The Harris functional applied to surface and vacancy formation energies in aluminium

The Harris functional applied to surface and vacancy formation energies in aluminium

M W Finnis

Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 1000 Berlin 33, Federal Republic of Germany

Abstract. Non-self-consistent calculations of total energies using the Hohenberg–Kohn–Sham functional and the approximate functional of Harris have been made for a three-layer (111) slab of aluminium and compared to fully self-consistent calculations. The norm-conserving pseudopotential of Bachelet, Hamann and Schlüter was used. The Harris total energy for the slab, which is calculated from an input charge density constructed by super-imposing free atomic charge densities, is closer to the self-consistent energy for reasons which are discussed in detail. Nevertheless, the corresponding Harris surface energies are not sufficiently accurate to be useful. By renormalising the atomic charge densities which are superimposed to form the input charge density, in particular by pushing charge from the tails towards the cores, the surface energy from a self-consistent calculation can be accurately reproduced by the Harris functional. Furthermore the Harris functional with the same renormalised atoms accurately re-produces the surface-layer contraction and also the for-mation energy of a vacancy (to within 0.01 eV) obtained from self-consistent calculations.

1. Introduction

It is now possible to calculate from first principles the total ground state energies of small clusters of atoms, or of periodic systems with up to a few tens of atoms per unit cell, using the Hohenberg–Kohn–Sham (HKS) density functional formalism. The only essential approximation involved is the local density approximation, according to which the exchange and correlation energy of an inhomogeneous electron gas takes the form $\int \mathrm{d}r n(r) \varepsilon_{xc}(n(r))$, where $\varepsilon_{xc}(n)$ is the exchange and correlation energy per electron in a uniform electron gas of density n, a quantity which is now well known for the range of densities of physical interest (Ceperley and Alder 1980). The effective one-electron potential is constructed from a charge density which is constructed from wavefunctions which are solutions of the one-electron Schrödinger equation, and this process is iterated to self-consistency. Furthermore, in favourable cases the self-consistent total energy may be minimised with respect to the atomic positions or the dynamics of the atoms can be followed, using for example the method devised by Car and Parrinello (1985). The relative energies and predicted structures obtained by self-consistent methods are generally reliable, but the calculations are computationally very demanding.

For this reason, and with a view also to obtaining physical insight, it is continuing interest to seek approximate schemes based on the Hohenberg–Kohn–Sham density functional formalism. Two possible non-self-consistent schemes which have been sug-gested by previous authors are investigated here, with particular attention to the one

suggested by Harris (1985) and studied independently by Foulkes (1987). The approximate schemes start with a guess for the electronic charge density, and solve the Kohn–Sham–Schrödinger equation for the electronic wave functions. The energy functionals we are considering are constructed from the results of this solution, with no iterations towards self-consistency.

The obvious initial charge density to choose is that obtained by superimposing the free atom charge densities. With this prescription Harris showed that his functional gives a good account of the bond lengths and vibrational frequencies of a number of homopolar dimers. Polatoglou and Methfessel (1988) tested the Harris (H) functional on the cohesive properties of a number of solids (Be, Al, V, Fe, Si and NaCl). They calculated cohesive energies, lattice constants and bulk moduli, and found that the deviations from the values obtained by the much more costly self-consistent calculations were quite small, e.g. up to 15% in the case of the bulk modulus of V, with generally better agreement for lattice constants than for bulk moduli. It is particularly noteworthy that the ionic crystal NaCl is no worse described than the metals by this model based on superimposing neutral atom charge densities.

Foulkes (1987) and Foulkes and Haydock (1989) also illustrated the accuracy of this functional for the dimers H–H, He–He, H–He and Ge–Ge, and studied in detail the derivation of a simple tight-binding model from it. The Harris model was used by Sutton *et al* (1988) as a basis for deriving the tight-binding bond model, which involves simplifications such as the two-centre approximation and a limited basis of localised orbitals. An important assumption in the practical application of this and other tight-binding models is that part of the total energy can be represented as a sum over pairs of atoms of a repulsive interatomic potential. The derivation of such a term follows from the assumption of superimposed spherical atomic-like charge densities from which the electrostatic energy is obtained within the Harris functional. It is therefore important to understand in what situations the Harris functional is reliable if one wants to justify schemes such as these which are based on further simplifications to it.

A more obvious approximation to the total energy than the Harris functional is simply given by the HKS functional without iterating beyond the first solution of the Schrödinger equation. This prescription, with the input charge density constructed by superimposing atomic charge densities, was successfully applied for example by Chelikowsky and Louie (1984) to calculate the bulk properties of diamond. It might be expected to perform better than the Harris model, since unlike the latter it uses the output wavefunctions in its construction as described in the following section. It has the further attraction of providing a variational upper bound to the self-consistent energy. However, in practice, as we shall see, the Harris model is often the more accurate.

The present paper puts these approximate schemes, which are described in detail in the following section, to a stringent test, by calculating the surface energy of a (111) three-layer slab of aluminium. This test case is of particular interest because there are already three published calculations which give somewhat different answers, notably in the sign of the surface relaxation (Payne *et al* 1989b, Feibelman 1983, Batra *et al* 1986). As will be seen, the Harris functional gives an excellent account of the surface energy and the surface relaxation, but only when the starting 'atoms' are renormalised. Furthermore, the renormalised atoms and the Harris functional also describe the vacancy formation energy accurately in comparison with a self-consistent calculation (the error is about 0.01 eV), and even give a more accurate energy for the dimer than the Harris functional with superimposed free atom charges. The physical reasons for these conclusions are discussed.

2. Approximate density functionals

The total energy of a system of electrons and ions is described by the minimum of the Hohenberg–Kohn–Sham energy functional

$$E^{\text{HKS}}[n] = T[n] + \int nV_{\text{ext}} + \tfrac{1}{2}\iint \frac{nn}{|r - r'|} + \int n\varepsilon_{\text{xc}}(n) + E_{\text{ion}}. \tag{1}$$

The first term is the kinetic energy of non-interacting electrons having the density n. V_{ext} is the potential of the ions. The third and fourth terms represent the Hartree and exchange–correlation energy of the electrons. The final term is the mutual electrostatic interaction energy of the ion cores which will not play any part in following discussion, since it is independent of the functional used to describe the electrons. The charge density is expressed as

$$n(r) = \sum_i f_i \psi_i^*(r)\psi_i(r) \tag{2}$$

where for the ground state the wavefunctions ψ_i are solutions of the single-particle Kohn-Sham Schrödinger equation

$$-\tfrac{1}{2}\nabla^2 \psi_i + V_{\text{eff}}\psi_i = \varepsilon_i \psi_i \tag{3}$$

with an effective potential

$$V_{\text{eff}}[n] = V_{\text{ext}} + \int \frac{n}{|r - r'|} + \mu_{\text{xc}}(n) \tag{4}$$

and $\mu_{\text{xc}}(n)$ is the local exchange–correlation potential

$$\mu_{\text{xc}}(n) = (\mathrm{d}/\mathrm{d}n)[n\varepsilon_{\text{xc}}(n)]. \tag{5}$$

We would like the occupancies f_i to take the value 2 for energies ε_i below the Fermi energy and 0 for higher energies. However, this sharp Fermi cut-off leads to well-known difficulties in the integration of quantities over the occupied states within the Brillouin zone. The solution employed here is to smear out the Fermi energy with a fictitious temperature T, which improves the accuracy of the sampling method of k-space integration. In the finite temperature theory which I have used for calculations the occupancies f_i are determined by the Fermi factors

$$f_i = 2/[\exp(\beta(\varepsilon_i - \varepsilon_F)) + 1] \tag{6}$$

where

$$\beta = 1/k_B T.$$

The self-consistent solution is obtained at a minimum of the free energy A defined by

$$A = E^{\text{HKS}}[n] - TS \tag{7}$$

where the entropy is defined by

$$S = -2k_B \sum_i [f_i \ln f_i + (1 - f_i)\ln(1 - f_i)]. \tag{8}$$

The estimated zero-temperature limit of the energy functional, which is the final result of the calculations is then

$$E^{\text{HKS}}(T = 0) \simeq E^{\text{HKS}}(T) - TS/2 \tag{9}$$

as discussed by Gillan (1989).

If we start with a guessed input charge density n^{in}, usually constructed by super-imposing the calculated charge densities of free atoms, we can solve the Kohn-Shan Schrödinger equation with the effective potential $V_{eff}[n^{in}]$ to obtain what we call output wavefunctions and eigenvalues, ψ_i^{out} and ε_i^{out}. We might then follow an iterative procedure of some kind to obtain the self-consistent charge density n^{sc}. The iteration would start by constructing a charge density from the output wavefunctions and their occupancies, let us call it n^{out}. With $n = n^{out}$ the Hohenberg-Kohn functional in (1) is an upper bound to the self-consistent energy. The functional $E^{HKS}[n^{out}]$ can be written in an alternative form, which we can get by substituting for the kinetic energy in terms of the eigenvalue sum, using the definition of the kinetic energy:

$$T[n^{out}] = \sum_i f_i^{out} \varepsilon_i^{out} - \int n^{out} V_{eff}[n^{in}]. \tag{10}$$

The result is

$$E^{HKS}[n^{out}] = \sum_i f_i^{out} \varepsilon_i^{out} - \int\int \frac{n^{out}n^{in}}{|r - r'|} + \tfrac{1}{2}\int\int \frac{n^{out}n^{out}}{|r - r'|} - \int n^{out}\mu_{xc}(n^{in})$$
$$+ \int n^{out}\varepsilon_{xc}(n^{out}) + E_{ion}. \tag{11}$$

The other approximation I will discuss is the functional introduced by Harris (1985) and Foulkes (1987). It is obtained from the form in equation (11) if we replace all the explicit charge densities by n^{in}. Denoting the Harris functional by E^H it is

$$E^H[n^{in}] = \sum_i f_i^{out} \varepsilon_i^{out} - \tfrac{1}{2}\int\int \frac{n^{in}n^{in}}{|r - r'|} - \int n^{in}\mu_{xc}(n^{in}) + \int n^{in}\varepsilon_{xc}(n^{in}) + E_{ion}. \tag{12}$$

This functional, regarded as a functional of n^{in}, is stationary at the self-consistent charge density n^{sc}. In practice it is a lower bound to the exact energy $E^{HKS}[n^{sc}]$. As discussed further below there is a plausible but non-rigorous argument to explain this.

Subtracting $E^H[n^{in}]$ from $E^{HKS}[n^{out}]$ we find after some algebra

$$E^{HKS}[n^{out}] - E^H[n^{in}] = \tfrac{1}{2}\int\int \frac{(n^{out} - n^{in})(n^{out} - n^{in})}{|r - r'|} - \int (n^{out} - n^{in})\mu_{xc}(n^{in})$$
$$+ \int n^{out}\varepsilon_{xc}(n^{out}) - \int n^{in}\varepsilon_{xc}(n^{in}). \tag{13}$$

Equation (13) is exact. We can expand the exchange-correlation energy ε_{xc} about n^{in} and make use of (5) so that to second order in $n^{out} - n^{in}$

$$E^{HKS}[n^{out}] - E_*^H[n^{in}] = \tfrac{1}{2}\int\int C(r, r')[n^{out}(r) - n^{in}(r)][n^{out}(r') - n^{in}(r')] \tag{14}$$

where we have written explicitly the r-dependence of the densities in order to make clear the meaning of the kernel C which is

$$C(r, r') = 1/|r - r'| + \{d\mu_{xc}[n(r)]/dn(r)\}|_{n = n^{in}} \delta(r - r'). \tag{15}$$

Now let us obtain expressions for the separate errors in $E^{HKS}[n^{out}]$ and $E^H[n^{in}]$ compared to the self-consistent energy $E^{HKS}[n^{sc}]$. The derivation closely follows Foulkes (1987) and Foulkes and Haydock (1989) but goes a bit further by obtaining completely explicit expressions for all the terms in the Taylor expansion of the energy. Considering the difference

$$E^{HKS}[n^{out}] - E^{HKS}[n^{sc}] = T[n^{out}] - T[n^{sc}] + \int (n^{out} - n^{sc})V_{ext} + \tfrac{1}{2}\int\int \frac{n^{out}n^{out}}{|r - r'|}$$
$$- \tfrac{1}{2}\int\int \frac{n^{sc}n^{sc}}{|r - r'|} + \int n^{out}\varepsilon_{xc}(n^{out}) - \int n^{sc}\varepsilon_{xc}(n^{sc}) \tag{16}$$

we would like to express the kinetic energy difference in a form which is amenable to calculation. This can be done by making use of two separate stationary conditions which correspond to the effective independent particle problem and to the self-consistent problem respectively, namely

$$(\delta T/\delta n)|_{n^{\text{out}}} + V_{\text{eff}}[n^{\text{in}}] = 0 \tag{17}$$

and

$$(\delta T/\delta n)|_{n^{\text{sc}}} + V_{\text{eff}}[n^{\text{sc}}] = 0. \tag{18}$$

Hence, expanding about the self-consistent density,

$$\int \int \left(\frac{\delta^2 T}{\delta n^2} (n^{\text{out}} - n^{\text{sc}}) + \tfrac{1}{2} \int \frac{\delta^3 T}{\delta n^3} (n^{\text{out}} - n^{\text{sc}})^2 + \frac{(n^{\text{in}} - n^{\text{sc}})}{|r - r'|} + \mu_{\text{xc}}(n^{\text{in}}) \right.$$

$$\left. - \mu_{\text{xc}}(n^{\text{sc}}) \right) (n^{\text{out}} - n^{\text{sc}}) = 0. \tag{19}$$

By combining equations (13), (16) and (19) the following results can be obtained:

$$E^{\text{HKS}}[n^{\text{out}}] = E^{\text{HKS}}[n^{\text{sc}}] + \tfrac{1}{2} \int \int \frac{(n^{\text{out}} - n^{\text{in}})(n^{\text{out}} - n^{\text{sc}})}{|r - r'|} + \int n^{\text{out}} \varepsilon_{\text{xc}}(n^{\text{out}}) - \int n^{\text{sc}} \varepsilon_{\text{xc}}(n^{\text{sc}})$$

$$- \tfrac{1}{2} \int [\mu_{\text{xc}}(n^{\text{in}}) + \mu_{\text{xc}}(n^{\text{sc}})](n^{\text{out}} - n^{\text{sc}}) - \tfrac{1}{12} \int \int \int \frac{\delta^2 T}{\delta n^3} (n^{\text{out}} - n^{\text{sc}})^3$$

$$+ \text{ higher-order kinetic energy terms} \tag{20}$$

$$E^{\text{H}}[n^{\text{in}}] = E^{\text{HKS}}[n^{\text{sc}}] + \tfrac{1}{2} \int \int \frac{(n^{\text{out}} - n^{\text{in}})(n^{\text{in}} - n^{\text{sc}})}{|r - r'|} + \int n^{\text{in}} \varepsilon_{\text{xc}}(n^{\text{in}}) - \int n^{\text{sc}} \varepsilon_{\text{xc}}(n^{\text{sc}})$$

$$- \tfrac{1}{2} \int \mu_{\text{xc}}(n^{\text{sc}})(n^{\text{out}} - n^{\text{sc}}) + \int \mu_{\text{xc}}(n^{\text{in}})(\tfrac{1}{2}n^{\text{out}} + \tfrac{1}{2}n^{\text{sc}} - n^{\text{in}})$$

$$- \tfrac{1}{12} \int \int \int \frac{\delta^3 T}{\delta n^3} (n^{\text{out}} - n^{\text{sc}})^3$$

$$+ \text{ higher-order kinetic energy terms.} \tag{21}$$

We note that the higher-order errors in (20) and (21) are entirely in $T[n]$ and as we can see from equation (13) these higher-order errors are the same for each functional.

To second order in the charge density differences (20) and (21) reduce to

$$E^{\text{HKS}}[n^{\text{out}}] = E^{\text{HKS}}[n^{\text{sc}}] + \tfrac{1}{2} \int C(r, r')[n^{\text{out}}(r) - n^{\text{in}}(r)][n^{\text{out}}(r') - n^{\text{sc}}(r')] \tag{22}$$

$$E^{\text{H}}[n^{\text{in}}] = E^{\text{HKS}}[n^{\text{sc}}] + \tfrac{1}{2} \int C(r, r')[n^{\text{out}}(r) - n^{\text{in}}(r)][n^{\text{in}}(r') - n^{\text{sc}}(r')]. \tag{23}$$

We see that the second-order error in the Harris energy differs from that in the HKS energy by the appearance of the factor $(n^{\text{in}} - n^{\text{sc}})$ in place of $(n^{\text{out}} - n^{\text{sc}})$.

3. Calculations

The surface calculations were performed with a trigonal unit cell containing three aluminium atoms, representing a (111) slab consisting of three close-packed atomic
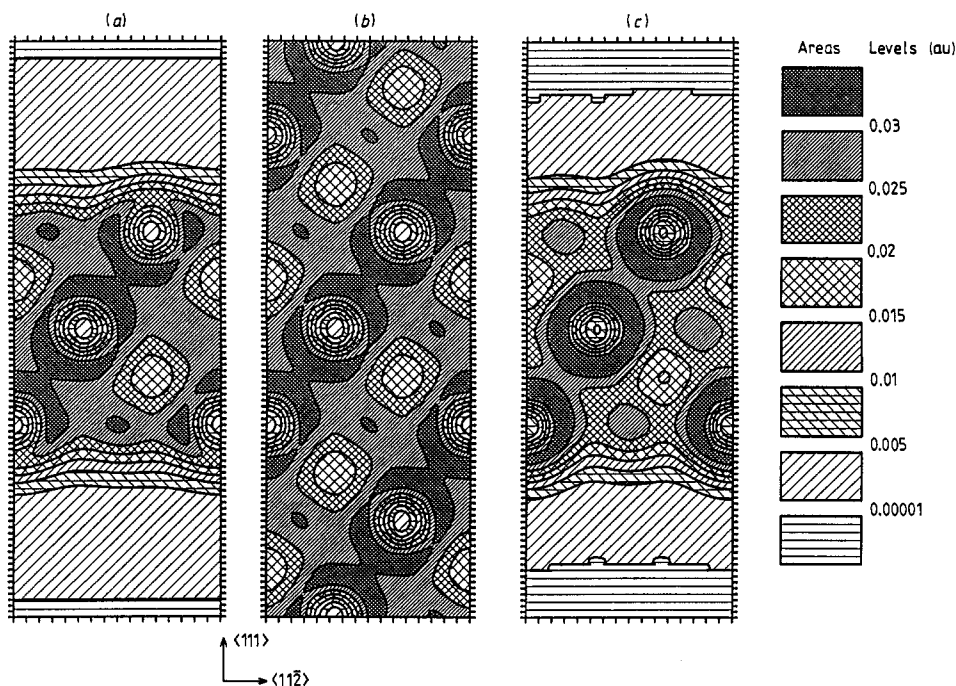
Figure 1. The valence electron charge density in aluminium. The atomic sites are identifiable by the concentric contours. The contours were obtained with the FARB-E-2D algorithm of Preusser (1989). (*a*) Self-consistent calculation on a three-layer slab. (*b*) Self-consistent calculation for the bulk. (*c*) Superposition of fitted atomic charge densities on a three-layer slab.

layers. The bulk lattice parameter of 7.6 au was assumed in all the calculations. The height of the cell was chosen so that it could be filled by a further three layers of aluminium. Periodic boundary conditions were applied. Bulk aluminium was calculated with the same cell containing all six layers of atoms. Thus the slab calculations were actually performed with a supercell in which three layers of missing atoms separated the periodically repeated slabs.

For the vacancy, a similar cell was used but doubled in the x and y dimensions and halved in the z direction. It contained 12 bulk atoms, of which one was removed to create the vacancy. Although this is a rather small supercell, certainly too small to relax the atoms around the vacant site, Gillan's calculations (Gillan 1989) suggest that the unrelaxed vacancy formation energy should be within about 0.1 eV of that of the isolated vacancy.

The non-local, norm-conserving pseudopotential of Bachelet *et al* (1982) was used, together with the exchange and correlation function of Ceperley and Alder (1980) as parameterised by Perdew and Zunger (1981). In order to save computation time, the non-local d component of the potential and all higher l components were set equal to the p̄ component. In this way only s non-locality has to be treated explicitly. The charge densities shown in figure 1 and the error analysis to be described used this approximation to the original pseudopotential. Including the d non-locality explicitly appeared to make no significant difference to the charge densities, but it changed the surface energy by 0.15 eV per surface atom. For this reason the subsequent calculations, including the relaxation of the interlayer spacing and the vacancy calculations were made with the full

non-local pseudopotential. A useful saving in time was also achieved by using a separable form of the pseudopotential, as discussed by Kleinman and Bylander (1982).

Brillouin zone sampling used 12 k-points, two sets of the 6 given by Chadi and Cohen (1973) for hexagonal symmetry, displaced in the z direction by $\frac{1}{4}$ and $\frac{3}{4}$ of the zone edge $2\pi/c$. The wavefunctions were expanded in plane waves up to a cutoff of 8 Ryd. Increasing the cutoff to 12 Rydbergs changed the calculated surface energy of 0.427 eV per surface atom by only 0.002 eV. The charge densities shown in figure 1 were calculated with the higher cutoff. For the self-consistent calculations, the wave-functions were relaxed by the modified steepest-descents version of the Car–Parrinello method described by Payne *et al* (1989a). If we denote by $E(N)$ the energy of the cell with N atoms, we have for the surface energy per surface atom

$$\gamma = \tfrac{1}{2}E(3) - \tfrac{1}{4}E(6) \tag{24}$$

and for the unrelaxed vacancy formation energy

$$E_v^f = E(11) - \tfrac{11}{12}E(12). \tag{25}$$

Total energies $E^H[n^{in}]$ were calculated with the Harris model assuming in the first instance a superposition of atomic charge densities. From the output wave functions the output charge density was constructed and the HKS energy was evaluated. Finally, the self-consistency procedure of Payne *et al* (1989) was followed, to minimise the HKS functional. In fact the function actually minimised was the fictitious free energy A as described in the previous section, from which the best estimate of the zero-temperature HKS functional was obtained. A second initial charge density was constructed by super-imposing what we refer to as renormalised atoms. These were first of all constructed by fitting the atomic charge densities with three Gaussians, the main effect being to reduce the size of the tails of the atomic charge densities. $E^H[n^{in}]$ and $E^{HKS}[n^{out}]$ were calculated as before. The results are shown in table 1.

The improvement in $E^H[n^{in}]$ and $E^{HKS}[n^{out}]$ achieved by renormalising the free atoms is dramatic. The fitting of the atomic charge density was not designed to produce this result, so a more systematic approach could be expected to do even better. The approach I followed was to multiply the free atom charge density by a factor

$$A/(e^{\beta(r-r_c)} + 1). \tag{26}$$

β and the cutoff radius r_c were varied to maximise the Harris functional for the three-layer slab $E^H(3)$. We are exploiting here the hypothesis that the Harris functional is a lower bound to the self-consistent energy. The maximisation does not need to be very accurate to find a renormalised atom with gives a very good E^H. We find with $\beta = 5$ and $r_c = 4.5$ (au) that $E^H(3) = -168.975$ eV compared to the self-consistent $E^{HKS}(3) = -168.973$ eV. The new renormalised atoms were then used with the fully non-local pseudopotential and the total energy of the three-layer slab was calculated as the interplanar spacing was varied. Results are shown in figure 2, which compares $E^H(3)$ for the free and renormalised atom densities with the self-consistent results. The variation of the H and HKS functionals with the initial charge density is illustrated by plotting the two approximations in figure 3.

We find that the Harris unrelaxed vacancy formation energy with the free atom charge densities is 0.863 eV. With the renormalised atoms it is 1.046 eV. The self-consistent result is 1.056 eV.

Finally, I calculated the energy of a dimer, with the bulk interatomic spacing. The periodic cell was chosen so that the repeated dimers were separated by 22.8 au. The self-consistent energy was -107.106 eV, the Harris energy with free atoms was -107.489 eV and the Harris energy with renormalised atoms was -107.255 eV.
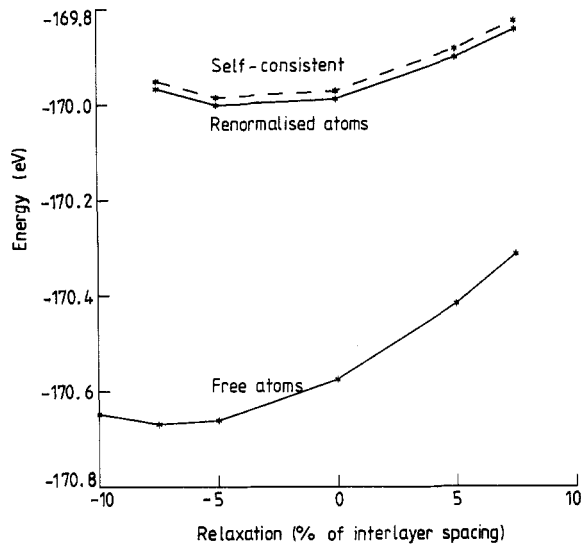
Figure 2. Total energy of the three-layer slab as a function of the surface relaxation.
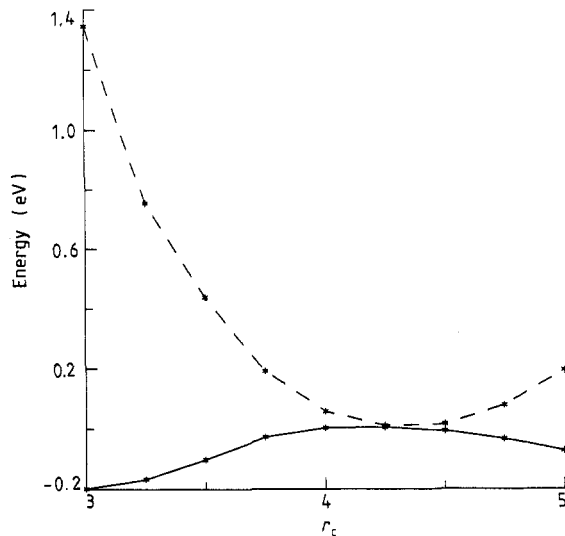

Figure 3. Errors in the Harris and HKS functionals as a function of r_c (equation (26)). Units are eV. Full curve: $E^H[n^{in}] - E^{HKS}[n^{sc}]$; broken curve: $E^{HKS}[n^{out}] - E^{HKS}[n^{sc}]$.

4. Discussion

Let us first consider the results shown in table 1. We see that neither of the non-self-consistent energy functionals $E^H[n^{in}]$ or $E^{HKS}[n^{out}]$ is reliable for calculating the surface energy using the free atom input charge density. The bulk energy is well described by either functional but there are significant errors in the calculated slab energies. Furthermore, the error in the slab energy is significantly greater with the HKS functional than with the H†. We first discuss these rather disappointing findings and then we discuss how the Harris functional can be greatly improved by using renormalised atoms.

† Very similar calculations and conclusions have recently been published by Read and Needs (1989).

Table 1. Energies calculated for the 3 atom cell (the slab) and the 6 atom cell (the bulk). The renormalised atom refers to the triple Gaussian fit to the atomic charge density. Units are eV per unit cell except for the surface energies γ which are in eV per surface atom. Note that these results were obtained with the non-local d component of the pseudopotential set equal to the p component.

		Free atoms	Renormalised atoms
3 atom slab	$E^{HKS}[n^{sc}]$	−168.977	−168.977
	$E^{HKS}[n^{out}]$	−167.192	−168.969
	$E^{H}[n^{in}]$	−169.455	−169.031
6 atom bulk	$E^{HKS}[n^{sc}]$	−339.054	−339.054
	$E^{HKS}[n^{out}]$	−339.050	−339.050
	$E^{H}[n^{in}]$	−339.151	−339.144
γ	$E^{HKS}[n^{sc}]$	0.275	0.275
	$E^{HKS}[n^{out}]$	1.167	0.278
	$E^{H}[n^{in}]$	0.060	0.271

Table 2. Errors in the H and HKS functionals for the three layer slab, starting from the superimposed free atom charge densities. Units are eV per unit cell. (*a*) Direct evaluation of error by subtraction of total energies. (*b*) Error to second order in charge density from (20)–(21). (*c*) Error omitting 3rd and higher orders in $T[n^{out}] - T[n^{sc}]$.

	$E^{H}[n^{in}] - E^{HKS}[n^{sc}]$	$E^{HKS}[n^{out}] - E^{HKS}[n^{sc}]$
(*a*)	−0.48	1.78
(*b*)	−0.43	1.81
(*c*)	−0.45	1.80

It is also noteworthy that the Harris energy always lies below the self-consistent energy. The HKS energy must always lie above the self-consistent energy from the variational principle, but we know in advance only that the Harris energy is stationary when the input charge density equals the self-consistent charge density.

We can understand these results by considering the expressions for the errors in H and HKS energies, equations (15) and (20)–(23). The Harris energy $E^{H}[n^{in}]$ differs from the HKS energy $E^{HKS}[n^{out}]$ only by the appearance of the factor $(n^{in} - n^{sc})$ in the integrand of equation (23) instead of the factor $(n^{out} - n^{sc})$. It is plausible that the sign of these factors is different over the important part of the range of integration because there is a tendency for the output charge density to overshoot the self-consistent charge density. We cannot make this argument into a rigorous proof that the Harris functional is a variational lower bound to the exact energy, but in all calculations reported here for the vacancy and surfaces and for other published calculations referred to previously it has turned out to be true. The question of why the H energy is better than the HKS energy for the surface can be answered with references to tables 2 and 3.

Firstly table 2 shows that the error in these functionals even with the free atom charge densities is well described by the second-order terms (22) and (23). Results (c) were calculated from equations (20) and (21) omitting third- and higher-order errors in the kinetic energy functional only, so the residual error, the difference between (a) and (c) is only due to kinetic energy terms.

Table 3 shows the electrostatic contribution to the errors given in row (a) of table 2. That is the part due to the Coulomb term in $C(r, r')$ (see equation (15)). It is helpful to

Table 3. Electrostatic contribution to the errors in the H and HKS functionals for the three layer slab. Units are eV. (*a*) Lowest Fourier component, $\pm 2\pi/c$. (*b*) Lowest two Fourier components, $\pm 2\pi/c$, $\pm 4\pi/c$. (*c*) All contributions.

	$E^H[n^{\text{in}}] - E^{\text{HKS}}[n^{\text{sc}}]$	$E^{\text{HKS}}[n^{\text{out}}] - E^{\text{HKS}}[n^{\text{sc}}]$
(*a*)	-0.37	1.21
(*b*)	-0.40	1.87
(*c*)	-0.77	1.97

think of these terms in reciprocal space, in which the Coulomb term scales as g^{-2}. The contributions of the two lowest Fourier components of the electrostatic error, which are normal to the slab, are shown in table 3. Comparing rows (a) of table 3 and table 2 we see that most of the error in the non-self-consistent energies comes from the longest wavelength electrostatic component ($g = \pm 2\pi/c$). Furthermore the Fourier components of factors involving n^{out} will be largest for long wavelengths, since they can be thought of as the response to an error in the input potential, which is the error in the input charge density multiplied by the Coulomb factor g^{-2}. The appearance of two n^{out} factors in $E^{\text{HKS}}[n^{\text{out}}]$ and only one in $E^H[n^{\text{in}}]$ explains the smaller error in $E^H[n^{\text{in}}]$. Figure 3 illustrates this for a particular variation of the charge density. For bulk crystal there are no such long-wavelength components of charge density or potential, so the errors are restricted to the contributions from higher Fourier components and are correspondingly much smaller.

The renormalised atoms have the property that these long-wavelength errors in the input density, and hence the input error in the potential, are much smaller. The variational principle gives us a much more accurate energy than one might have expected simply by looking at the charge densities shown in figure 1. The superimposed atomic charge density, even renormalised, is quite different to the self-consistent charge density for the slab, which within the surface layer actually looks very similar to that of the bulk crystal.

Although these may be the first total energy calculations which make use of renormalised atoms, the idea of renormalising atoms at the surface is not new. Weinert and Watson (1984) calculated the surface dipole contribution to work functions by using overlapping spherical atomic charge distributions and found them to be systematically too large. They found that renormalised atoms with contracted tails gave much improved values for the dipole barriers. Their renormalised atoms were constructed by solving for the atomic charge density with the atom inside a spherical constraining potential.

We see in figure 2 that the renormalised atoms give a good account not only of the total energy but also of the surface relaxation, which is inwards for this slab. This direction of relaxation is in agreement with the sign of the forces in the pseudopotential calculations of Batra *et al* (1986) but disagrees in sign with the calculation of Feibelman (1983) who used a LCAO method with Gaussian basis functions. On the other hand our surface energy of 0.43 eV/surface atom (unrelaxed) is close to Feibelman's 0.39 eV (relaxed) and lower than the 0.5 eV of Batra *et al* (unrelaxed). The most accurate calculation of the surface energy of the Al(111) surface is probably that of Payne *et al* (1989b), who used a 15 Ryd cutoff, 6 layers of atoms, and obtained 0.42 eV/surface atom (unrelaxed). Also shown in figure 2 is the energy versus interplanar spacing calculated with the Harris functional and *free*-atom charge densities. The error in the latter is not great in percentage terms, but too great to be useful for an estimate of the surface energy, which is obtained by subtracting large numbers. Furthermore the estimate of -7% for

the inward surface relaxation with the free-atom Harris functional is significantly greater than the -3% predicted with both the self-consistent and renormalised atom Harris calculations.

Finally we should ask if the renormalised Harris atoms are transferable; do they also work for defects other than the (111) surface. As further evidence we so far have the vacancy calculations. The formation energy we have obtained (1.056 eV) is somewhat higher than experiment (0.66 eV) and the only previous realistic calculation (Gillan 1989) (0.56 eV). Apart from the neglect of relaxation, the main reason for this is probably the small unit cell. It may also be the case that a better convergence with cell size could be obtained by calculating at constant volume per atom, as Gillan did, rather than at constant lattice parameter, but this remains to be investigated. For the present purpose of assessing the energy functionals these systematic errors are not important. The vacancy calculations show that the error in using the renormalised atoms is only 0.01 eV, whereas the free-atom charge densities give an error of 0.2 eV. This leads us to suggest, with some caution, that the renormalised atoms might give an acceptably accurate energy for a range of defect calculations.

Even the dimer energy is better described by the renormalised atoms (error = -0.15 eV) than the free atoms (error = -0.38 eV). The renormalisation has the effect of setting up more charge between the atoms, along the nearst-neighbour bonds. The self-consistent charge density is also more heaped up in the bonds than that of the free atoms, as bonding orbitals are occupied. In this way we can understand how the renormalised atoms reduce the error. In general the lowering of the energy eigen-value of a bonding orbital compared to that of separate atomic orbitals means that the molecular orbital decays more rapidly into the space away from the atoms. The renormalised atoms are designed to mimic this effect.

Conclusions

I have made self-consistent calculations and compared the results with the Harris and HKS functionals for the cases of an aluminium surface and vacancy. The charge density constructed by overlapping free atomic charge densities leads to rather good total energies with the approximate functionals. However, the errors in the surface energy and vacancy formation energy with reference to self-consistent calculations are unacceptably large with both functionals, and greater in the case of the HKS functional. Explicit expressions have been derived for the errors in these functionals which were evaluated for the test case of a three layer slab. It was found that the errors are associated with the electrostatic energy of the longest-wavelength Fourier components of the errors in the charge density. The HKS energy is less accurate because its error involves two factors of n^{out} compared to one in the case of the Harris functional. The error in n^{out} is especially important at long wavelength (small g-vector) since it is a response to the Coulomb potential g^{-2} of the error in the initial guessed charge density. These long-wavelength components are absent in bulk crystal calculations.

The errors can be made negligible if we multiply the free atomic charge density by a function which decreases the size of the tail. The charge density must then be rescaled to conserve the total charge. A scaled Fermi function works very well for this purpose. The energy as a function of surface relaxation is well reproduced by the renormalised atoms and the Harris functional (figure 2). I initially thought that these renormalised atoms which seem to describe correctly the long-wavelength components of the charge

normal to the (111) slab would be very crystallography specific and perhaps of no use in other situations. However, the same renormalised atoms and the Harris functional also reproduce the result of a self-consistent calculation of the vacancy formation energy with an error of 0.01 eV. Even for the case of a dimer, the error is more than halved by using the renormalised atoms. It seems that if the renormalised atoms are optimised for a given defect, in the present case the unrelaxed (111) surface, then the variational principle is enough to make them work for other defects or atomic configurations.

The conclusion is that these or similar renormalised atoms can be useful in the calculation of defect energies in other metallic systems, and especially in systems in general for which pseudopotentials and plane-wave basis sets are inappropriate, for example in transition metals, where self-consistent calculations may still be too time consuming compared to non-self-consistent calculations. The rule that the Harris functional is bounded above by the self-consistent energy suggests that ehe best shape for the renormalised atoms might be found by maximising the Harris functional with respect to some shape parameters (such as the β and r_c used to scale the free atomic charge) for a given set of atomic coordinates. One could then use them to explore the total energy as a function of atomic positions.

References

Bachelet G B, Hamann D R and Schlüter M 1982 *Phys. Rev.* B **26** 4199
Batra I P, Ciraci S, Srivastava, G P, Nelson J S and Fong C Y 1986 *Phys. Rev.* B **34** 8246
Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
Ceperley D M and Alder B J 1980 *Phys. Rev. Lett.* **45** 566
Chadi D J and Cohen M L 1973 *Phys. Rev.* B **8** 5747
Chelikowski J R and Louie S G 1984 *Phys. Rev.* B **29** 3470
Feibelman P J 1983 *Phys. Rev.* B **27** 1991
Foulkes W M C 1987 *PhD Thesis* University of Cambridge
Foulkes W M C and Haydock R 1989 *Phys. Rev.* B **39** 12520
Gillan M J 1989 *J. Phys: Condens. Matter* **1** 689
Harris J 1985 *Phys. Rev.* B **31** 1770
Kleinman L and Bylander D M 1982 *Phys. Rev. Lett.* **48** 1425
Payne M C, Joannopoulos J D, Allan D C, Teter M P and Vanderbilt, D H 1989a *Phys. Rev. Lett.* **56** 2656
Payne M C, Roberts N, Needs R J, Needels M and Joannopoulos J D 1989b *Surf. Sci.* **211–212** 1
Perdew J P and Zunger A 1981 *Phys. Rev.* B **23** 5048
Polatoglou H M and Methfessel M 1988 *Phys. Rev.* B **37** 10403
Preusser A 1989 *ACM Trans. Math. Software* **15** 79
Sutton A P, Finnis M W, Pettifor D G and Ohta Y J 1988 *J. Phys. C: Solid State Phys.* **21** 35