Cyclic cluster model for calculating defects in solids using the local density approximation

J Miró†, P Deák†, C P Ewels‡ and R Jones‡

† Department of Atomic Physics, Technical University Budapest, Budafoki út 8, H-1111 Budapest, Hungary
‡ Department of Physics, University of Exeter, Stocker Road, Exeter EX4 4QL, UK

Abstract.   The cyclic cluster model (CCM) was introduced earlier for defect calculations in the framework of semi-empirical quantum chemical methods with only two-centre interactions. The CCM represents a good balance between an approximate description of localized and extended states, offering most of the advantages of supercell methods at the cost of a molecular cluster calculation. This paper presents the adaptation of the CCM to an *ab initio* density functional Hamiltonian, incorporating the local density approximation, norm-conserving pseudopotentials and a Gaussian basis. Test results on perfect diamond cyclic clusters illustrate the advantages of the CCM.

1. Introduction

Structural modelling of defects and surfaces in semiconductors is one of the most important practical challenges to computational solid state physics. The local density approximation (LDA) to density functional theory [1, 2] has proven very successful in describing periodic many-electron systems. Imperfections of the crystal, however, lift the ideal periodicity and, thereby, the principal basis of conventional band theory. In the supercell model (SCM) the periodicity is artificially restored in a sufficiently large unit cell (LUC) around the defect [3]. However, in the case of local defects, the commonly used and very convenient plane wave expansion for one-electron states may prove to be very time consuming, especially for light impurities with incompletely screened cores. As an alternative, one-electron states may be expanded over a basis of localized orbitals (e.g., linear combination of Gaussian type orbitals, LCGTO), but such an SCM calculation becomes prohibitively expensive [4, 5]. As a consequence, LCGTO–LDA calculations are usually applied to a simple molecular cluster model (MCM [6]) of the defective solid [7]. The crystalline background is accounted for by hydrogen atoms saturating the dangling bonds on the cluster surface. This has several disadvantages.

— The 'crystalline' states are poorly described. The electronic states of the perfect MCM cannot be assigned to states of the perfect crystal band structure, and the convergence of the HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital) energies to the valence band (VB) and conduction band (CB) edges with the size of MCM is slow and oscillatory upon increasing the size of the MCM [10].

— The point group symmetry is reduced even without introducing the defect (from O_h to at least T_d in the case of group IV and III–V semiconductors), and the uniformity of the

sites is lost. The environment of an impurity in different sites may change considerably (due to the limitation of the cluster size). The comparison of the energy in the different configurations then becomes problematic.

— The bonds to hydrogen atoms on the cluster surface are poor mechanical substitutes for the bonds in the crystalline background. They also create a dipole layer on the surface of the MCM, shifting ionization energies and heats of formation in a manner which is hard to account for. Also, this dipole layer induces charge transfer between concentric shells of the cluster, even if it consists of nominally equivalent atoms.

This paper presents a solution to these problems by adapting the cyclic cluster model (CCM) [8] to LCGTO–LDA calculations. Section 2 describes the CCM and section 3 gives its adaptation to LCGTO–LDA theory. In section 4, bench-mark results are given for diamond to show that CCM calculations are an acceptable alternative to SCM ones even in reproducing the extended states of the crystal.

2. The cyclic cluster model

Several attempts have been made to eliminate the disadvantages of the MCM (while retaining its simplicity) by applying the Born–Kármán cyclic boundary conditions directly to a cluster, i.e. to a 'crystal' with $N = 1$ unit cell (see e.g. [9–11]). This can be done by modifying the matrix of interatomic distance vectors while the calculations are still done in direct space. The total point group symmetry of the perfect crystal is preserved; the system remains uniform and free from surface dipoles. The coincidence of the cyclic boundary condition with the Bloch condition for $N = 1$ means that all one-electron states belong to the centre of the Brillouin zone (BZ), i.e., $K = 0$. It must be realized that choosing an LUC means a reduced, small Brillouin zone (SBZ). The points of the SBZ represent a set of primitive k vectors. It was recognized [12] that choosing an LUC as a symmetric multiple of the primitive unit in an appropriate way may reduce the full star of primitive k_q vectors into the centre of the SBZ so that they satisfy the condition

$$\sum_q \omega(k_q) \sum_{|R|=R_m} e^{ik_q \cdot R} = 0 \qquad n = 1, 2, \ldots m_{max} \tag{1}$$

with the $\omega(k_q)$ weighting factors equal to the number of branches in their star. The inner summation in (1) goes over lattice vectors of equal length. (1) was given by Chadi and Cohen [13] to select so-called special k point sets which give a good approximation to the electron density by substituting the integration over the *entire* BZ by a finite weighted sum,

$$n(r) = \int_{BZ} n_k(r) \, dk \approx \sum_q \omega(k_q) n_{k_q}(r) \tag{2}$$

where

$$n_k(r) = \sum_{ij} b_{ij}(k) \varphi_i(r) \varphi_j(r) \qquad b_{ij}(k) = 2 \sum_\lambda^{occupied} c_{i\lambda}^*(k) c_{j\lambda}(k) \tag{3}$$

is the electron density if the one-electron states are expanded as

$$\psi_\lambda = \sum_i c_{i\lambda}(k) \varphi_i(r). \tag{4}$$

The approximation in (2) improves with increasing m_{max}. (N.B. Usually, those k points for which the inner summation yields zero are referred to as special. However, it is sufficient

if a set of k_q vectors satisfies the entire equation.) This concept was used in semi-empirical SCM calculations [12, 14], but the idea suits the cyclic cluster model equally well [15].

The cyclic cluster model (CCM) is defined as the application of cyclic boundary conditions to an LUC chosen in the way described above to represent a special k point summation performed implicitly [8]. The CCM is practically a $K = 0$ SCM calculation with interactions neglected beyond the boundary of the Wigner–Seitz (WS) cell corresponding to the LUC. On the one hand, this limits the accuracy of reproducing extended states but, on the other, it excludes most of the artificial defect–defect interactions of the SCM. Therefore, the CCM is a well balanced approximation, unifying advantages of SCM (full point group symmetry, homogeneous environment for the defect, no surface states or dipole layer, unambiguous assignment of background states to real extended states) with the computational simplicity of MCM.

The main difficulty with CCM is the handling of interactions involving atoms on the surface of the WS cell without loss of symmetry. In semi-empirical quantum chemical approximations, where the basis is assumed to be orthogonal and only two-centre interactions are involved, this problem was solved by a simple averaging procedure [8, 11, 15]. Such CCM calculations were successfully applied in describing the behaviour of H and O in crystalline silicon (see, e.g., [16] and [17]). In the case of the LDA, however, the GTO basis is not orthogonal, and three-centre interactions have to be dealt with.

3. The CCM in the LCGTO–LDA formalism

The basic equation of the LDA is the Kohn–Sham equation [2]:

$$\left[-\frac{1}{2}\Delta + V_{ion}(\boldsymbol{r}) + \int \frac{n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\,\mathrm{d}^3 r' + \mu_{xc} \right] \psi_\lambda(\boldsymbol{r}) = \varepsilon_\lambda \psi_\lambda(\boldsymbol{r}) \tag{5}$$

with the exchange–correlation potential

$$\mu_{xc} = \frac{\mathrm{d}[n(\boldsymbol{r}) E_{xc}(n(\boldsymbol{r}))]}{\mathrm{d}[n(\boldsymbol{r})]}. \tag{6}$$

The exchange–correlation energy density is expressed as (see, e.g., [18])

$$E_{xc} = A[n(\boldsymbol{r})]^{S_0}. \tag{7}$$

Using expansion (4), (5) can be turned into a set of algebraic equations which has to be solved iteratively (because of (2) and (3)) until the total energy becomes self-consistent (SCF). The total energy can be expressed as a functional of the electron density [1],

$$E[n(\boldsymbol{r})] = T[n(\boldsymbol{r})] + \int V_{ion}(\boldsymbol{r})\,\mathrm{d}^3 r + \frac{1}{2}\int \frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\,\mathrm{d}^3 r'\,\mathrm{d}^3 r + \int n(\boldsymbol{r}) E_{xc}\,\mathrm{d}^3 r. \tag{8}$$

The potential energy of the valence electrons in the field of the ionic cores, $V_{ion}(\boldsymbol{r})$, is usually given by norm-conserving pseudopotentials but that is not necessary. If full potentials are used, of course, the core electrons needed to be included. The third term in (8), the Hartree energy, contains, in principle, four-centre integrals but those can be avoided by fitting the electron density, in each step of the SCF iteration, by an independent expansion,

$$\tilde{n}(\boldsymbol{r}') = \sum_k c_k g_k(\boldsymbol{r}') \tag{9}$$

so that the Hartree energy for the difference $|n(\boldsymbol{r}) - \tilde{n}(\boldsymbol{r}')|$ is minimized [7].

Let both the basis functions for the wave function in (4) and the basis functions of the fitted electron density in (9) be expressed as sums of GTOs:

$$\varphi_i(\boldsymbol{r}) = \sum_l \chi_i(\boldsymbol{r} - \boldsymbol{R}_i - \hat{A}l) \tag{10}$$

$$g_k(\boldsymbol{r}') = \sum_l h_k(\boldsymbol{r}' - \boldsymbol{R}_k - \hat{A}l) \tag{11}$$

where i and k run over atoms in the WS cell corresponding to the LUC, while l gives all possible repetition of the LUC defined by the matrix of unit vectors $\hat{A}$.

In the case of two-centre terms, such as the overlap integral, the cyclic boundary condition is realized (assuming one of the centres at the centre of the real WS cell, i.e. at $l = \boldsymbol{0}$) by

$$\langle \varphi_i \mid \varphi_j \rangle = \sum_l^{min|\boldsymbol{R}_i - (\boldsymbol{R}_j + \hat{A}l)|} \langle \chi_i(\boldsymbol{r} - \boldsymbol{R}_i) \mid \chi_j(\boldsymbol{r} - \boldsymbol{R} - \hat{A}l) \rangle \tag{12}$$

where the summation contain only those terms for which the distance is minimal (one term, or several of equal value). In three centre terms of the CCM, like those related to the ionic potential (and similarly, the ones related to the Hartree term with functions g_k instead of the potential),

$$V_{ion}(\boldsymbol{r}) = \sum_k V_k = \sum_k \sum_l V(\boldsymbol{r} - \boldsymbol{R}_k - \hat{A}l) \tag{13}$$

$$\langle \varphi_i \varphi_j \mid V_k \rangle = \sum_{l,l'}^* \langle \chi_i(\boldsymbol{r} - \boldsymbol{R}_i) \chi_j(\boldsymbol{r} - \boldsymbol{R}_j - \hat{A}l) \mid V(\boldsymbol{r} - \boldsymbol{R}_k - \hat{A}l') \rangle \tag{14}$$

where the summation $\sum^*$ is carried out in such a way that only terms with

$$\min_{l'} |\boldsymbol{R}_i - (\boldsymbol{R}_k + \hat{A}l')| \cap \min_{l'} |(\boldsymbol{R}_j + \hat{A}l) - (\boldsymbol{R}_k + \hat{A}l')| \tag{15}$$

remain, with l chosen again to satisfy

$$\min_l |\boldsymbol{R}_i - (\boldsymbol{R}_j + \hat{A}l)|. \tag{16}$$

Conditions (15) and (16) have been incorporated within a variant of the AIMPRO program code by Jones and Briddon [20].

By implementing CCM in the overlap matrix, in fact, a small overlap of an atomic orbital with another on a distant atom in the real LUC is replaced by a bigger overlap with one on an equivalent atom in a repeated LUC. If the cluster is not large enough this may cause the overlap matrix to become non-positive definite, especially if the expansions (10) and (11) have been chosen to describe free atomic wave functions (as in the AIMPRO code). Therefore, the orbital exponents of the GTOs (defined in [20]) have been refitted to ensure that

$$\langle \chi_i(\boldsymbol{r} - \boldsymbol{R}_i) \mid \chi_j(\boldsymbol{r} - \boldsymbol{R}_j - \hat{A}l) \rangle \leqslant 10^{-14} \tag{17}$$

for every l allowed by the condition (16). This is justified by the common knowledge that atomic wave functions 'contract' in bonding. Sometimes, the GTO expansion is determined by using a localization potential in the atomic calculation (see e.g. [19]). A similar procedure to optimize the basis functions is under development.

4. Results

Test calculations have been performed for diamond in order to compare results of a perfect CCM to an MCM calculation on a cluster with comparable size on one hand, and with the best LCGTO–LDA band structure on the other. A 32-atom cyclic cluster has been used, with

$$\hat{A} = a \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \tag{18}$$

(a is the lattice constant), representing the stars of the wave vectors $\Gamma[000]$, $X[100]$, and $\Sigma[\frac{1}{2}\frac{1}{2}0]$ (in $2\pi/a$ units). These satisfy (1) up to $m_{max} = 6$.

Comparison was made between a tetrahedral MCM, $C_{35}H_{36}$, and the CCM C_{32} (figure 1). In both cases, four GTOs for every carbon valence shell orbital (i.e. 16 for each atom) were used with a contracted basis for hydrogens (two GTOs for the s and p orbitals but with fixed ratios, i.e. in effect four orbitals per hydrogen). The calculations have been carried out on an IBM RISC 6000/32H work station with 64 Mbytes of RAM. The comparison of electronic structures and the time required to obtain them can be seen in figure 2 and table 1. The one-electron states of the CCM were assigned to the appropriate primitive k vectors (crosses). The band structure (continuous lines) was obtained by graphic interpolation, in order to guide the eye. The one-electron energies in both calculations have been rigidly shifted to have the HOMO at zero energy. Beside the bands being narrower by $\sim$2 eV and the gap being wider by 0.6 eV in the MCM (cf the real bands in table 2), the CCM offers the advantages of correct symmetry and degeneracy of electronic states as well as zero charge on every atom, while the computational time is substantially less (!) than that of the MCM.

Table 1. Comparison of the running times for C_{32} CCM and $C_{35}H_{36}$ MCM calculations on an IBM RISC 6000/32H work station using four GTOs for both s and p orbitals of carbon, and a contracted basis for hydrogens.

	C_{32} (CCM)	$C_{35}H_{36}$ (MCM)
Basis set	C(4, 4)	C(4, 4) + H(2, 3)$_{contracted}$
Setup of first $\hat{H}$	447 s	593 s
1 step in SCF procedure	190 s	378 s
No of cycles till conv.	13	13
Total time	2916 s	5507 s

With respect to computational time, the CCM is obviously even more favourable if compared to an SCM calculation. It is also better suited for describing localized orbitals of isolated defects, since direct interactions between repeated defects are neglected. A remaining question is how accurately the CCM is able to describe extended states compared to SCM calculations. Therefore, our CCM results were also compared to the band structure obtained from a calculation [4] with the same pseudopotential but with a much larger basis set (ten GTOs for every valence shell orbital, i.e. 40 for each atom). Since no basis set with polarization functions is yet available for the AIMPRO code, the effect of the basis size was tested by contracting our basis set, i.e. reducing the effective variational freedom to four orbitals per carbon atom. The results are shown in table 2.

The C_{32} CCM calculation with 16 AO/atom reproduces the valence band rather well. The minimum of the conduction band, however, is obtained at Γ. It is known that the
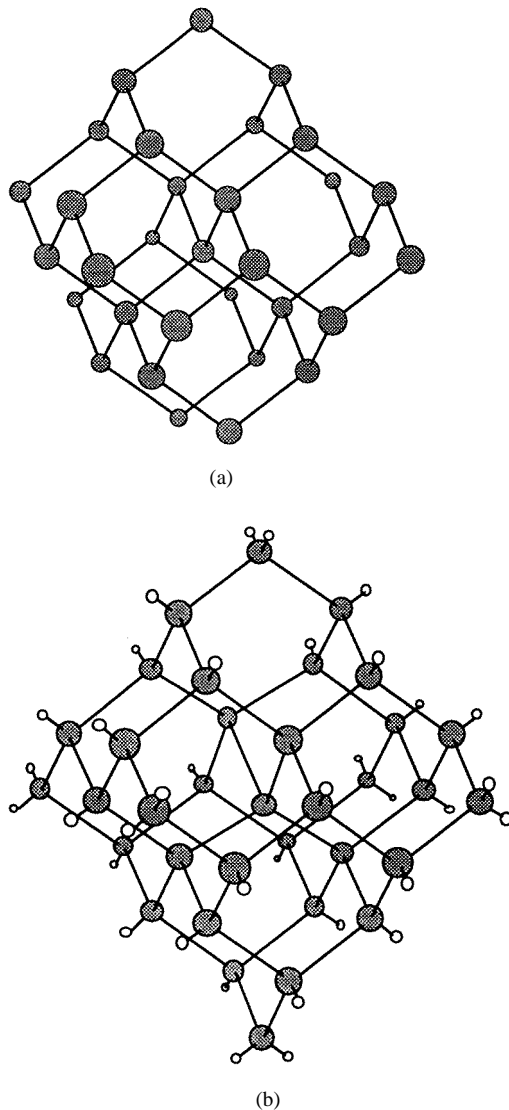
(a)

(b)

Figure 1. The C_{32} CCM (a) and the $C_{35}H_{36}$ MCM (b) of diamond. Small empty circles represent the hydrogen atoms.

reproduction of the indirect gap requires a large basis. If results with the 4 AO/atom basis are compared to the one with 16 AO/atom, it can be seen that the change in the energy of the X_{1C} state is by far the biggest. It shifts downwards by 6.95 eV in the case of the 16 AO/atom basis relative to the 4 AO/atom case, while Γ_{15} shifts down by only 2.38 eV. From that it is clear that the direct gap in our CCM calculation is only a basis set effect. Since increasing the cluster size increases also the basis for the extended states, a 64-atom [22] CCM calculation has been performed with the 16 AO/orbital basis set, using the SGI computer at the Supercomputer Center of the Florida State University. The only major change, relative to the 32-atom CCM, is a further improvement at the bottom of the CB. Since other values change by less than 30% of the change at X_{1c}, this proves that the latter
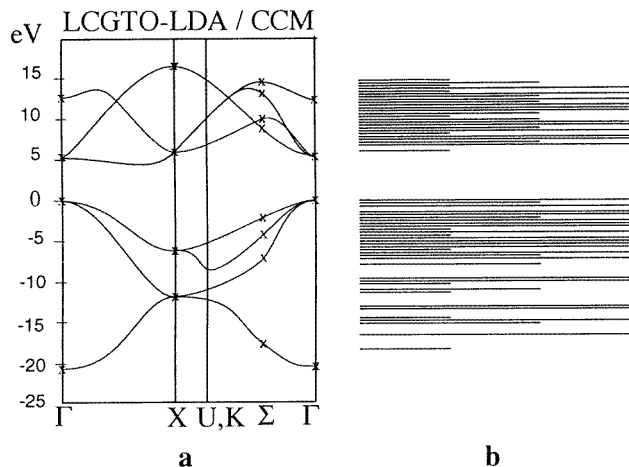
Figure 2. (a) The energies of one-electron states from the C_{32} CCM (denoted with $\times$) can be assigned to particular primitive $\boldsymbol{k}$ vectors (Γ, X, Σ). The bands are interpolated graphically to facilitate comparison with the band structure. (b) In the case of MCM the one-electron states cannot be assigned to states in the band structure. Due to the lower symmetry (T_{d}), only singly, doubly and triply degenerate states appear. The degeneracy is indicated by the length of the lines.

Table 2. Comparison of the CCM results to those of a full band structure calculation [4] and of experiments [21] for diamond. Calculations were performed for the experimental lattice constant, 3.57 Å.

Symm. point	Expt. [20] diamond	LDA			
		Full band struct. calc. [4] 40 AO/atom	C_{64} 16 AO/atom	C_{32} 16 AO/atom	C_{32} 4 AO/atom
Γ_1	-21.03	-21.68	-21.21	-21.12	-17.97
Γ'_{25}	0	0	0	0	0
Γ'_{15}	6.02	5.59	5.60	5.37	7.75
Γ'_2	13.41	13.21	13.55	13.85	14.94
X_{1v}	-12.43	-12.90	-12.66	-12.80	-11.4
X_{4v}	-6.27	-6.43	-6.36	-6.48	-6.33
X_{1c}	5.91	4.65	5.34	6.25	13.20
X_{4c}	—	16.87	16.49	16.50	17.13
L'_2	—	-15.79	-15.52	—	—
L_{1v}	-13.09	-13.73	-13.25	—	—
L'_3	-2.82	-2.86	-2.92	—	—
L_3	9.23	8.47	9.13	—	—
L_{1c}	9.58	8.90	9.46	—	—

depends primarily on the basis set.

In conclusion, we have shown that adapting the cyclic cluster model to the LCGTO–LDA method provides a framework for defect calculations which is more economic than full supercell or even molecular cluster model calculations, but still provides a good description of the extended states of the host while preserving the advantage of LCGTO expansions in reproducing localized states.

Acknowledgment

References

[1] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[2] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[3] Louie S A, Schlüter M, Chelikowsky J R and Cohen M L 1976 *Phys. Rev.* B **13** 1654
[4] Bachelet G B, Greenside H S, Baraff G A and Schlüter M 1981 *Phys. Rev.* B **24** 4745
[5] Briddon P R *PhD Thesis* University of Exeter
[6] Stoneham A M 1975 *Theory of Defects in Solids* (Oxford: Clarendon)
[7] Jones R 1992 *Phil. Trans. R. Soc.* A **67** 905
[8] Deák P and Snyder L C 1987 *Phys. Rev.* B **36** 9619
[9] Bennett A J, McCaroll M and Messmer R P 1971 *Phys. Rev.* B **3** 1397
[10] Zunger A 1974 *J. Phys. C: Solid State Phys.* **7** 76 and 96
     Zunger A 1975 *J. Chem. Phys.* **63** 1713
[11] Deák P, Kazsoki J and Giber J 1978 *Phys. Lett.* A **66** 395
[12] Evarestov R A 1975 *Phys. Status Solidi* b **72** 569
     Evarestov R A and Smirnov V P 1983 *Phys. Status Solidi* b **119** 9
[13] Chadi D J and Cohen M L 1973 *Phys. Rev.* B **8** 5747
[14] Evarestov R A and Lovchikov V A 1977 *Phys. Status Solidi* b **79** 743
     Harker A H and Larkins F P 1979 *J. Phys. C: Solid State Phys.* **12** 2487
     Harker A H and Larkins F P 1979 *J. Phys. C: Solid State Phys.* **12** 2497
     Smith J E, Symanszki P V and Matthew J A D 1985 *Philos. Mag.* B **51** 193
[15] Deák P 1981 *Acta Phys. Hung.* **50** 247
     Deák P 1981 *Phys. Lett.* **83A** 39
[16] Deák P, Snyder L C and Corbett J W 1988 *Phys. Rev.* B **37** 6887
[17] Deák P, Snyder L C and Corbett J W 1991 *Phys. Rev.* B **45** 747
[18] Ceperley D M and Alder B J 1980 *Phys. Rev. Lett.* **45** 566
[19] Porezag D, Frauenheim Th and Köhler Th 1995 *Phys. Rev.* B **51** 12 947
[20] Jones R and Briddon P R *Identification of Defects in Semiconductors (Semiconductors and Semimetals)* ed M Stavola (New York: Academic) at press
[21] Madelung O (ed) 1991 *Data in Science and Technology: Semiconductors (Group IV Elements and II–V Compounds)* (Springer)
[22] The 64-atom CCM is a $2 \times 2 \times 2$ multiple of the Bravais cell, representing the k_q set: $\Gamma[000]$, $\Delta[\frac{1}{2}00]$, X[100], W[$1\frac{1}{2}0$], $\Sigma[\frac{1}{2}\frac{1}{2}0]$, and L[$\frac{1}{2}\frac{1}{2}\frac{1}{2}$], which satisfies (1) up to $m_{max} = 8$.