CHAPTER SIX

Electronic Band Structure Calculations

6.1 Introduction

For a given crystalline solid the *BC* program of the *CAESAR* package carries out electronic band structure calculations using the EHTB method [1] as a function of wave vector for a set of wave vector points. Its input file, *filename.BI*, has the structural information and computational options about a solid to examine. This file is generated by running the *NEWBI* program interactively or by editing an existing *filename.BI* file.

The *BC* program produces output files *filename.BO*, *filename.BE*, *filename.BW* and *filename.BV*. The text file *filename.BO* records the input, output and the computational history. Crystal orbital energies are stored in a binary file *filename.BE*, gross populations of atomic orbitals in a binary file *filename.BW*, and overlap populations between atoms in a binary file *filename.BV*. The file *filename.BE* is used to obtain the Fermi level, total DOS plot and Fermi surfaces. The file *filename.BW* is used to obtain PDOS plots, and the file *filename.BV* to obtain COOP plots.

6.2 Creating an input file *filename.BI*

A *filename.BI* file can be constructed by using the *NEWBI* program. When this program is launched, it opens the *New .BI File* dialog box (*Figure 6.1*). In the following, we describe how to use the *New .BI File* dialog box.

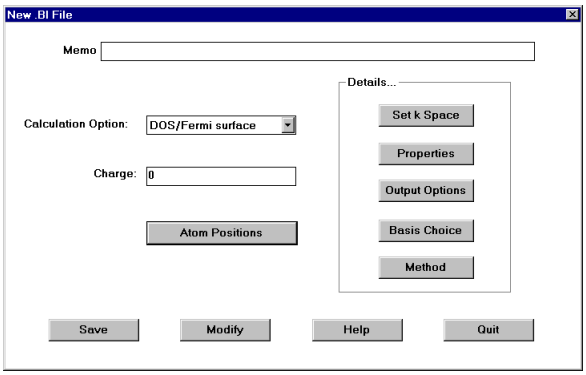

Figure 6.1. Dialog box *New .BI File*.

6.2.1 Basic parameter-setting for *filename.BI*

The *New .BI File* dialog box has five parameter fields to consider, namely, *Memo*, *Calculation Option*, *Charge*, *Atom Positions* and *Details...* In this section, we examine the first four items.

•*Memo*

This is a text string used for adding a short comment about a current calculation.

•*Calculation Option*

The arrow at the right corner of this field activates the drop-down list, which provides two options, *Dispersion only* and *DOS/Fermi surface*. The default option is *DOS/Fermi surface*.

Dispersion only:

This option is used when the sole purpose of electronic band structure calculations is to generate band dispersion relations for certain line segments of the FPZ. With this option, *BC* performs electronic structure calculations for k-points lying on the line segments the user specify.

DOS/Fermi surface:

This option is used to run *BC* for a set of k-points that are designed to sample the whole FPZ. It is needed when the purpose of electronic band structure calculations is not only to obtain band dispersion relations but also to calculate the properties that require the sampling of the k-points covering the whole FPZ (e.g., the Fermi level, DOS, PDOS, Fermi surfaces, etc.).

•*Charge*

This refers to the overall charge on the unit cell of a solid. The charge does not affect crystal orbitals and their energies calculated by the EHTB method, but it determines the Fermi level corresponding to the total number of valence electrons of the unit cell.

•*Atom Positions*

Electronic band structure calculations for a solid require the specification of not only the atoms and their positions of a unit cell but also the repeat vectors. A structure file *filename.CI* contains this information, which is copied into a *filename.BI* file when the user creates the latter using the *NEWBI*

program. The atom positions thus prepared are given in Cartesian coordinates. However, the atom positions of a *filename.BI* file can be specified either in the Cartesian or in the Z-matrix format.

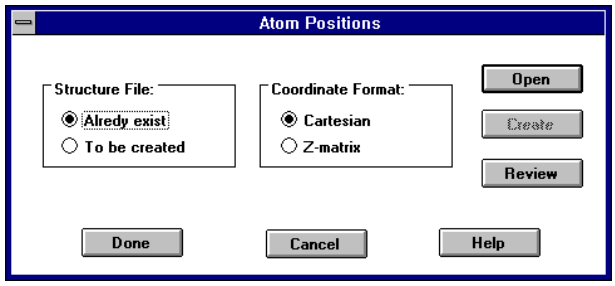

Figure 6.2. Dialog box *Atom Positions*.

The *Atom Positions* command of the *New. MI File* box opens the *Atom Positions* dialog box (*Figure 6.2*). The use of this box was described in detail in *Section 5.2.1*. When a proper *filename.CI* file does not exists, choose the *To be created* option of the *Structure File* field and select either the *Cartesian* or the *Z-matrix* option of the *Coordinate Format* field. Then click the *Create* button to open the atom-defining box, where the user should first specify the atoms and their positions of a unit cell. Then specify the positions of the VECT-A, VECT-B and VECT-C points to be used to determine the repeat vectors ***a, b*** and ***c***, respectively. Any atom of the unit cell can be used as the origin of the three repeat vectors. The positions of VECT-A, VECT-B and VECT-C refer to the end points of the vectors ***a, b*** and ***c***, respectively, with respect to the position of the atom chosen as the origin. It is necessary to specify only two of the VECT-A, VECT-B and VECT-C points for a 2D system, and only one of the VECT-A, VECT-B and VECT-C points for a 1D system. When a proper *filename.CI* file is present, the required structural information can be copied into a *filename.BI* file by selecting the *Already exist* button of the *Structure File* field and clicking the *Open* command.

When the *Done* button of the *Atom Positions* box is checked, the *Direct Space Setting* dialog box opens (*Figure 6.3*). When the Cartesian format is used, the *Atom* field of this box shows the *x*, *y* and *z* coordinates of VECT-A, VECT-B and VECT-C with respect to the coordinate origin, OR-0001. When the Z-matrix format is used, the *x*, *y* and *z* fields appear as blanks.

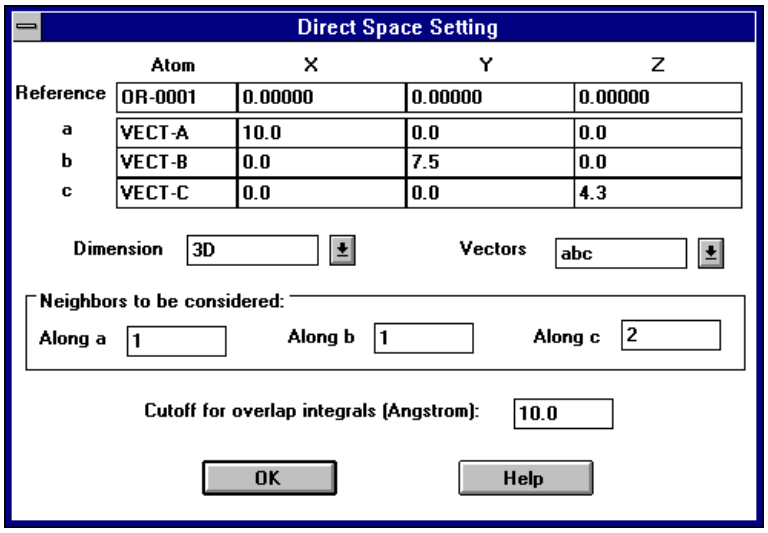

Figure 6.3. Dialog box *Direct Space Setting*.

When a 3D lattice consists of 2D layers or 1D chains, it is important to perform calculations for an isolated 2D layer or an isolated 1D chain, respectively. For this purpose, the dimensionality of the system can be re-defined by selecting an option of the *Dimension* field (*1D*, *2D* or *3D*). The options of the *Vectors* field depend on the selection made from the *Dimension* field. For calculations of a 3D solid there is only one choice, i.e., *abc*, which means that the three vectors ***a***, ***b*** and ***c*** are needed to define the 3D lattice. There are three planes (*ab*-, *bc*- or *ac*-plane) to choose for a 2D layer, and three directions (*a*-, *b*- or *c*-direction) to choose for a 1D chain. In most cases, these choices are sufficient. However, there occur cases when 2D layers are parallel to none of the *ab*-, *bc*- and *ac*-planes, and also when 1D chains are parallel to none of the *a*-, *b*- and *c*-directions. For such cases, the user should employ a new unit cell (with new repeat vectors **a'**, **b'** and **c'**) for the solid such that its 2D layer is contained in the *a'b'*-, *b'c'*- or *a'c'*-plane, or its 1D chains are parallel to the *a'*-, *b'*- or '*c*-direction. This can be done easily by using the *Builder* program.

The *Neighbors to be considered* field of the *Direct Space Setting* dialog box determines how many nearest neighbor unit cells, *n*, to include in the lattice sums along each repeat direction. To ensure that the lattice sums are converged, an adequate *n* value should be used for each direction. Along a given direction, $n = 1$ is sufficient, if the repeat distance along that direction is longer than 7 Å. If the length of a repeat vector, *d*, is shorter than 7 Å, then the *n* value along that direction should be such that $nd > 7$ Å.

Finally, the user should specify the cutoff distance for overlap integral calculations (default value = 10 Å). When the distance between two orbitals is greater than this distance, the integrals between them are set to zero.

After completing the *Direct Space Setting* box, click the *OK* button to return to the *New .BI File* dialog box.

6.2.2 Advanced parameter-setting for *filename.BI*

Certain computational parameters may not change for routine electronic band structure calculations, and *CAESAR* uses default values for them. Examples include the k-point sets, output options, basis sets, and options for property calculations. The *Details...* field of the *New .BI File* dialog box has five command buttons by which to specify these parameters. The use of these buttons is allowed only after the *Atom Positions* field is completed.

•*Set k Space*

The *Set k Space* button opens the *k Space Setting* dialog box (*Figure 6.4*). According to the dimensionality of a system (i.e., 1D chain, 2D layer, or 3D lattice), this dialog box defines the FPZ appropriate for the system. The dialog box has four fields to consider.

***k point set*:**

This field provides two options, i.e., *Default* and *User-defined*. The *Default* option gives three sets of k-points, i.e., *Small*, *Medium* and *Large*. The *User-defined* option gives three cases of defining k-points, i.e., *Area/volume*, *Line segments*, and *One point*. Unless requested otherwise by the user, the *NEWBI* program uses the *Small* k-point set of the *Default* option.

A set of k-points sampling the FPZ can be defined on the grid points covering an appropriate part of the FPZ, which depends upon the symmetry of the real-space unit cell:

$-0.5 \le k_a \le 0.5$
$-0.5 \le k_b \le 0.5$ triclinic & trigonal cells
$-0.5 \le k_c \le 0.0$

$-0.5 \le k_a \le 0.5$
$0.5 \le k_b \le 0.0$ monoclinic & hexagonal cells

$-0.5 \le k_c \le 0.0$

$-0.5 \le k_a \le 0.0$
$-0.5 \le k_b \le 0.0$ orthorhombic, tetragonal & cubic cells
$-0.5 \le k_c \le 0.0$

The default k-point sets are obtained by defining k-points on the grids that result when half the length of each reciprocal vector starting from the origin (e.g., $0.0 \le k_a \le 0.5$, $0.0 \le k_b \le 0.5$, and $0.0 \le k_c \le 0.5$) is divided into *n* equi-distant points. For the *Small*, *Medium* and *Large* sets, *n* = 6, 8 and 11, respectively. Therefore, the choice of the *Small* set for a 3D solid generates 11x11x6 k-points for a triclinic or trigonal system, 11x6x6 k-points for a monoclinic or hexagonal system, and 6x6x6 k-points for an orthorhombic, tetragonal or cubic system. The program *BC* checks the symmetry of the unit cell and constructs the k-point set accordingly.

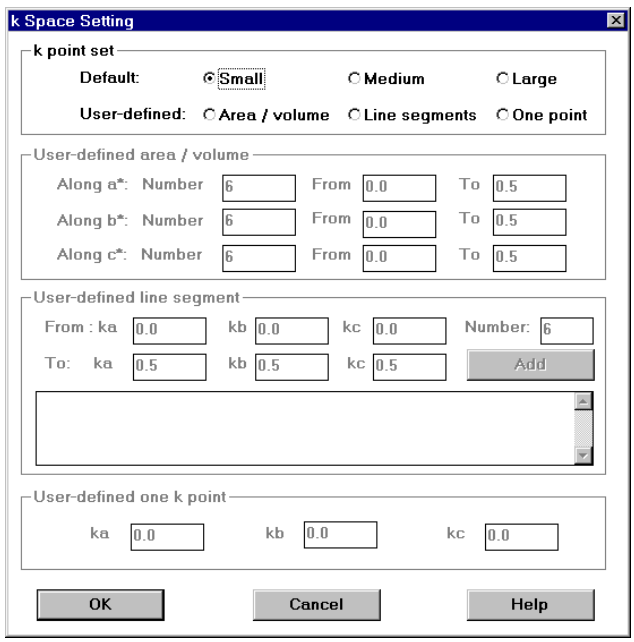

Figure 6.4. Dialog box defining a 3D FPZ.

For highly symmetric systems such as hexagonal and cubic, a much smaller part of the FBZ can be used to determine a k-point set sampling the whole reciprocal space. Such a set saves the computational time in particular for first principles calculations. *BC* uses the k-point sets as defined above, because the orbital energies $e_n(k)$ calculated for a set of k points can be easily fitted by the cubic spline fitting method to generate crystal orbital energies at other k points for which electronic band calculations are not actually carried out. In addition, EHTB calculations are fast.

The *Area/volume* button of the *User-defined* option activates the *User-defined Area/volume* box, which allows the user to define a customized set of k-points. For this purpose, specify the number *n* in the *Number* field in each reciprocal vector direction, so that any length of a reciprocal vector, defined by the values in the *From* and *To* fields, can be divided into *n* equi-distant points. For example, the *k Space Setting* in *Figure 6.5* is concerned with a 2D layer lying in the crystallographic bc-plane. This example has more k-points along the b*-direction than along the c*-direction. The use of such a k-point set is recommended when electronic energy bands are more dispersive along

one direction (in this case, along the b*-direction) because all k-points are equally probable in the reciprocal space.

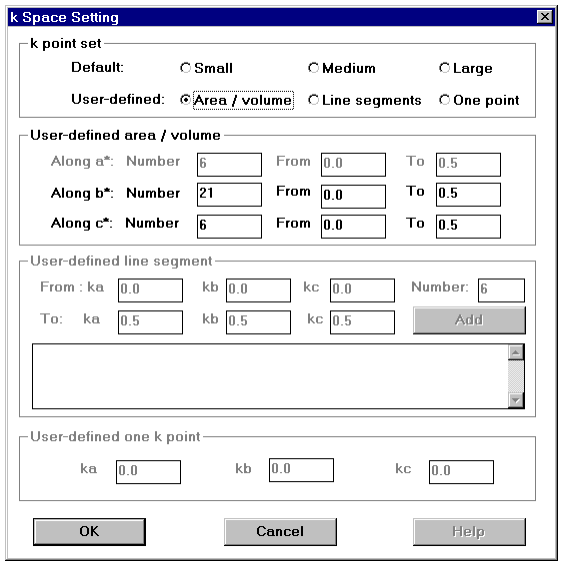

Figure 6.5. Dialog box defining a 2D FPZ.

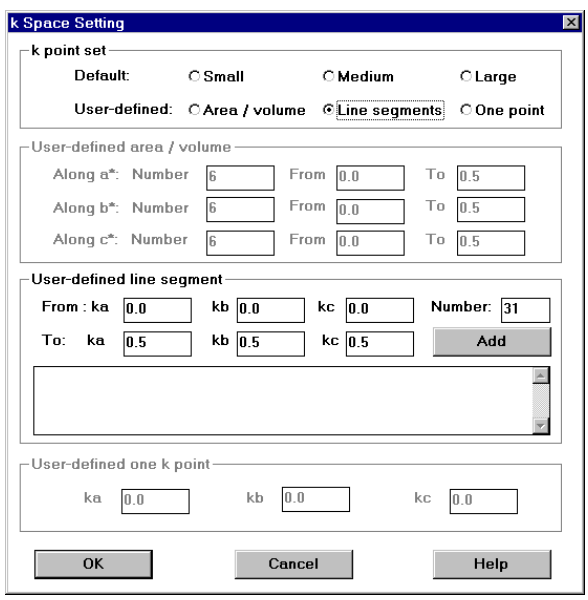

Figure 6.6. Dialog box defining k points along a line segment.

It is often necessary to calculate dispersion relations of certain bands only for the k-points lying along specific line segments. This is achieved by clicking the *Line segment* button of the *User-defined* option. This activates the *User-defined line segment* box. Then specify the k_a, k_b and k_c components of the starting and the end k-points of a line segment and the number of the total k points on the segment. Then click the *Add* button to list the line segment information in the space lying

below. The user may define as many line segments as needed. For example, *Figure 6.6* defines the body-diagonal segment of a 3D FPZ with 31 k-points on it.

Sometimes one may need to calculate electronic band structure calculations only for a specific k-point. In such a case, click the *One point* button of the *User-defined* option to activate the *User-defined one point* box and specify the k_a, k_b and k_c components of that point.

•*Properties*

To calculate gross populations of atomic orbitals as well as overlap populations between atoms, click the *Properties* button of the *New .BI File* box. This opens the *Property Calculation Options* dialog box (*Figure 6.7*), which provides two options:

Atomic orbital population:

If this option is checked, the parameter fields are activated and the *All COs* button is automatically selected as the default. This signals the *BC* program to perform calculations of gross atomic orbital populations for every crystal orbital separately (at each k point of a given k point set) and store the results to prepare for PDOS analysis. If desired, the user may limit the range of the crystal orbitals to be used for the calculations of gross atomic orbital populations. Unless requested by the user, gross population analysis is not carried out.

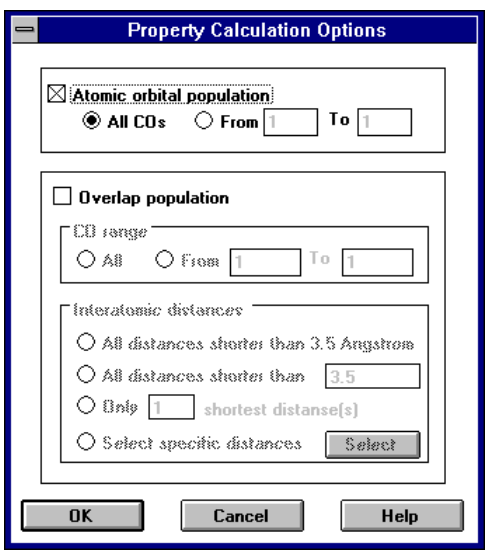

Figure 6.7. Dialog box *Property Calculation Options*.

Overlap population:

To calculate overlap populations between atoms, click this option (*Figure 6.8*) to activate the *CO range* and *Interatomic distances* fields. If desired, the user may limit the range of the crystal orbitals to be used for overlap population calculations. For example, *Figure 6.8* shows the default setting, for which overlap populations are calculated for every crystal orbital for all "bonds" (i.e., interatomic contacts) shorter than 3.5Å. There are four ways of selecting the kinds of distances for which to calculate overlap populations:

(1) all bonds shorter than 3.5Å
(2) all bonds shorter than a distance the user specifies
(3) all bonds belonging to the n shortest distances
(4) all bonds lying within a specific distance range

The last option activates the *Select* button, which is used to open the *Define Distances for COOP* dialog box (*Figure 6.9*). The user can define up to 5 different kinds of bonds. For each kind of bond, the user should provide the names of the atoms defining the bonds and the range of their distances. For example, *Figure 6.9* defines two kinds of bonds. The first set includes all bonds between the atoms C and O whose lengths lie between 1.00 and 2.00 Å. The second set includes all bonds between the atoms C-01 and C-03 whose lengths lie between 1.40 and 1.45 Å.

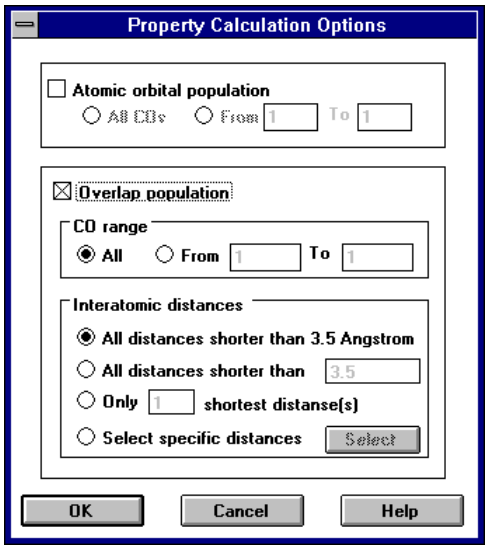

Figure 6.8. Parameter setting for overlap population calculations.

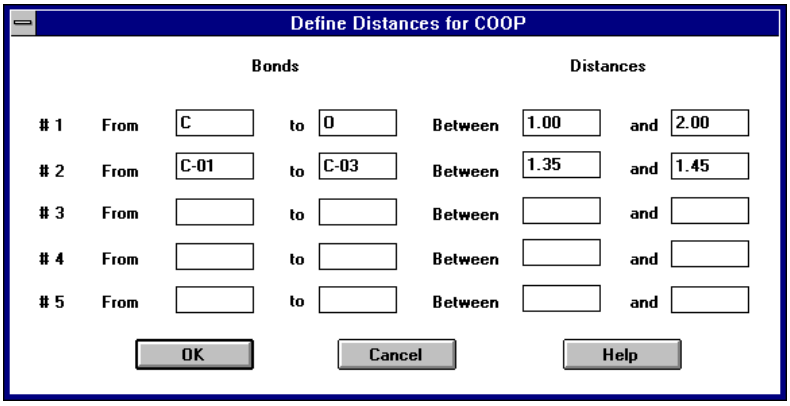

Figure 6.9. Dialog box *Define Distance for COOP*.

•*Output Options*

To control output options, click the *Output Options* button (*Figure 6.1*) to open the *Output Options* dialog box. None of the option fields of this box is activated, because the default is to print none. Each option field is activated by checking the small square box.

Overlap integrals (Sij) and ***Hamiltonian elements (Hij)***:

These are used to print the matrix elements between atomic orbitals. If the button *All AOs* is checked, *BC* prints the whole S_{ij} or H_{ij} matrix. It is possible to print only a certain part of the matrix by selecting the range of the atomic orbitals. For this purpose, the sequence numbers of the starting and end atomic orbitals of the chosen range should be given. These numbers are determined by the atom sequence in the input file *filename.BI* and then by the orbital sequence within each atom, i.e.,

s for H

s, x, y, z for non-transition metal elements

x^2-y^2, z^2, xy, xz, yz, s, x, y, z for transition metal atoms

Energy levels* and *Crystal orbitals:

These determine the output options for each k-point. By checking the option *Energy levels*, one can print all energy levels of the crystal orbitals or only a part of them. If the *Crystal orbitals* button is checked, it is necessary to determine the range of the crystal orbitals to be printed and the range of the atomic orbitals to be printed for the crystal orbitals chosen. It is recommended to reduce the amount of output by properly limiting the print ranges.

Unless the output options are defined for individual k points, there will be no output of crystal orbitals and crystal orbital energies. The *Output for k points* button activates the two drop-down lists lying below, in which the commonly used k points and the printing options are listed. The numbers in the parenthesis following a symbol in the *k-point list* field represent the k_a, k_b and k_c values of the k-points.

When a desired k-point is clicked from the *k-point list*, the symbol of the selected k-point is displayed in the parameter field lying above the list. The output options for this selected k-point can be chosen from the *Option list* by clicking them. Then the selected options are displayed in the parameter field lying above the list. It is possible to choose more than one option for each k-point selected. Finally, the selected items are sent to the large drop-down list on the left by clicking the *Select* button. It is possible to modify or delete the items by using the mouse as in a normal editor. A specific k point, once selected and moved into the list, cannot be selected again. For each k-point selected, the options to be chosen must be all different. If the symbol *A* is selected, one cannot select any other k points since all k points have already been selected.

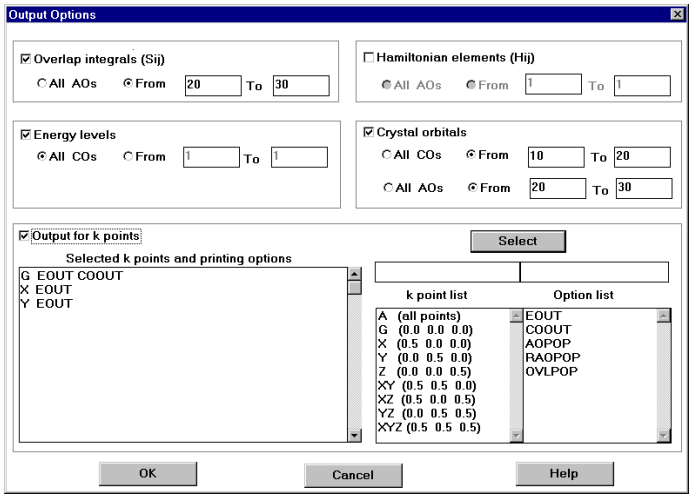

Figure 6.10. Sample of output parameter setting.

It is important to remember that the selected k-points are part of the k-point set for which to perform electronic band structure calculations. Outputs for k points other than those given in the *k-*

point list can be printed by specifying their three components in the parameter field located above the list. An example of an output option is given in *Figure 6.10*. This example will print the overlap integrals for the atomic orbitals from #20 to #30, but will not print the Hamiltonian matrix elements. For all selected k points, the energy levels of all crystal orbitals will be printed, but the orbital vectors will be printed only for the crystal orbitals from #10 to #20. For each of these crystal orbitals, the coefficients will be printed only for the atomic orbitals from #20 to #30. The selected k points are the Γ point, for which both the crystal orbitals and crystal orbital energies will be printed. For the points X and Y, only the crystal orbital energies will be printed.

•*Basis Choice*

This command opens the *Basis Set Options* dialog box, which is used to specify the STO's needed for EHTB electronic structure calculations. Chapter 5 describes in detail the STO basis sets stored in the *CAESAR* package and how to use them.

•*Method*

This command opens the *Method* dialog box, whose functions are described in Chapter 5.

A *filename.BI* is complete after completing the basic and advanced parameter settings. To run the *BC* program using this input file, it should be saved before exiting the *New .BI File* dialog box.

6.2.3 Editing *filename.BI*

The user may generate a new *filename.BI* file by editing an existing *filename.BI*. This can be achieved by using any text editor or by using the *Modify* command of the *New .BI* dialog box. When clicked, it opens the *Open* dialog box to allow the user to locate the existing file and copy its contents into the parameter fields of *New .BI File* dialog box. Then the user can edit the file by going through the appropriate fields of the *New .BI File* dialog box where changes have to be made.

6.3 Output files of the *BC* program

In reading a *filename.BO* file of a 2D layer or a 1D chain, it is important to notice the conventions for k-points adopted in *BC*. In the outputs for any 2D calculations, the components of a k-point are always k_a and k_b. Suppose that a 3D solid consists of 2D layers parallel to the bc-plane, and one calculates for an isolated layer using its repeat vectors defined by their end positions,

POS VECT-B x1 y1 z1
POS VECT-C x2 y2 z2

where *VECT-B* is the end atom of the translational vector along the b-direction located at (*x1, y1, z1*), and *VECT-C* is the end atom of the translational vector along the c-direction located at (*x2, y2, z2*). In the file *filename.BO*, each k-point is represented as (k_a, k_b, 0) instead of (0, k_b, k_c). For calculations of a 2D layer system, k_a and k_b simply represent the reciprocal vectors corresponding to the first and second repeat vectors defined in the input file *filename.BI*, respectively. A similar situation arises for calculations for an isolated 1D chain. For example, imagine that this chain is parallel to the c-direction, and its repeat vector is defined by its end position,

POS VECT-C x1 y1 z1

In the file *filename.BO*, each k-point is represent as (k_a, 0, 0) instead of (0, 0, k_c). For calculations of an isolated 1D chain, k_a simply represents the reciprocal vector corresponding to the repeat vector defined in the input file *filename.BI*. This convention holds throughout the whole *CAESAR*.

A binary file for crystal orbital energies, *filename.BE*, is always created after running *BC*. It contains all orbital energies calculated for every k-point of a chosen k-point set. When requested, *BC* generates gross atomic orbital populations for every crystal orbital at each k-point of the given k-point set. Results of these calculations are stored a binary file *filename.BW*. The file is large, so its deletion is recommended after completing the intended calculations (i.e., PDOS plots). When requested, *BC* generates overlap populations between atoms and stores them in a binary file *filename.BV*. This file is also large, so its deletion is recommended after completing the intended calculations (i.e., COOP plots).

6.4 Limitations of *BC*

The dynamic memory management is employed in *BC*. There are a few limitations. The maximum number of atoms allowed for a unit cell is 500. The maximum numbers of atomic orbitals and chemical bonds allowed for a unit cell are 1,500 and 2,000, respectively. The number of nearest neighbor unit cells to include for lattice sums should be less than 11. The number of elements in a unit cell can be up to 40. If the overlap populations are calculated by specifying distances, the maximum number of the distance sets is 5.

Figure 6.11. Warning message calling for the cancellation of a huge computational job.

During the execution of *BC*, a number of temporary files named *BC.n* are generated, where n is an integer between 1 and 99. The sizes of these files dramatically increase with increasing the number of atoms in a unit cell and the number of k points to be calculated. *CAESAR* provides a utility program to compare the size requirements for the requested computation with the available memory and space. If the requirements are not met, a warning message is displayed on the screen (*Figure 6.11*). In such a case, the computational job should be canceled.

6.5 Examples

Examples for the *BC* program will be presented after we describe the *PC* and *PP* programs in Chapter 7.

References

1. (a) Whangbo, M. -H.; Hoffmann, R. *J. Am. Chem. Soc*. **1978**, *100*, 6093.
(b) Whangbo, M. -H.; Evain, M.; Hughbanks, T.; Kertesz, M.; Wijeyesekera, S.; Wilker, C.; Zheng, C.; Hoffmann, R. *QCPE* **1989**, *9*, 61.