

The $k \cdot p$ method in pseudopotential total energy calculations: error reduction and absolute energies

I J Robertson and M C Payne

Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

Received 25 July 1990, in final form 28 February 1991

Abstract. The scope and potential of the $k \cdot p$ total energy method introduced in a previous paper is further investigated. The method provides for the generation of rapid but approximate solution of the Kohn–Sham equations at many k -points from the exact solution at a few k -points. The method is applied to three diverse aluminium structures and the errors are partitioned into those which are due to the $k \cdot p$ method and those which are general to any finite sampling method. Both of the errors that are associated with any finite sampling technique are shown to be significant even for dense sampling of k -space. One of the $k \cdot p$ errors is shown to be insignificant. The other $k \cdot p$ error is significant. However, a method is introduced which allows the magnitude of the error to be reduced to the level of insignificance. The resulting $k \cdot p$ total energy method is shown to be immune from any additional errors beyond those associated with any finite sampling method. Thus it is a quick and accurate method for the calculation of *absolute* total energies.

1. Introduction

In principle a pseudopotential total energy calculation of an infinite solid within the density functional formalism involves an infinite number of electrons, each described by a solution of a Kohn–Sham equation (Kohn and Sham 1965). If the structure is periodic then one may label the infinity of Kohn–Sham eigenstates by a k -point index which takes all values within a finite and well defined volume of reciprocal space. Properties of a state such as the eigenvalue vary continuously and relatively slowly with respect to k -point index so that an infinite number of electrons may be represented by a finite number of electrons whose k -point index forms a grid within the first Brillouin zone of the structure concerned.

In order to achieve this simplification even for aperiodic structures, one chooses to impose periodicity by choice of a suitable supercell. The precise number of k -points required for a given degree of convergence will vary from structure to structure. Semiconductors and insulators may be adequately treated using a relatively small number of suitably chosen special k -points (Chadi and Cohen 1973), however, calculations on metals show that for these structures, significantly larger sets of k -points are required.

In calculating energy differences between similar structures, errors due to insufficiently dense sampling within the Brillouin zone tend to cancel (Cheng *et al* 1987), but in comparing dissimilar structures e.g. a surface and bulk, the same degree of cancellation will not be present. For both these reasons there is often a need to sample at a large number of k -points. However this is both time consuming and extremely demanding on

computer memory with the result that adequately converged results are often unobtainable.

In a previous paper (Robertson and Payne 1990), hereafter referred to as 1, we described a method which used exact solutions of the Kohn–Sham equations at a small number of k -points as a basis to generate quick and approximate solutions at a much larger number of other k -points. (Here and throughout the rest of the paper we use the word ‘exact’ to describe solutions of a given Hamiltonian in a full space of plane waves that are reciprocal lattice vectors of the structure and that have energy below a given cut-off. The word ‘inexact’ will refer to solutions obtained with a subspace of this space as is the case for instance in the $k \cdot p$ method). The usefulness of this $k \cdot p$ method was demonstrated by performing calculations on FCC aluminium using a relatively small cut-off energy of 100 eV. The procedure was shown to be up to several orders of magnitude quicker than an exact solution, and subject to an extra error of only 2% in the calculation of certain energy differences. The figure of 2% is relatively small, reflecting the promise of the $k \cdot p$ method, but was helped by a large degree of error cancellation between similar structures rather than being a truly low error in an *absolute* energy.

This paper has three major purposes:

- (i) We show how to reduce any extra error in calculating eigenvalues by the $k \cdot p$ method by up to two orders of magnitude so that they are reduced to the level of insignificance. This ensures that it is no longer necessary to rely on the error cancellations which are only present when comparing similar structures.
- (ii) We show the general applicability of the $k \cdot p$ method by using it to calculate energies of aluminium in three diverse structures. These are the FCC structure, a line of atoms and a square lattice.
- (iii) We demonstrate that the effect of any error in the electronic charge density and potential due to the use of $k \cdot p$ eigenstates is insignificant.

The combined effect of these three advances is to show that for a diverse range of structures, the $k \cdot p$ method can generate total energies using a given grid of k -points with the introduction of no significant additional errors and at a fraction of the computational cost of ‘exact methods’. This transformation to a quick and accurate method allows one to use very large k -point sets, effectively eliminating those errors due to inadequate sampling of the Brillouin zone thus opening the path to calculating truly absolute energies.

The layout of the paper is as follows. In section 2 we describe four errors that may arise in a total energy calculation. These are errors in evaluating the charge density and in evaluating the eigenvalue sum, and may arise due to either finite sampling of the Brillouin zone or due to the use of the $k \cdot p$ method. In sections 3 to 5 we evaluate three of these errors for the structures that we have chosen to study. In section 6 we evaluate the most serious error of the four and show how the magnitude of the error may be reduced by up to two orders of magnitude. In section 7 we present our conclusions.

The three structures that we will be concerned with are all formed from pure aluminium. In each case the nearest-neighbour atomic separation is 2.85 Å. The first structure is the face centred cubic structure with a four atom cubic unit cell of lattice parameter 4.0305 Å. The second structure is a line of atoms. The unit cell is primitive and tetragonal with dimensions $5.7 \times 5.7 \times 2.85$ Å. The final structure is a square lattice. Again the unit cell is primitive and tetragonal with dimensions $5.7 \times 2.85 \times 2.85$ Å. The pseudopotential used to represent the aluminium ions is that of Goodwin *et al* (1990) and we use the exchange and correlation energy function of Ceperley and Alder (1980)

as parametrized by Perdew and Zunger (1981). Unless stated otherwise, the cut-off energy used in the exact calculations is 190 eV. In the $k \cdot p$ calculations the 19 lowest solutions at either $k = 0$ or $k = (1/2a, 1/2b, 1/2c)\pi$ are used as a basis, where a , b and c are the dimensions of the particular unit cell. The k -point grids used throughout the paper are those of Monkhorst and Pack (1976) where the notation 6^3 refers to a uniform division of the first Brillouin zone into a $6 \times 6 \times 6$ mesh of k -points. We generate the density of states and hence the Fermi level and the occupancies of the eigenstates using the Gaussian broadening technique developed by Fu and Ho (1983).

2. Definition of errors relevant to the $k \cdot p$ total energy method

In performing a pseudopotential total energy calculation, errors may arise from a multitude of sources: from the use of an incomplete basis set for expanding the electronic wavefunctions; from local density approximations to exchange and correlation or from an inadequate pseudopotential. It is not these errors that we are concerned with in this work. The errors that we are concerned with are those which arise from the consideration of too few k -points in the Brillouin zone (finite sampling errors) and those which arise from the use of the $k \cdot p$ approximation to generate eigenstates and eigenvalues ($k \cdot p$ errors).

Throughout this paper, as in our previous paper, we will estimate the total energy of a structure by use of the Harris energy functional (Harris 1985), defined by

$$H = \sum_i \varepsilon_i w_i - E_H[n_{\text{in}}(r)] + E_{\text{xc}}[n_{\text{in}}(r)] - \int d^3r \mu_{\text{xc}}[n_{\text{in}}(r)] n_{\text{in}}(r) + E_{\text{ion-ion}} \quad (1)$$

where E_H is the Hartree energy, E_{xc} is the exchange-correlation energy, $E_{\text{ion-ion}}$ is the ion-ion interaction, μ_{xc} is the exchange correlation potential, $n_{\text{in}}(r)$ is the input charge density, ε_i is the i th eigenvalue and w_i is the corresponding occupation probability. Errors in the Harris expression for the energy are second order with respect to deviations in the input charge density from the self-consistent density.

Even if the Harris expression for the energy is evaluated using a self-consistent charge density, it may still be in error because either the self-consistent charge density used in equation (1) is wrong or because, in spite of using the correct charge density, there is an error in the eigenvalue sum. The error in the charge density may arise because too few k -points are used to generate a self-consistent charge density; this will be called a "finite sampling error in charge density"; or it may arise because the eigenstates that are used to generate the charge density are approximate states generated by $k \cdot p$ theory; this will be called a " $k \cdot p$ error in charge density". The error in the eigenvalue sum may also arise from two sources. It may either be due to the fact that the eigenvalues are calculated at too few k -points, "a finite sampling error in eigenvalue sum", or it may be due to the fact that the eigenvalues used are not exact but those obtained by the $k \cdot p$ method, i.e. "a $k \cdot p$ error in the eigenvalue sum".

More formally one should define the following.

(i) *The exact sum.* For a given input potential, a given cut-off energy and a given grid of k -points, the exact sum is defined as the eigenvalue sum obtained by solving at each k -point using the full basis of allowed plane waves with energies below the energy cut-off. The solutions will not be self-consistent in general.

(ii) *The $k \cdot p$ sum.* For a given input potential, a given cut-off energy and a given grid of k -points, the $k \cdot p$ sum is defined as the eigenvalue sum obtained by solving at each k -point using the $k \cdot p$ method rather than the exact solution referred to above.

(iii) *The finite sampling error in eigenvalue sum.* For a given potential, k -point grid and cut-off energy, this error is defined as the difference between the exact sum for that k -point grid and the exact sum in the case of an infinitely fine k -point grid.

(iv) *The $k \cdot p$ error eigenvalue sum.* For a given potential, k -point grid and cut-off energy, this is defined as the difference between the exact sum for that k -point grid and the $k \cdot p$ sum for that k -point grid.

(v) *The $k \cdot p$ error in charge density.* In general, the number of k -points used to generate the charge density (N_1) and the number of k -points used to generate the eigenvalue sum (N_2) need not be equal. It is also possible to generate either, neither or both the N_1 and N_2 solutions exactly or by the $k \cdot p$ method. The $k \cdot p$ error in charge density is the error which arises because the N_1 solutions used to generate the charge density are generated by the $k \cdot p$ method. Formally, for a given value N_2 of k -points solved exactly to generate the eigenvalue sum, it is the error introduced by generating the self-consistent charge density by N_1 $k \cdot p$ eigenstates rather than by N_1 exact eigenstates.

(vi) *The finite sampling error in charge density.* The finite sampling error in the charge density is the error that arises because the charge density is generated by too few (N_1) exact eigenstates. More formally, for a given number of (N_2) exact solutions used to generate the eigenvalue sum, it is the error introduced by using a finite number (N_1) rather than an infinite number of exact solutions in order to generate the charge density.

The two finite sampling errors arise purely due to coarse sampling of the Brillouin zone. They are the errors that are general to total energy pseudopotential calculations and they can be systematically reduced by finer sampling in k -space. The two $k \cdot p$ errors are new errors that the $k \cdot p$ method has introduced. Computationally, the $k \cdot p$ method is several orders of magnitude faster than previous methods which allows one to use much larger k -point sets than would otherwise be possible. It consequently allows finite sampling errors to be reduced. It is important to establish that the $k \cdot p$ errors can be made smaller than the finite sampling errors or the use of such a scheme will not reduce the magnitude of the error in the total energy. In the next four sections we evaluate each of these four errors. We show that if uncorrected the largest of these errors is generally the $k \cdot p$ error in the eigenvalue sum. Without correction this error would limit the scope of $k \cdot p$ calculations. However, we show how one may reduce this error significantly with the result that for k -point grids of interest, the $k \cdot p$ errors become negligible in comparison with the errors due to finite k -point sampling. This allows accurate *absolute* total energies to be calculated using the $k \cdot p$ method.

3. $k \cdot p$ errors in the charge density

FCC aluminium is rather free-electron like. Hence, the Kohn-Sham eigenstates at different k -points generate very similar charge densities. Any problems that are associated with error in the charge densities either due to finite sampling or due to $k \cdot p$ errors are far more likely to show up in a less free-electron-like structure. It is for this reason that in the next two sections we concentrate on the linear structure and the square lattice.

In order to evaluate the $k \cdot p$ error in the charge density we perform two calculations on each of the two structures. First we perform an exact Car-Parrinello (1985) calculation

Table 1. The $k \cdot p$ eigenvalue sum of a 2^3 grid of k -points and the corresponding total energy of the line and square lattice using exact and $k \cdot p$ generated potentials.

	Eigenvalue sum for 2^3 grid (eV)	Total energy (eV)
<i>Line</i>		
Using exact potential	-18.311 92	-55.662 31
Using $k \cdot p$ potential	-18.434 19	-55.662 39
<i>Square lattice</i>		
Using exact potential	-17.886 96	-56.429 45
Using $k \cdot p$ potential	-17.931 20	-56.429 44

using the modified algorithm of Payne *et al* (1986) with a Monkhorst Pack grid of 2^3 , using an energy cut-off of 190 eV for the plane wave basis set and iterating to self-consistency. We then perform the same calculation using the 19 lowest bands at $k = 0$ to generate $k \cdot p$ solutions over the same 2^3 Monkhorst Pack grid, again iterating to self-consistency. In each case we store the final self-consistent charge density, one exact and one produced by the $k \cdot p$ method. In order to achieve a fair comparison between these charge densities it is important to treat them in an identical manner. In particular the eigenvalue sums must be calculated in the same way. The $k \cdot p$ charge density is used to generate a potential from which the 2^3 eigenvalue sum is evaluated exactly. The Harris expression in each case may then be evaluated by adding an exact eigenvalue sum to the other contributions. By this apparently tortuous route we ensure that any difference between the two values for the energy is due to differences in charge density and not to differences in the methods of generating the eigenvalue sum. The results are shown in table 1. Given the excellent agreement between the two sets of results we may confidently conclude that $k \cdot p$ error in the potential produces an error in the total energy of the order of 10^{-4} eV and is consequently negligible.

4. Finite sampling error in the charge density

For the reason outlined in the previous section, we again restrict ourselves to the two non-close-packed structures. Strictly speaking the finite sampling error in the charge density is defined as the difference between exact calculations as the number of k -points used to generate the charge density is changed. In order to save time we estimate this error as the difference between $k \cdot p$ calculations as the number of k -points used to generate the charge density is changed. In the previous section it was noted that the $k \cdot p$ error in the charge density was negligible. The $k \cdot p$ error in eigenvalue sum is primarily dependent on the number of points used to generate the eigenvalue sum, N_2 . For a fixed N_2 any $k \cdot p$ errors will largely cancel. Given these two facts we proceed as follows.

For each structure we have performed $k \cdot p$ calculations using a cut-off energy of 190 eV, using 19 bands at $k = 0$ and Monkhorst Pack grids of 1^3 , 2^3 , 4^3 , 6^3 , 8^3 , 10^3 . In each case we iterate to self-consistency in the manner referred to in our previous paper. In each case the final self-consistent charge density and potential is stored. The potentials are then used to generate the 19 lowest exact non-self-consistent solutions at $k = (1/2a, 1/2b, 1/2c)\pi$ and these are used to evaluate $k \cdot p$ solutions over a Monkhorst Pack grid of 16^3 .

Table 2. The $k \cdot p$ eigenvalue sum of a 16^3 k -point grid and the corresponding total energy for our linear structure (see text), as a function of the size of the grid of $k \cdot p$ k -points used to generate the potential.

Grid for potential	Eigenvalue sum for 16^3 grid (eV)	Energy (eV)
1^3	-18.134 28	-56.233 78
2^3	-19.038 64	-56.266 84
4^3	-19.318 46	-56.283 57
6^3	-18.766 58	-56.255 77
8^3	-18.204 38	-56.246 30
10^3	-18.101 05	-56.246 73

Table 3. The $k \cdot p$ eigenvalue sum of a 16^3 k -point grid and the corresponding total energy for our square structure (see text), as a function of the size of the grid of $k \cdot p$ k -points used to generate the potential.

Grid for potential	Eigenvalue sum for 16^3 grid (eV)	Energy (eV)
1^3	-18.188 96	-57.397 76
2^3	-19.014 78	-57.513 02
4^3	-16.677 25	-57.247 37
6^3	-16.164 88	-57.236 81
8^3	-16.403 46	-57.239 68
10^3	-16.616 08	-57.244 73

The eigenvalue sums obtained in each case can be used in equation (1) with the charge densities which produced them to generate a series of estimates of the energy. The energies are calculated using a fixed number of k -points to generate the eigenvalue sum but a variable number of k -points to generate the charge density. The variation of the energy with respect to the size of the grid used to generate the charge density gives a measure of the finite sampling errors in the charge density.

Table 2 shows the results for the line of atoms and table 3 the results for the square lattice. For the square lattice there is a sharp change in energy as the grid size is increased from 1^3 to 4^3 . After that the convergence is rather more erratic and for a 10^3 grid the finite sampling error is still of the order of 0.01 eV. For the line the finite sampling error in charge density is less but there is still evidence of significant error for k -point grids as large as 10^3 .

The conclusion of this and the previous section is that use of the $k \cdot p$ method is fully justified for generating charge densities. Any extra errors introduced by the $k \cdot p$ method in charge density are negligible compared to the finite sampling errors that the method is attempting to overcome.

5. Finite sampling error in the eigenvalue sum

This error was considered in paper 1. It is important to reconsider it here since the magnitude of the error was only studied for one structure and it was evaluated with

Table 4. The finite sampling error in eigenvalue sum for the three structures. The potential in each case is that self-consistent potential generated by an 8^3 grid of $k \cdot p$ k -points. The 'grid' column refers to the grid dimension over which the eigenvalue sum is evaluated.

Grid	Eigenvalue sum for line (eV)	Eigenvalue sum for square lattice (eV)	Eigenvalue sum for FCC (eV)
2	-17.639 13	-15.787 47	-35.127 44
4	-18.120 05	-16.112 74	-35.695 91
6	-18.045 86	-16.350 61	-35.912 94
8	-18.193 95	-16.418 68	-35.983 56
10	-18.228 86	-16.378 36	-35.983 55
12	-18.237 73	-16.389 87	-35.955 65
14	-18.220 22	-16.404 95	-35.955 02
16	-18.204 38	-16.403 46	-35.961 36
18	-18.192 47	-16.389 51	-35.962 52
20	-18.205 34	-16.397 95	-35.960 22
22	-18.214 50	-16.402 37	-35.955 40
24	-18.215 93	-16.397 92	-35.957 02
26	-18.213 21	-16.396 35	-35.958 66
28	-18.207 53	-16.399 46	-35.958 79
30	-18.203 63	-16.399 78	-35.957 77
32	-18.207 85	-16.397 20	-35.957 61
40	-18.208 51	-16.397 33	-35.958 00
48	-18.211 35	-16.398 17	-35.957 43
56	-18.209 06	-16.398 08	-35.957 61
64	-18.209 29	-16.398 26	-35.957 50

relatively small k -point sets. The error is defined as the difference in the exact eigenvalue sum for finite and infinite grid.

Finite sampling errors in the eigenvalue sum are simply the result of numerically integrating the expression

$$\text{eigenvalue sum} = \int \varepsilon N(\varepsilon) w(\varepsilon) d\varepsilon. \quad (2)$$

Where $N(\varepsilon)$ is the density of states at energy ε and $w(\varepsilon)$ is the Fermi function which should in principle be evaluated at absolute zero but in practice is evaluated nearer room temperature.

Although the precise value of an eigenvalue sum depends on the exact $N(\varepsilon)$, its general convergence with respect to number of sampling points should be insensitive to small changes in that density of states. For this reason it is possible to investigate the finite sampling errors in the exact eigenvalue sum by the significantly faster procedure of investigating the finite sampling errors in the $k \cdot p$ sum. This has been done for the three structures under consideration. In each case we take the self-consistent charge density produced by an 8^3 grid of $k \cdot p$ points (see section 4) and use this potential to generate the 19 lowest exact solutions at $k = (\frac{1}{2}a, \frac{1}{2}b, \frac{1}{2}c)\pi$. These solutions are then used to generate the $k \cdot p$ eigenvalue sum over grids ranging from 2^3 to 64^3 . The results are displayed in table 4.

The rate of convergence differs for each of the structures studied. This demonstrates that relying on finite sampling errors to cancel for different structures is unlikely to be particularly successful. The convergence is worst for the line of atoms with the finite sampling error for a grid of 16^3 k -points is still of one order of 0.01 eV.

Table 5. The variation of $k \cdot p$ sum, exact sum and $k \cdot p$ error with cut-off energy for the FCC structure. The number of allowed plane waves (NPW) is also shown for each energy.

E_{cut} (eV)	NPW	$k \cdot p$ sum (eV)	Exact sum (eV)	$k \cdot p$ error (meV)
30	27	-29.855 31	-29.897 77	42.46
40	33	-30.124 06	-30.173 64	49.58
50	57	-31.491 17	-31.590 48	99.31
60	81	-32.212 95	-32.317 75	104.80
70	81	-32.212 95	-32.317 75	104.80
80	93	-32.731 58	-32.839 76	108.18
90	123	-33.549 22	-33.659 04	109.82
100	147	-34.101 74	-34.213 51	111.77
110	171	-34.471 71	-34.583 89	112.18
120	179	-34.537 47	-34.649 46	111.99
130	251	-35.572 79	-35.689 07	116.28
190	389	-35.983 56		

6. $k \cdot p$ error in the eigenvalue sum

Investigations of this error in 1 were restricted to the case of FCC aluminium using a cut-off energy of only 100 eV. Results of that investigation suggested that the error was of the order of 0.03 eV per atom (0.12 eV for a four-atom unit cell). This is larger than any of the other errors dealt with in the previous sections and, as it stands, puts an upper limit on the accuracy of the $k \cdot p$ method which restricts it to being a quick method for obtaining approximate results. In this section we show how to reduce this error by up to two orders of magnitude. With this modification the $k \cdot p$ total energy method is transformed to a quick method for calculating *absolute* energies to high accuracy.

Our method is to calculate the $k \cdot p$ eigenvalue sum, the exact eigenvalue sum and hence the $k \cdot p$ error in eigenvalue sum, for a series of cut-off energies far below the cut-off energy that we actually wish to use. (Given the low cut-off energy, the time taken to do this is very small.) We then try to relate the $k \cdot p$ error in eigenvalue sum and the $k \cdot p$ eigenvalue sum by a function of the form

$k \cdot p$ error in eigenvalue sum

$$= A + B(k \cdot p \text{ eigenvalue sum}) + C(k \cdot p \text{ eigenvalue sum}^2 + \dots). \quad (3)$$

(Note that the $k \cdot p$ error as a function of the plane wave cut-off is a series of steps and as such would not be amenable to fitting to a polynomial form). We use this function (3) to estimate the $k \cdot p$ error in eigenvalue sum for the high cut-off energy that we actually wish to use. By adding this correction term to the $k \cdot p$ eigenvalue sum for that cut-off energy, we hope to obtain a very accurate value for the exact eigenvalue sum.

The viability of this method rests on the extent to which the function (3) fitted at low cut-off energies may be extrapolated to the high cut-off energies which we require. To assess this method we demonstrate its use on the three structures under consideration.

For each structure we start by calculating the self-consistent charge density generated by an 8^3 grid of $k \cdot p$ k -points as described in section 4. We use this charge density and a series of energy cutoffs below 190 eV to generate the 19 lowest exact solutions at $k =$

Table 6. The variation of $k \cdot p$ sum, exact sum and $k \cdot p$ error with cut-off energy for the square lattice structure. The number of allowed plane waves (NPW) is also shown for each energy.

E_{cut} (eV)	NPW	$k \cdot p$ sum (eV)	Exact sum (eV)	$k \cdot p$ error (meV)
30	17	-14.369 65	-14.253 84	-115.81
40	29	-14.688 28	-14.697 19	8.91
50	39	-15.013 40	-15.032 52	19.12
60	47	-15.236 71	-15.261 19	24.48
70	55	-15.395 85	-15.421 90	26.05
80	77	-15.739 98	-15.770 41	30.43
90	77	-15.739 98	-15.770 41	30.43
100	117	-16.120 70	-16.155 23	34.53
110	117	-16.120 70	-16.155 23	34.53
120	151	-16.297 41	-16.332 90	35.49
130	151	-16.297 41	-16.332 90	35.49
190	297	-16.418 68		

Table 7. The variation of $k \cdot p$ sum, exact sum and $k \cdot p$ error with cut-off energy for the line structure. The number of allowed plane waves (NPW) is also shown for each energy.

E_{cut} (eV)	NPW	$k \cdot p$ sum (eV)	Exact sum (eV)	$k \cdot p$ error (meV)
30	39	-16.045 79	-16.058 59	12.80
40	51	-16.242 34	-16.265 47	23.13
50	79	-16.667 09	-16.711 15	44.06
60	87	-16.774 61	-16.820 31	45.70
70	119	-17.154 46	-17.201 79	47.33
80	157	-17.475 17	-17.523 55	48.38
90	169	-17.561 07	-17.609 93	48.86
100	225	-17.849 05	-17.899 04	49.99
110	233	-17.882 40	-17.932 44	50.04
120	277	-18.003 25	-18.053 52	50.27
130	301	-18.054 32	-18.104 72	50.40
190	583	-18.193 95		

$(1/2a, 1/2b, 1/2c)\pi$ and hence $k \cdot p$ eigenvalue sums over an 8^3 grid. For the same charge density, the same 8^3 grid and the same series of cut-off energies we evaluate the exact eigenvalues sums. The difference between these two values for the sums is of course the $k \cdot p$ error in eigenvalue sum. The results of this process are displayed in tables 5, 6 and 7 and in figures 1, 2 and 3.

Figures 1, 2 and 3 shows the variation of the $k \cdot p$ error as a function of the $k \cdot p$ sum. An increase in the cut-off energy decreases the $k \cdot p$ sum as a direct consequence of the variational theorem. Similarly an increase in the cut-off energy tends to increase the $k \cdot p$ error since the number of bands used to generate $k \cdot p$ solutions is fixed at 19 whereas the number of plane waves used to generate the exact solutions is enlarged. The magnitude of both of these effects tends to progressively decrease at higher energies. The $k \cdot p$ error tends to decrease most rapidly, reflecting the diminishing importance of high-energy bands in the $k \cdot p$ sum, and thus producing a saturation in the curve. All

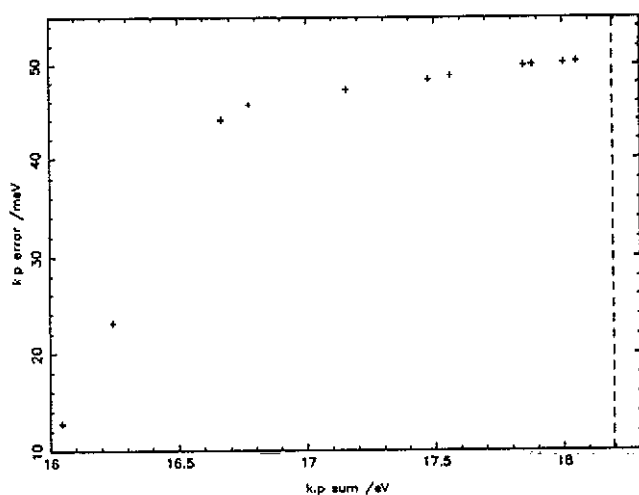


Figure 1. A plot of $k \cdot p$ error against $k \cdot p$ sum at various cut-off energies for a line of aluminium atoms. The vertical dotted line indicates the $k \cdot p$ sum corresponding to 190 eV.

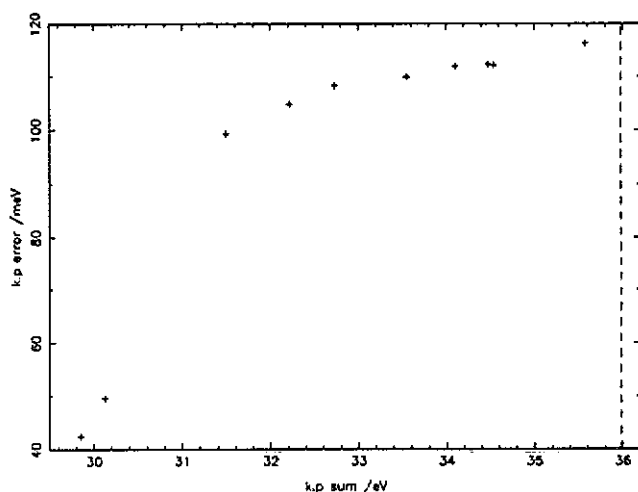


Figure 2. A plot of $k \cdot p$ error against $k \cdot p$ sum at various cut-off energies for FCC aluminium. The vertical dotted line indicates the $k \cdot p$ sum corresponding to 190 eV.

three curves (figures 1, 2, 3) show these effects, although the degree to which saturation has occurred does differ. The saturation is the most pronounced in figure 1 with the $k \cdot p$ error using a cut-off of 60 eV within 0.005 eV of the $k \cdot p$ error using a cut-off of 130 eV. In fact the convergence of the $k \cdot p$ error *per atom* is best for the four atom FCC (figure 2) and even for the square lattice (figure 3) the $k \cdot p$ error at 80 eV is within 0.005 eV of the $k \cdot p$ error at 130 eV.

For each of the figures 1–3, the abscissa of the point generated by the cut-off energy of 190 eV is known since it is the $k \cdot p$ eigenvalue sum for this energy, and is recorded in table 4. It is the corresponding ordinate that we wish to estimate by extrapolation. The

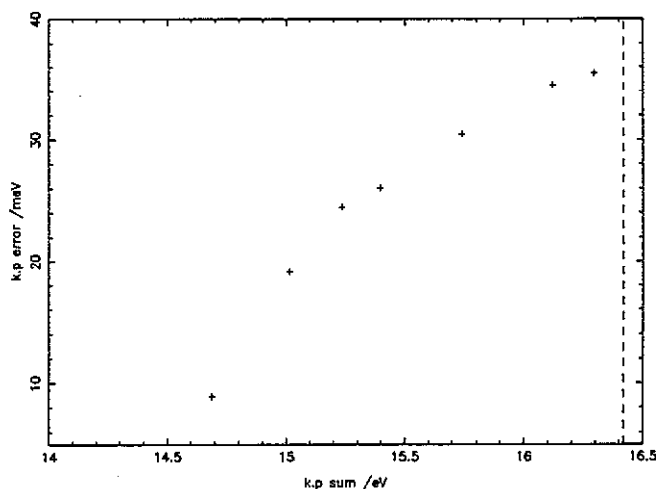


Figure 3. A plot of $k \cdot p$ error against $k \cdot p$ sum at various cut-off energies for a square lattice of aluminium atoms. The vertical dotted line indicates the $k \cdot p$ sum corresponding to 190 eV.

briefest glance at the smoothness of each of the curves reveals the viability of an extrapolation scheme. Our choice of the $k \cdot p$ eigenvalue sum as the argument in our expansion (3), also ensures that the distance over which we must extrapolate is relatively small.

The crudest method of estimating the $k \cdot p$ error is a zeroth order approximation. That is simply using the $k \cdot p$ error at some low energy to estimate the $k \cdot p$ error at 190 eV. We generate a better estimate by use of a first-order approximation.

For the line of atoms use of the data for cut-offs of 60 and 70 eV to linearly extrapolate to the $k \cdot p$ error for a cut-off of 130 eV gives a $k \cdot p$ error for that energy of 51.19 meV as compared to the true value of 50.40 meV. That is a discrepancy of around 1 meV. Extrapolating from the point generated by the 70 eV cut-off to that generated by the 130 eV cut-off represents an extrapolation in the $k \cdot p$ sum of 0.85173 eV ($= 18.05352 \text{ eV} - 17.20179 \text{ eV}$). The extrapolation of the $k \cdot p$ sum required to reach the point generated by the energy cut-off of 190 eV is only slightly further at 0.99216 eV ($= 18.19395 \text{ eV} - 17.20179 \text{ eV}$) so we may take a value of 1 meV as a fair indication of the error involved in linearly extrapolating to the 190 eV cut-off.

For the FCC structure use of the data for cut-offs of 80 and 90 eV to linearly extrapolate to the $k \cdot p$ error for a cut-off of 130 eV gives a $k \cdot p$ error for that energy of 113.87 meV as compared to the true value of 116.28 meV, i.e. a discrepancy of under 3 meV for a four-atom unit cell. Again extrapolation to an energy cut-off of 190 eV is only slightly further so we may take this error of under 1 meV per atom as the error involved in linearly extrapolating to the 190 eV cut-off.

The square lattice is the most difficult case. Linear extrapolation using the data at cut-offs of 60 and 70 eV gives a $k \cdot p$ error at 130 eV of 34.94 meV as compared to the true value of 35.49 meV. The level of this agreement is slightly fortuitous since the extrapolated value using the data at 70 and 80 eV gives the slightly inferior estimate of 37.52 meV.

In all cases then, even a crude linear extrapolation scheme of data at a low cut-off energy can reduce the $k \cdot p$ error in eigenvalue sum to around 0.001 eV per atom i.e. by

up to two orders of magnitude. This is an order of magnitude less than the finite sampling errors in the eigenvalue sum. In this context it may be concluded that the $k \cdot p$ error in eigenvalue sum has been reduced to a level of relative insignificance.

It should be stressed that our comments here are intended to demonstrate the general feasibility of an extrapolation strategy. The precise values of the lower cut-off energies that are used will vary from situation to situation. In practice one would like a method which allowed the predetermination of two cut-off energies at which evaluation of the $k \cdot p$ error would provide adequate data for linear extrapolation. That is two cut-off energies for which the $k \cdot p$ error is near saturation. One quick way to achieve this would be to evaluate the $k \cdot p$ error in the eigenvalues at a *single* k -point at a series of cut-off energies. The cut-off energy at which this $k \cdot p$ error saturates being indicative of the cut-off energy for which the full $k \cdot p$ error saturates. Since evaluating eigenvalues at a single k -point is much quicker than evaluating them over a whole grid, this procedure requires negligible extra time. The time taken for the whole $k \cdot p$ error determination is therefore dominated by that time to generate two exact sums over the Brillouin zone non-self-consistently and at a low cut-off energy. As has been said before, this time is insignificant in comparison with the time taken for a full self-consistent solution at the high cut-off energy.

Finally we note that this $k \cdot p$ error evaluation ran most smoothly for the structure for which the plane wave basis set was largest. This is an extremely desirable state of affairs given that it is these large structures for which the $k \cdot p$ method is required, traditional methods being too time consuming.

7. Conclusions

The two finite sampling errors general to any method of total energy calculation have been evaluated as a function of sampling density for three diverse aluminium structures. The errors have been shown to be significant even for relatively large k -point sets confirming that extremely time consuming calculations would be required to obtain accurate absolute energies. One error associated with the $k \cdot p$ method has been shown to be negligible. The other error associated with the $k \cdot p$ method has been shown to be large but a technique has been developed which allows even this error to be reduced to a level at which it is insignificant. The modified $k \cdot p$ total energy method introduces no significant errors and given the fast computational speed of the method it is possible to use k -point sets several orders of magnitude larger than previously possible, thus allowing finite sampling errors to be eliminated.

Acknowledgments

The authors thank the Science and Engineering Research Council for a studentship (IJR) and for 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

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
Cheng C, Needs R J, Heine V and Churcher N 1987 *Europhys. Lett.* **3** 475
Fu C L and Ho K M 1983 *Phys. Rev. B* **28** 5480
Goodwin L, Needs R J and Heine V 1990 *J. Phys.: Condens. Matter* **2** 351
Harris J 1985 *Phys. Rev. B* **31** 1770
Kohn W and Sham L J 1965 *Phys. Rev. A* **140** 1133
Monkhorst H J and Pack J D 1976 *Phys. Rev. B* **13** 5188
Payne M C, Joannopoulos J D, Allan D C, Teter M P and Vanderbilt D H 1986 *Phys. Rev. Lett.* **56** 2656
Pewder J P and Zunger A 1981 *Phys. Rev. B* **23** 5048
Robertson I J and Payne M C 1990 *J. Phys.: Condens. Matter* **2** 9837