Finite basis set corrections to total energy pseudopotential calculations

J. Phys.: Condens. Matter **2** (1990) 4395–4404. Printed in the UK

Finite basis set corrections to total energy pseudopotential calculations

G P Francis and M C Payne
Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

Received 28 December 1989, in final form 1 March 1990

Abstract. A means of correcting total energy pseudopotential calculations performed using a fixed cut-off energy for the plane waves in the basis set is presented. The use of a finite set of special k-points in such a calculation will introduce errors in the total energies which decrease only slowly with increasing cut-off energy. In particular, total energy differences are not accurate unless the cut-off energy used is sufficiently large that the total energies are themselves converged. This would not be the case if a truly constant cut-off energy could be used. Unfortunately this can only be achieved by using an infinite k-point set. We have derived a correction which will explicitly eliminate these errors to give total energies which can correspond to a strictly constant cut-off energy. In this way, total energy differences and hence many physical properties can be accurately calculated using cut-off energies significantly lower than otherwise possible, with substantial savings in computational time.

1. Introduction

Total energy pseudopotential calculations can be used to determine a wide variety of physical properties of materials. Calculations are performed on periodic supercells thereby allowing the electronic wavefunctions to be expanded in terms of a discrete set of plane waves at each of an infinite set of k-points in the Brillouin zone. This in turn allows the application of the following two approximations. Firstly, a small number of carefully chosen k-points can be used to accurately represent the wavefunction at all k-points (Chadi and Cohen 1973, Monkhorst and Pack 1976), and secondly, by truncating the basis set the wavefunctions at each k-point can be expanded in terms of a finite basis set. In principle by increasing the number of k-points and the size of the basis set it is possible to achieve absolute energy convergence. However, even in the case of very small systems, this proves to be extremely computationally expensive. In order to perform calculations on larger, more complex systems it is necessary to be able to use smaller plane wave basis sets at each k-point without reducing the accuracy of the calculation.

It is known that differences in the total energies of systems of the same size can be accurately calculated for numbers of plane waves and of k-points very much smaller than those required to ensure convergence of the absolute energies provided that identical basis sets are used for each calculation (Cheng *et al* 1988). However, when computing energy differences between systems of varying size it is impossible to use identical plane wave basis sets unless an infinite number of k-points are used in the calculation. We must choose instead either to use a constant number of plane waves in

0953-8984/90/194395 + 10 $03.50

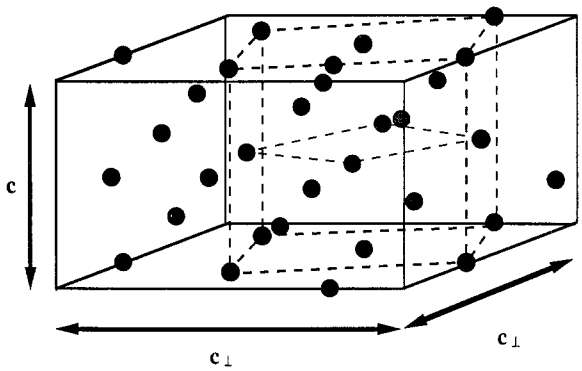

Figure 1. The 20-atom unit cell of germanium used in the calculation. The darker atoms and dashed lines define the conventional FCC unit cell when lattice parameters c and $c_\perp$ take the experimental equilibrium values.

the basis set or a constant kinetic energy cut-off for the plane wave basis set. The latter is considered to be the more physical and to have better convergence properties (Dacosta *et al* 1986). However when used with a finite number of k-points this truncation of the basis set introduces discontinuities in total energies whenever the number of plane waves used changes abruptly. The purpose of this paper is to illustrate how these systematic discontinuities and associated errors in the total energy can be dealt with straightforwardly. The technique to be described here is primarily one of correcting total energies for the errors due to the use of finite k-point sets in finite cut-off energy calculations. It will be shown that physical properties can then be determined from calculations employing severely reduced basis sets, rather than having to make basis sets sufficiently large that the discontinuities are effectively eliminated. Previous attempts to reduce these errors have been made by Yin (1985) and by Denteneer and van Haeringen (1986) but these, in the light of the present work, represent somewhat unsatisfactory approaches. This work uses perfect bulk germanium as an illustrative example. A unit cell containing twenty atoms will be used for the calculations. We consider the total energy as a function of one lattice parameter so that a total energy *curve* can be explored, allowing comparison of equilibrium lattice constants and elastic moduli with experimental values.

In sections 2 and 3 we describe the detailed effects of truncation of the plane wave basis set on stress and total energy calculations respectively. We proceed in section 4 to illustrate how a quantitative knowledge of these effects leads to improved convergence of calculated physical properties with increasing cut-off energy. Finally, we present our conclusions in section 5.

2. Stress calculations

The twenty-atom cell of Ge used for the calculations is shown in figure 1. We have simultaneously computed the stress $\sigma(c)$ and the total energy per unit cell $E_{\text{tot}}(c)$ as we vary the single lattice parameter c in the [100] crystallographic direction while keeping the other two lattice parameters ($c_\perp$) fixed. The calculations were performed using an *ab initio*, norm-conserving local pseudopotential (Starkloff and Joannopoulos 1977) and the local density approximation with the density-functional theory. The Perdew and

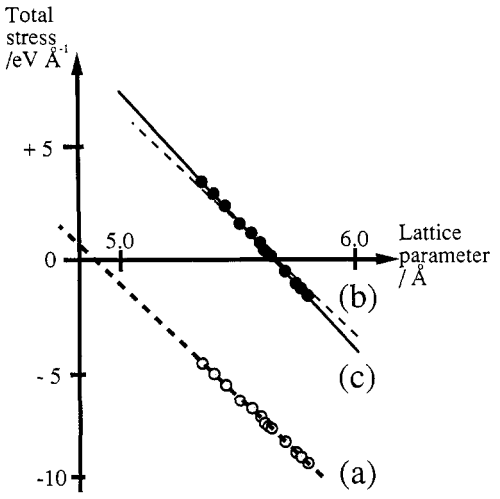

Figure 2. The quantum-mechanical stress determined as a function of lattice parameter, using $E_{max} = 125$ eV and 14 equally weighted k-points. Curve (a) shows the calculated (uncorrected) stress. Curve (b) shows the stress corrected by adding a constant Pulay stress. Curve (c) shows the stress corrected by adding a $1/c$ Pulay stress.

Zunger (1981) parametrisation of the Ceperley and Alder (1980) form of the exchange–correlation potential was employed. The value of $c_\perp$ was first determined as that corresponding to the total energy minimum under isotropic volume strain. Figure 2 shows the calculated stress as a function of lattice parameter.

The calculation of the quantum-mechanical stress at each value of the lattice constant was performed using a locally constant number of plane waves per k-point, corresponding to a fixed cut-off energy $E_{max} = 125$ eV and 14 equally weighted k-points. The plane wave $|\boldsymbol{k} + \boldsymbol{G}\rangle$ will be included in the basis set for k-point $\boldsymbol{k}$ if

$$(\hbar^2/2m_e)\,|\boldsymbol{k} + \boldsymbol{G}|^2 \leq E_{max}. \tag{1}$$

The number of such states included in the basis set for k-point $\boldsymbol{k}$ will be denoted $N_k^{PW}(c, E_{max})$. If k-point $\boldsymbol{k}$ carries a weighting f_k then we will define a weighted number of basis states as

$$N^{PW}(c) = \prod_k (N_k^{PW}(c))^{f_k} \qquad \sum_k f_k = 1. \tag{2}$$

Elementary considerations of the average density of states in $\boldsymbol{G}$-space give the approximate relation

$$N^{PW}(c) = g_\eta c \tag{3}$$

where

$$g_\eta = (c_\perp^2/6\pi^2)(2m_e/\hbar^2)^{3/2} E_{max}^{3/2}. \tag{4}$$

However, this expression takes no account of the discrete nature of N^{PW} inherent in a finite k-point set calculation. Figure 3 shows the fluctuations in the actual $N^{PW}(c)$ curve for the 20-atom unit cell using $E_{max} = 125$ eV and 14 equally weighted k-points about the continuum limit expression (3). While the number of plane waves $N^{PW}(c, E_{max})$ is well-defined for given E_{max} via equations (1) and (2), the discrete nature of $N^{PW}(c)$ means that E_{max} is not well-defined for given N^{PW}. The total energy is assumed to be a differentiable function of lattice parameter and cut-off energy in a calculation employing a complete k-point set. In the case of a finite k-point set we must define a strictly constant well-defined cut-off energy $E_{max}(c, N^{PW})$ according to equations (3) and (4). We can then distinguish the calculated stress $\sigma(c)$ corresponding to locally constant N^{PW} and the corrected stress corresponding to strictly constant E_{max}.

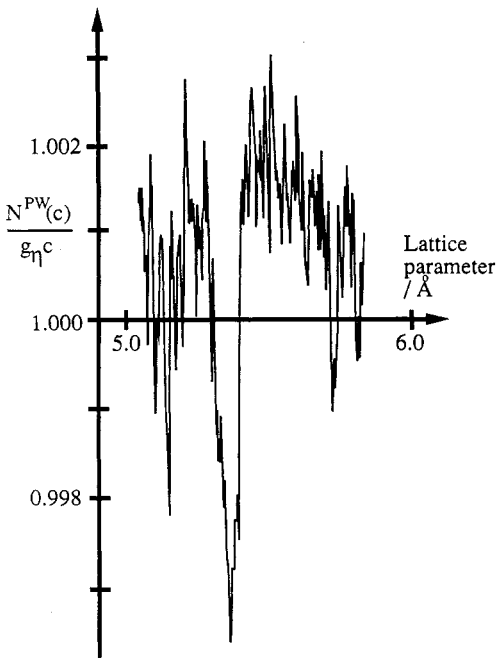

Figure 3. The ratio of the actual number of plane waves in the basis set to the corresponding number in the continuum limit, using $E_{max} = 125$ eV and 14 equally weighted k-points, determined as a function of the lattice parameter. The fluctuations in the plotted ratio around unity are only of the order of 0.2% but this leads to significant errors in the calculated total energies.

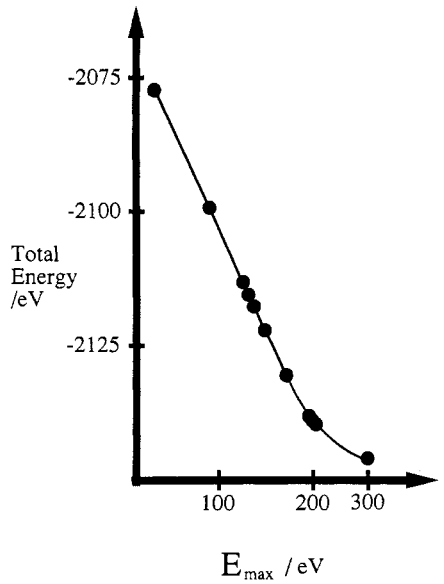

Figure 4. The convergence of total energy with increasing cut-off energy E_{max} for a value of lattice parameter close to the equilibrium value.

We will define the difference between these stresses as the 'Pulay stress' $\sigma_{Pulay}(c)$, by analogy with the Pulay force (Pulay 1969). It has been shown (Froyen and Cohen 1986, Vanderbilt 1987) that

$$\sigma_{Pulay}(c, E_{max}) = -(\partial E_{tot}(c, E_{max})/\partial c)_{E_{max}} + (\partial E_{tot}(c, N^{PW}(c, E_{max}))/\partial c)_{N^{PW}} \quad (5)$$

$$= (\partial E_{max}(N^{PW}, c)/\partial c)_{N^{PW}} (\partial E_{tot}(c, E_{max})/\partial E_{max}))_c$$

$$= -(2E_{max}/3c)(\partial E_{tot}(c, E_{max})/\partial E_{max})_c$$

$$\simeq -(2E_{max}/3c)(\partial E_{tot}(c, N^{PW}(c, E_{max}))/\partial E_{max})_c \quad (6)$$

where the final approximation is valid to within the other uncertainties of the calculation. In this form the value of the Pulay stress can be calculated explicitly for all values of c by calculating the change of E_{tot} with increasing ln E_{max}. Figure 4 shows a logarithmic plot of E_{tot} as a function of E_{max} for a typical lattice parameter. Since the total energy is a monotonically decreasing function of cut-off energy the Pulay stress is a positive tensile stress. σ_{Pulay} will only vanish when E_{max} is sufficiently large that absolute total energy convergence is achieved. However the c dependence of $\partial E_{tot}/\partial$ ln E_{max} diminishes with increasing E_{max}, with the Pulay stress becoming inversely proportional to c when total energy *differences* have converged. With $E_{max} = 125$ eV calculated total energy differences have not converged, but $(c\sigma_{Pulay}(c))$ is nevertheless constant to within about

Table 1. The comparison of equilibrium properties determined from the behaviour of the quantum-mechanical stress, as a function of the lattice parameter c, with the experimental values. The stresses were calculated using $E_{max} = 125$ eV and 14 equally weighted k-points, and were corrected by including a Pulay term. The experimental data is from *Landolt–Börstein* (1982).

Method	Equilibrium lattice constant (Å)	Equilibrium Young modulus (GPa)
Experiment	5.65	131
Calculated stress	4.89 ± 0.01	99 ± 2
$1/c$ corrected stress	5.56 ± 0.01	127 ± 2
Constant corrected stress	5.67 ± 0.01	112 ± 2

1%. Figure 2 shows the stress obtained by correcting the calculated stress with the $1/c$ Pulay term of equation (6). Table 1 compares the equilibrium properties thus obtained with their experimental values. It is clear that the inclusion of this correction yields physical properties much closer to the experimental values than is otherwise possible at this cut-off energy. In order to emphasise the need to use the full c-dependence of the Pulay stress if accurate physical properties are to be obtained, figure 2 and table 1 include the results of correcting the calculated stress with a constant Pulay stress corresponding to a single evaluation of equation (6) at the equilibrium lattice parameter.

3. Total energy calculations

The simultaneous calculation of the total energy $E_{tot}(c, N^{PW}(c, E_{max}))$ provides an independent determination of the equilibrium properties. Figure 5 shows the calculated total energy curve using a fixed cut-off energy $E_{max} = 125$ eV and 14 equally weighted k-points. The following two features of the figure, which become more striking at lower cut-off energies (Dacosta *et al* 1986), merit particular attention.

Firstly, while one would expect a real material to possess a smooth total energy curve as a function of lattice parameter, this curve is very ragged. The reason for this is qualitatively well understood (Dacosta *et al* 1986). Whenever the number of basis states increases discontinuously, as it must if a finite k-point set is used, the total energy will fall abruptly due to the increased variational freedom of the wavefunction. The second important feature of figure 5 is that in ranges of c where the number of plane waves is constant, $E_{tot}(c, N^{PW})$ exhibits a locally increased gradient, corresponding to the Pulay stress defined in section 2. We wish to correct the total energy curve by eliminating the errors introduced by these two features.

At each k-point $\boldsymbol{k}$ we must relate the calculated energy $E^{\boldsymbol{k}}_{tot}(c, N^{PW}_{\boldsymbol{k}}(c, E_{max}))$ to the required energy $E^{\boldsymbol{k}}_{tot}(c, E_{max})$, assumed differentiable, where $N^{PW}_{\boldsymbol{k}}(c, E_{max})$ is defined according to equation (1). The total differential of the calculated energy is

$$\mathrm{d}E^{\boldsymbol{k}}_{tot}(c, N^{PW}_{\boldsymbol{k}}(c, E_{max})) = (\partial E^{\boldsymbol{k}}_{tot}/\partial c)_{N^{PW}_{\boldsymbol{k}}}\,\mathrm{d}c + (\partial E^{\boldsymbol{k}}_{tot}/\partial N^{PW}_{\boldsymbol{k}})_c\,\mathrm{d}N^{PW}_{\boldsymbol{k}} \tag{7}$$

and we have also that

$$(\partial E^{\boldsymbol{k}}_{tot}(c, N^{PW}_{\boldsymbol{k}}/\partial N^{PW}_{\boldsymbol{k}})_c\,\mathrm{d}N^{PW}_{\boldsymbol{k}} = (\partial E^{\boldsymbol{k}}_{tot}(c, E_{max})/\partial E_{max})_c\,(\partial E_{max}/\partial N^{PW}_{\boldsymbol{k}})_c\,\mathrm{d}N^{PW}_{\boldsymbol{k}} \tag{8}$$

where $E_{max}(c, N^{PW})$ is defined by equations (3) and (4). We can now integrate equation (7) by using equations (5), (6) and (8) to give

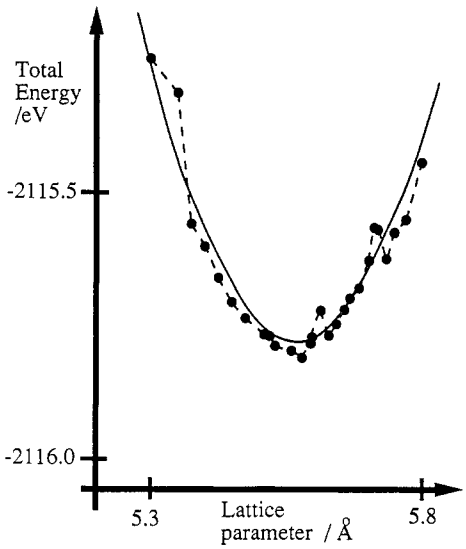

Figure 5. The calculated (uncorrected) total energy determined as a function of lattice parameter, using a constant cut-off energy E_{max} = 125 eV and 14 equally weighted k-points. The full curve represents the least-squares parabolic fit to the data.

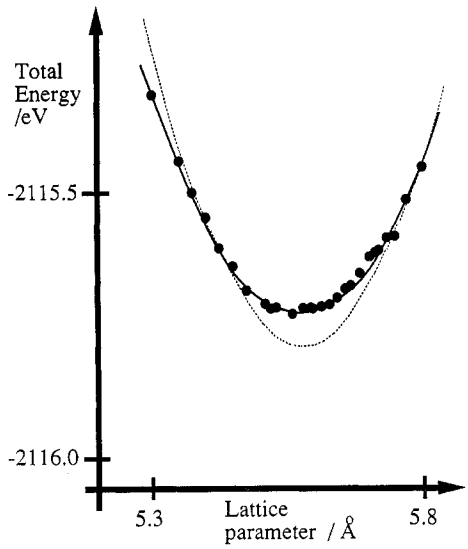

Figure 6. The corrected total energy determined as a function of lattice parameter, using a constant cut-off energy E_{max} = 125 eV and 14 equally weighted k-points, shown with its least-squares parabolic fit. The uncorrected best fit curve (broken curve) is included for comparison.

$$E^k_{tot}(c, N^{PW}_k(c, E_{max})) = \int^c \left(\frac{\partial E^k_{tot}(c', E_{max})}{\partial c'}\right)_{E_{max}} dc' + \frac{2E_{max}}{3}\left(\frac{\partial E^k_{tot}(c, E_{max})}{\partial E_{max}}\right)_c \ln\left(\frac{N^{PW}_k(c, E_{max})}{g_\eta c}\right) \tag{9}$$

where we have assumed only that $\partial E^k_{tot}/\partial E_{max}$ is independent of c. This approximation is easily checked and found to be valid to within about 1% at 125 eV over a 10% range of lattice parameter.

Summing over weighted special k-points we finally find

$$E_{tot}(c, E_{max}) = \sum_k f_k E^k_{tot}(c, E_{max}) = E_{tot}(c, N^{PW}(c, E_{max})) - \frac{2E_{max}}{3}\left(\frac{\partial E_{tot}(c, E_{max})}{\partial E_{max}}\right)_c \ln\left(\frac{N^{PW}(c, E_{max})}{g_\eta c}\right) \tag{10}$$

where we have further assumed that $\partial E^k_{tot}/\partial E_{max}$ is also independent of $\boldsymbol{k}$. Clearly the closer we are to convergence of energy differences the better will be our approximation. At E_{max} = 125 eV the assumed $\boldsymbol{k}$-independence is found to be valid to within a few percent. When we are close to convergence of energy differences, the derivative in (10) will be very nearly independent of c and so can be determined for all values of c from

Table 2. The comparison of equilibrium properties determined from the behaviour of the total energy, as a function of the lattice parameter c, with the experimental values. The total energies were calculated using $E_{max} = 125$ eV and 14 equally weighted k-points, and were corrected according to equation (10).

Method	Equilibrium lattice constant (Å)	Equilibrium Young modulus (GPa)
Experiment	5.65	131
Calculated energy	5.57 ± 0.01	160 ± 10
Corrected energy	5.57 ± 0.01	129 ± 2

calculations of $E_{tot}(E_{max})$ at one or two values of c. Alternatively, when a total energy curve is available, the derivative may be treated as a fitting parameter and estimated by optimising the smoothness of the corrected energy curve.

A test of the validity of these assumptions at any cut-off energy is whether the parameter-free expression (10) yields smooth calculated $E_{tot}(c, E_{max})$ curves. Figure 6 shows the curve obtained for $E_{max} = 125$ eV. The curve is indeed remarkably smooth to within the other uncertainties of the calculation, and moreover yields a value of the elastic constant which is within 2% of the experimental value as shown in table 2. In contrast, the corresponding uncorrected $E_{tot}(c, N^{PW}(c, E_{max}))$ curve yields an elastic constant some 25% from the experimental value. That the curves in figure 6 differ significantly clearly shows that the earlier schemes of Yin (1985) and Denteneer and van Haeringen (1986) are unsatisfactory, since they rely on the 'noise' in the uncorrected curve having a random nature.

A further test of the validity of the assumptions is provided by a comparison of properties calculated from the corrected energy curve with values derived from the corresponding corrected stress curve. Comparison of such results in tables 1 and 2 confirms their validity.

4. Convergence properties

We have considered the convergence of total energy E_{tot} with increasing plane-wave cut-off energy E_{max}. In most cases we are interested in the energy differences between different structures, rather than their absolute energies. The convergence of energy differences occurs at a much smaller cut-off energy than that required for absolute convergence. This is understood to be due to the higher plane waves changing the electronic wavefunctions only within the ion cores and not in the bonding regions. Once the cut-off energy is sufficiently large that this regime is reached, further increasing E_{max} will merely decrease the total energy of each structure by precisely the same amount.

The important point here is that even when the cut-off energy is sufficiently large that energy differences in $E_{tot}(c, E_{max})$ have converged, with $\partial E_{tot}(c, E_{max})/\partial \ln E_{max}$ therefore independent of c, energy differences in the calculated $E_{tot}(c, N^{PW}(c, E_{max}))$ will not have converged since the correction term in equation (10) may still be large. The logarithm will always fluctuate from zero, albeit with slowly decreasing magnitude as E_{max} increases, but the constant prefactor will only vanish when absolute energy convergence is achieved. Hence unless the correction is applied, calculated energy differences will not converge until the calculated energies themselves have converged!

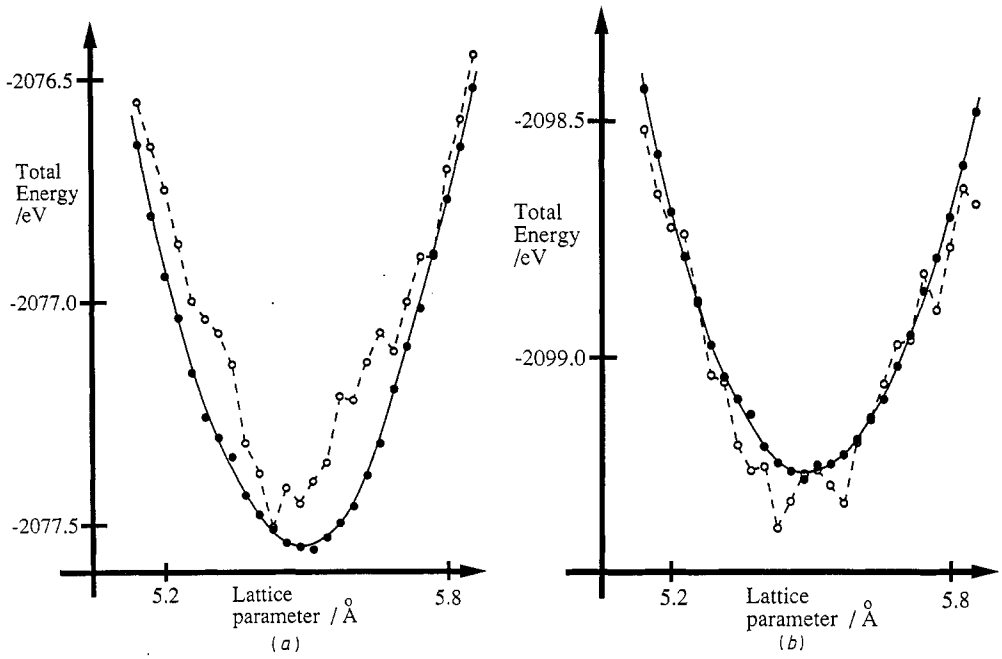

Figure 7. The calculated and corrected total energies determined as functions of lattice parameter using a constant cut-off energy and 14 equally weighted k-points. Figures 7(a) and 7(b) show the curves for cut-off energies of 62.5 eV and 93.75 eV respectively. In both cases, the filled data points represent the corrected values while the empty ones are the original calculated values and the full curves are the best fit parabolae.

In this way, the use of the correction significantly reduces the number of plane waves required for calculating total energy differences.

The validity of equation (10) is measured by the smoothness of the corrected energy curve which it predicts. In our example, $E_{\max} = 125$ eV is adequate for the approximation to be valid to within the other small uncertainties of the calculation. Figure 7 shows the corrected curves using cut-off energies of 62.5 eV and 93.75 eV. The validity of (10) clearly worsens as the cut-off is reduced, but nevertheless the 'noise' in the corrected energy remains an order of magnitude less than the 'systematic noise' in the uncorrected energy. In any case it is clear that equation (10) becomes valid at much lower cut-off energies than those required for convergence of energy differences.

It should be noted that while equation (10) corrects the errors due to the use of a finite k-point set, within the Monkhorst and Pack (1976) special k-point summation approximation to the Brillouin zone integration, it is necessary to perform the calculation with a k-point set sufficient for the assumptions inherent in equation (10) to be valid. As for cut-off energy, this validity may be checked for a given k-point set by determining the smoothness of the corrected total energy curve so predicted. In addition, it remains necessary to test that the k-point set is sufficient for the special k-point scheme to be a valid approximation.

Other quantities such as elastic moduli will be less sensitive to small changes in energy differences, so that convergence of such properties may be achieved at very much reduced cut-off energies but only if the correction is applied. In our example good values of elastic moduli can be achieved with a reduction in computational time of one order of magnitude. Table 3 compares the convergence of an elastic modulus determined from corrected and uncorrected energies respectively.

Table 3. The convergence of equilibrium properties with increasing cut-off energy E_{max}. Note that (i) the consistency of the stress and energy corrections improves with increasing cut-off energy, and that (ii) the corrected properties may take values outside the already large uncertainties in the corresponding uncorrected properties. The number of lattice parameters used for the stress and energy evaluations at each cut-off energy is shown in the 'number of data points' column.

Cut-off energy (eV)	Number of data points	Equilibrium Young modulus (GPa)			Equilibrium lattice constant (Å)		
		Uncorrected energy	Stress (corrected)	Energy (corrected)	Uncorrected energy	Stress (corrected)	Energy (corrected)
						(±0.01)	
62.5	27	157 ± 8	161 ± 1	170 ± 2	5.49	5.48	5.48
93.75	26	130 ± 7	143 ± 1	138 ± 2	5.50	5.49	5.50
125	27	160 ± 10	127 ± 2	129 ± 2	5.57	5.56	5.57
130	6	127 ± 9	126 ± 2	122 ± 2	5.61	5.58	5.60
200	7	146 ± 15	133 ± 2	136 ± 6	5.62	5.67	5.63
Experiment			131			5.65	

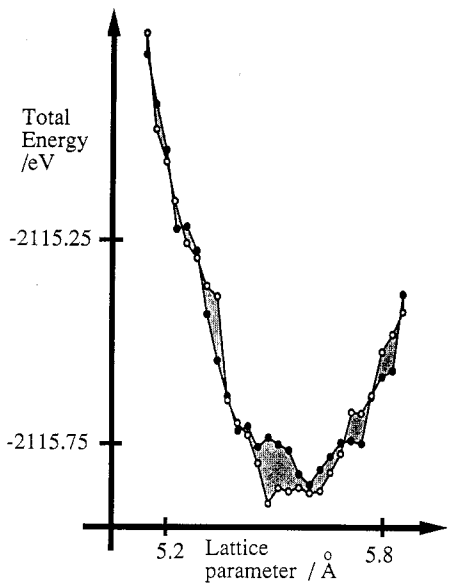

Figure 8. The predicted total energies for cut-off energies of 125 eV (full circles) and 126 eV (empty circles) determined as functions of lattice parameter. The latter curve has been shifted slightly in energy such that the difference between the curves is due entirely to the change in the finite k-point set effect resulting from a change in the cut-off energy. The shaded region emphasises the varying energy differences of the two curves. The elastic moduli calculated from the two curves differ by about 15%.

To emphasise the magnitude of the correction, we have calculated the total energies using cut-off energies of 125 eV and 126 eV respectively, assuming that their corrected energies are identical, in order that the difference between the curves is purely due to the now correctable errors resulting from the finite k-point sets used and not due directly to the finiteness of the cut-off energy. Figure 8 shows these curves and thereby demonstrates the critical sensitivity of uncorrected energy differences and elastic moduli to the precise value of the cut-off energy used.

5. Conclusions

We have shown that at finite plane wave cut-off energy E_{max}, we can explicitly correct calculated energies by eliminating the errors due to the use of the finite k-point set

(within the special k-point scheme). The corrected energy is more physical than the original total energy, but will not be equal to the true total energy of the material unless very high E_{max} is used in the calculation. However, energy differences in the corrected energy will always converge at a much smaller cut-off energy than that required for energy differences in the uncorrected energy to have converged. We can therefore perform accurate total energy calculations in order to determine total energy differences and physical properties using significantly smaller basis sets than are possible without applying the correction, allowing substantial savings in computer time.

Acknowledgments

The authors thank the Science and Engineering Research Council for a studentship (GPF) and for the use of the CRAY XMP at the Rutherford Appleton Laboratory under grant GR/E 91790. One of us (MCP) thanks the Royal Society for financial support.

References

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
Cheng C, Needs R J and Heine V 1988 *J. Phys. C: Solid State Phys.* **21** 1049
Dacosta P G, Nielsen O H and Kunc K 1986 *J. Phys. C: Solid State Phys.* **19** 3163
Denteneer P J H and van Haeringen W 1986 *Phys. Rev.* B **33** 2831
Froyen S and Cohen M L 1986 *J. Phys. C: Solid State Phys.* **19** 2623
Landolt-Börstein New Series 1982 *Numerical Data and Functional Relationships in Science and Technology* (Berlin: Springer)
Monkhorst H J and Pack J D 1976 *Phys. Rev.* B **13** 5188
Perdew P and Zunger A 1981 *Phys. Rev.* B **23** 5048
Pulay P 1969 *Mol. Phys.* **17** 197
Starkloff Th and Joannopoulos J D 1977 *Phys. Rev.* B **16** 5212
Vanderbilt D 1987 *Phys. Rev. Lett.* **59** 1456
Yin M T 1985 *Proc. 17th Int. Conf. on the Physics of Semiconductors (San Francisco, 1984)* (New York: Springer)