of the gap pressure coefficients and the metallisation pressures have attracted much attention. Some authors (Rodrigues *et al* 1985, Van Camp *et al* 1986, and others) state that despite the error in the gap itself, its pressure coefficients dE_g/dp and even d^2E_g/dp^2 are reproduced well by the LDA calculation. Others believe that the metallisation pressure p_M is correct in the LDA (Satpathy *et al* 1985, and others). Obviously, both statements cannot be true simultaneously. For example, if p_M is correct one would expect an error in dE_g/dp of the order $\Delta E_g/p_M$,

where ΔE_g is the error of E_g . However, these two statements have never been applied to the same system: pressure coefficients are usually calculated for classical semiconductors (Si, GaAs, etc.) while p_M calculations were performed for such materials as Xe, CsI, etc.

We have calculated pressure coefficients for Si using equation (11) and quasi-classical HF and have found a very small correction to the LDA. This was due to the fact that ϵ_0 for Si is large and does not depend greatly on pressure. The difference between the pressure dependence of the HF and LDA gaps was also small. So for Si (as well as for other classical semiconductors) it is possible to use LDA for pressure coefficients. This is not the case, for instance, for large-gap insulators. We have calculated the pressure dependence of the minimal gap in Xe, up to metallisation, using a simple fit to available experimental points for ϵ_0 (figure 1). In this case $E_g^{\text{HF}} - E_g^{\text{LDA}}$ does depend on pressure, as well as ϵ_0 . Hence, dE_g/dp differs considerably from dE_g^{LDA}/dp . Moreover, the metallisation pressure for Xe is underestimated in LDA: at 920 kBar Xe becomes metallic in LDA but a 5.4 eV gap remains in the HF approximation, and our approach gives a 1.2 eV gap at the same pressure. We estimate the metallisation pressure for Xe as 1200 kbar.

6. Conclusions

The analyses of the structure of the exchange–correlation part of the self-energy $\Sigma_{xc}(r, r', E)$ of the Green function allows us to make the following conclusions.

In metals all parts of Σ_{xc} are comparatively short-ranged and thus the electron spectrum may be reasonably described by use of the potential of density-functional theory.

In insulators there is a long-range contribution that corresponds to the statically screened exchange interaction between the states $|k\lambda\rangle$ and $|k'\lambda'\rangle$ where $\lambda = \lambda'$ and $k \approx k'$. This contribution cannot be described by any local potential.

This non-local (long-range) part of Σ_{xc} will coincide with the analogous part of the Σ_x^{HF} in the Hartree–Fock approximation if the latter is divided by ϵ_0 , the macroscopic dielectric constant.

The practical results that stem from these facts are as follows.

Correct description of the electron spectrum, particularly of the dielectric gap, in insulators and semiconductors is impossible without taking into account the long-range part of Σ_{xc} .

In the first approximation the correction to a DFT spectrum may be derived by equation (11), which is in fact a kind of interpolation between the DFT and the Hartree–Fock method. The accuracy of this prescription is an order higher than that of the HF or of the DFT methods themselves. This is true for semiconductors, wide-gap insulators and ionic crystals.

The necessary HF spectrum may be derived accurately enough by the very simple quasi-classical technique. The total computer time for the calculation of the correction to the band structure is negligible; the corresponding program is very simple.

This method does not make use of any adjustable parameters; it is necessary, however, to know from an experiment the dielectric constant ϵ_0 .

References

- Carlsson A E 1985 *Phys. Rev. B* **31** 5178–82
 Dagens L and Perrot F 1972 *Phys. Status Solidi* **52** 163–72

- Godby R W, Schlüter M and Sham L J 1987 *Phys. Rev. B* **36** 6497–500
- Guinea F and Tejedor C 1980 *J. Phys. C: Solid State Phys.* **13** 5515–27
- Heaton R E, Harrison J G and Lin C C 1983 *Phys. Rev. B* **28** 5992–6006
- Hedin L 1965 *Phys. Rev.* **139** A796–803
- Hedin L and Lundqvist S 1969 *Solid State Phys.* **23** 1–189 (New York: Academic)
- Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864–72
- Horsch P 1985 *Solid State Commun.* **54** 741–4
- Hybertsen M S and Louie S G 1986 *Phys. Rev. B* **34** 5390–413
- Itie J P and Le Toulec R 1984 *J. Physique Coll.* **45** C8 53–6
- Kohn W and Sham L J 1965 *Phys. Rev.* **136** A1133–41
- Levine Z H and Louie S G 1982 *Phys. Rev. B* **25** 6310–6
- Liberman D A 1971 *Phys. Rev. A* **3** 1224–9
- Lipari N O and Kunz A B 1971 *Phys. Rev. B* **3** 491–9
- MacDonald A H 1980 *J. Phys. F: Met. Phys.* **10** 1737–44
- Makarenko I and Weil G 1982 *Phys. Rev. B* **26** 7113–6
- Manghi F, Riegler G, Bertoni C M and Bachelet G B 1985 *Phys. Rev. B* **31** 3680–8
- Ohkosi I 1985 *J. Phys. C: Solid State Phys.* **18** 5415–27
- Pantelides S T, Mickish D S and Kunz A B 1974 *Phys. Rev. B* **10** 5203–12
- Penn D R 1962 *Phys. Rev.* **128** 2093–6
- Perdew J P 1986 *Int. J. Quantum Chem., Quantum Chem. Symp.* **19** 497–523
- Perdew J P and Levy M 1983 *Phys. Rev. Lett.* **51** 1884–7
- Perdew J P and Zunger A 1981 *Phys. Rev. B* **23** 5048–79
- Perrot F 1972 *Phys. Status Solidi* **52** 163–9
- Pickett W E 1985 *Comment. Solid State Phys.* **12** 1–9
- Pickett W E and Wang C S 1984 *Phys. Rev. B* **30** 4719–33
- Rodrigues C O, Peltzer Y Blanca E L and Cappannini O M 1985 *Solid State Commun.* **56** 575–9
- Satpathy S, Christensen N E and Jepsen O 1985 *Phys. Rev. B* **32** 6793–9
- Schreiber M and Bross U 1983 *J. Phys. F: Met. Phys.* **13** 1895–901
- Sham L J and Kohn W 1966 *Phys. Rev.* **145** B561–9
- Sham L J and Schlüter M 1983 *Phys. Rev. Lett.* **51** 1888–91
- Sterne P A and Inkson J C 1984 *J. Phys. C: Solid State Phys.* **17** 1497–510
- Strinati G, Mattausch H G and Hanke W 1982 *Phys. Rev. B* **25** 2867–88
- Svane A 1987 *Phys. Rev. B* **35** 5496–502
- Syassen K 1982 *Phys. Rev. B* **25** 6548–51
- Trickey S B, Ray A K and Worth J P 1981 *Phys. Status Solidi b* **106** 613–20
- Van Camp P E, Van Doren V E and Devreese J T 1986 *Phys. Rev. B* **34** 1314–6