Theor Chim Acta (1987) 72: 277-289

Theoretica Chimica Acta

Problems and prospects in the *ab initio* treatment of pure and defective crystals*

Cesare Pisani and Roberto Dovesi

Institute of Theoretical Chemistry, University of Torino, Via Giuria 5, I-10125 Torino, Italy

(Received 6 June/Accepted June 18, 1987)

Accurate Hartree-Fock LCAO calculations for moderately complex crystalline systems are now feasible; a number of important applications may be envisaged in the areas of material science and technology. Some critical aspects of the corresponding computer schemes are discussed which are of fundamental importance in determining the cost of the computation. Data are provided concerning actual computations which are indicative of the kind of periodic systems that can (or cannot) be treated at present. The result of a perfect-crystal *ab initio* HF study can be used as an input for treating with the same approximation local-defect problems, by use of suitable embedding techniques. A scheme of this kind is presented, and its computational implications are discussed: due to the intrinsic complexity of this problem, it may be foreseen that the study of defects in crystals will be a typical application of supercomputers in the area of quantum chemistry.

Key words: *Ab initio*-Hartree Fock — Perfect crystals — Defective crystals

1. Introduction

In the last ten years, a number of computer programs have been implemented for the *ab initio* calculation of total energy and ground state properties of crystalline systems. The development of increasingly powerful computers justifies this effort, since it makes it possible for such programs to gather new and valuable information on the chemical properties of condensed matter. Wider and wider classes of periodic systems, and a variety of problems can be studied that are of

* This paper was presented at the International Conference on 'The Impact of Supercomputers on Chemistry', held at the University of London, London, UK, 13-16 April 1987

interest from a technological or simply scientific point of view. As will be discussed in more detail in the following, having at hand the electronic wavefunction of the perfect crystal also allows local defects to be treated with the use of suitable embedding techniques.

A few recent *ab initio* computations for periodic systems may be cited that are indicative of the state of the art and of the areas of application:

(i) Dacarogna and Cohen [1] have studied the stability of lithium, sodium, and potassium in various crystal phases at different pressures. Total energies are calculated using an *ab initio* local-density (LD) pseudopotential method [2], and using as basis functions plane waves up to a cutoff energy of 5 or 7 a.u.

(ii) A very accurate study of the structural and electronic properties of beryllium has been performed by Blaha and Schwarz [3] by means of the linearized-augmented-plane-wave method [4]. A LD exchange-and-correlation potential is adopted; however, no shape approximations to the charge density or to the potential are made, and all electrons are taken into account.

(iii) The powerful and simple linear-muffin-tin-orbital method has been used by Lambrecht and Andersen [6] in conjunction with a LD approximation for studying the diamond-structure crystals C, Si, and Ge. It is shown that a minimal basis set is sufficient for an accurate calculation of a number of ground-state properties.

(iv) The Hartree-Fock (HF)-LCAO program CRYSTAL (developed at our Institute in collaboration with Saunders, of SERC laboratories [7, 8]) has been used by Dovesi et al. for studying the properties of MgO (100) surfaces, both bare and CO covered [9, 10]. A slab model was adopted, comprising up to four atomic layers; comparison with the perfect crystal results allowed the authors to test the validity of the model, and to calculate surface formation energies.

The examples above are illustrative of the variety of approaches that are currently adopted. In most cases, the Kohn-Sham [11] one-electron hamiltonian is used, with an LD approximation for the exchange and correlation effective potential. The local character of the approximate hamiltonian, and the inclusion of the correlation term in the self-consistent procedure makes these schemes to appear preferable to HF ones in the opinion of most solid state physicists. On the other hand, at variance with LD wavefunctions, the HF solution represents a well defined quantity with well understood characteristics; in principle, it may indefinitely be improved by using one of several schemes for removing the correlation error. Furthermore, if localized atomic orbitals (AO) are used as basis functions, one can exploit all the experience gathered in molecular quantum chemistry, and adopt the powerful algorithms developed in that context for the evaluation of one- and two-electron integrals [12]. For all these reasons, *ab initio* HF schemes are gaining importance in the quantum chemistry of condensed matter.

The purpose of this communication is threefold. First, we shall discuss some critical aspects of the CRYSTAL program in order to show that sophisticated and highly specific algorithms are needed to make such computations economically feasible. Second, computational data concerning a few selected

crystalline systems are commented upon, so as to give an idea of which kind of systems can (or cannot) be treated at present, and what impact supercomputers may be expected to have in this field. Third, we shall describe the essential characteristics of an *ab initio* HF embedding scheme for the treatment of local defects in crystals which is currently being implemented in our laboratory, and exploits much of the algebraic machinery embodied in CRYSTAL, plus, of course, the solution for the perfect host crystal obtained by means of that program.

2. The treatment of integral series in a LCAO-HF scheme for crystals

Figure 1 depicts the general mode of operation of CRYSTAL. The classification and evaluation of one- and two-electron integrals, and the calculation of the Fock, overlap, and density matrix are carried out in direct space; transformation to reciprocal space is performed only in order to obtain one-electron eigenvalues and eigenvectors, by carrying the Fock and overlap matrices to block-diagonal form. Four steps, labelled in Fig. 1 as α, β, γ, δ, characterize each stage of the self-consistent (SC) procedure:

α) *Reconstruction of the Fock matrix in real space*

Because of translational symmetry, a matrix element such as $\langle \mu \boldsymbol{m} | \mathscr{F} | \nu \boldsymbol{n} \rangle$, between the μth AO in the crystal cell identified by the lattice vector $\boldsymbol{m}$ and the νth AO in cell $\boldsymbol{n}$, depends only on μ, ν, and on the relative position $\boldsymbol{g} = \boldsymbol{m} - \boldsymbol{n}$ of the two

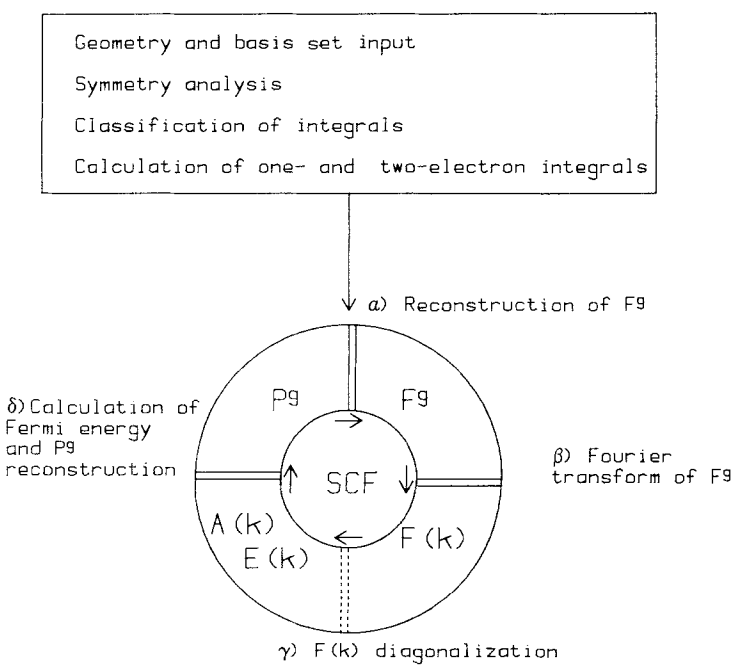

Fig. 1. Scheme of the CRYSTAL program for perfect crystals. The computational steps α, β, γ, δ of the SC stage are discussed in the text. The symbols included in the "SC wheel" are representative of the matrices that are obtained from the various computational steps: direct space submatrices $F^{\boldsymbol{g}}$, $P^{\boldsymbol{g}}$, and reciprocal space submatrices $F(\boldsymbol{k})$, $A(\boldsymbol{k})$, $E(\boldsymbol{k})$

cells. The general matrix element of the Fock (or overlap, or density) matrix may therefore be labelled as $F^g_{\mu\nu} = \langle \mu \mathbf{0} | \mathcal{F} | \nu \mathbf{g} \rangle$; each F^g submatrix is of order p, the number of AOs in each cell. Later in this section, indications are given concerning the algorithms adopted for treating efficiently this computational step.

β) Fourier transformation of F from direct to reciprocal space

For each $\boldsymbol{k}$ vector within the Brillouin zone (BZ), the Fock submatrix of order p: $F(k) = \sum_g F^g \exp(i\boldsymbol{k} \cdot \boldsymbol{g})$ is evaluated by direct summation.

γ) Diagonalization of the $F(\boldsymbol{k})$ submatrices

This is the only step where submatrices of order p are individually treated, all other steps requiring (in principle) a summation over an infinite number of other submatrices. By solving the matrix equation $F(\boldsymbol{k})\, A(\boldsymbol{k}) = S(\boldsymbol{k})\, A(\boldsymbol{k})\, E(\boldsymbol{k})$, one obtains the eigenvalues $E(\boldsymbol{k})$ and the eigenvectors $A(\boldsymbol{k})$ which define the crystalline orbitals exhibiting the translational symmetry properties described by the wavevector $\boldsymbol{k}$.

δ) Determination of the Fermi energy and reconstruction of the density submatrices P^l

The problem of self-consistently determining the Fermi energy ε_F, and of reconstructing the density submatrices in direct space:

$$P^l = 2 \int_{\mathrm{BZ}} d\boldsymbol{k}\, A(\boldsymbol{k})^* \theta(\varepsilon_F - \varepsilon(\boldsymbol{k})) \tilde{A}(\boldsymbol{k}) \exp(i\boldsymbol{k} \cdot \boldsymbol{l}) \tag{1}$$

is a standard one in solid physics (see for instance [13]). The choice of the sampling $\boldsymbol{k}$ points where the $F(\boldsymbol{k})$ matrices are evaluated (step β) and diagonalized (step γ) is strictly related to the techniques employed for performing the integrals over the BZ involved in this step.

As was stated in the introduction, it is not the purpose of this communication to analyze in detail all the steps of the computation. The characteristics and the general philosophy of the procedure are perhaps best recognized by closer examination of the way the F^g matrix is reconstructed in step α.

The expression for $F^g_{\mu\nu}$ is [7, 8]:

$$F^g_{\mu\nu} = T^g_{\mu\nu} + Z^g_{\mu\nu} + C^g_{\mu\nu} + X^g_{\mu\nu}, \tag{2}$$

where T and Z are the kinetic and nuclear attraction terms, while C and X are the Coulomb and exchange terms, to be determined self-consistently since they depend on the density matrix P (note that each F^g matrix depends in principle on *all* P^l submatrices!):

$$C^g_{\mu\nu} = \sum_{\lambda\rho} \sum_{l}^{\infty} P^l_{\lambda\rho} \sum_{h}^{\infty} (\mu \mathbf{0}\ \nu \boldsymbol{g} | \lambda \boldsymbol{h}\ \rho \boldsymbol{h} + \boldsymbol{l}) \tag{3}$$

$$X^g_{\mu\nu} = -0.5 \sum_{\lambda\rho} \sum_{l}^{\infty} P^l_{\lambda\rho} \sum_{h}^{\infty} (\mu \mathbf{0}\ \lambda \boldsymbol{h} | \nu \boldsymbol{g}\ \rho \boldsymbol{h} + \boldsymbol{l}). \tag{4}$$

A basis set of local functions (AO's) is adopted, each AO being described as a linear combination of gaussian type functions, so that standard quantum chemistry packages can be used for the evaluation of the integrals [12, 14, 15]. However, the infinite nature of the system and its high symmetry (translational and, in many cases, point symmetry) require and allow some new features with respect to molecular codes. Schematically, six points are worth mentioning, concerning the reconstruction of the Fock operator and the organization of the integral package:

(i) Use of the point symmetry of the Fock operator

Only an irreducible subset [16] $F^G_{\mu\nu}$ of matrix elements are explicitly evaluated by summation of the terms appearing in Eq. (1), the others being obtained by rotation.

(ii) Multipolar expansion of the shell charge distributions for the evaluation of long range Coulomb interactions

The $\boldsymbol{l}$ and $\boldsymbol{h}$ sums appearing in the Fock matrix (as well as the $\boldsymbol{g}$ one required for the Fourier transformation of the F matrix in step β and for the evaluation of total energy) extend in principle to the infinite set of translation vectors; in practice, the evaluation of the long range Coulomb interactions can be simplified according to the following scheme:

a) define Mulliken shell charge distributions:

$$\rho_s(\boldsymbol{r}) = \sum_{\lambda \in s} \sum_{\rho} \sum_{\boldsymbol{l}} P^{\boldsymbol{l}}_{\lambda\rho} \chi^{\boldsymbol{0}}_{\lambda}(\boldsymbol{r}) \chi^{\boldsymbol{l}}_{\rho}(\boldsymbol{r}) - z_s \delta(\boldsymbol{r} - \boldsymbol{r}_s). \tag{5}$$

b) reorder the Z and C contributions:

$$Z^{\boldsymbol{g}}_{\mu\nu} + C^{\boldsymbol{g}}_{\mu\nu} = \sum_{s} \sum_{\boldsymbol{h}} \int \chi^{\boldsymbol{0}}_{\mu}(\boldsymbol{r}) \chi^{\boldsymbol{g}}_{\nu}(\boldsymbol{r}) |\boldsymbol{r} - \boldsymbol{r}' - \boldsymbol{h}|^{-1} \rho^{\boldsymbol{h}}_{s}(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}'. \tag{6}$$

c) apart from a few $\boldsymbol{h}$ vectors (say 1 to 100, according to the cell size, dimensionality of the system, exponents of the involved AOs) for which all the bielectronic integrals are evaluated individually according to Eq. (3) ρ_s is external to $\chi^{\boldsymbol{0}}_{\mu}\, \chi^{\boldsymbol{g}}_{\nu}$: it can then be expanded in multipoles and the series can be evaluated "analytically" to infinity, using Ewald techniques [17] combined with recursion formulae of the kind proposed by McMurchie and Davidson [14, 8].

As a result, the long range Coulomb contributions are evaluated through three center integrals (in the case of bulk MgO and using a split valence basis set [18] more than 1000 terms contribute to each shell distribution ρ_s).

(iii) Truncation of the exchange series

Equation (4) shows that the value of the bielectronic integrals decay exponentially with increasing modulus of the vectors $\boldsymbol{h}$ and $\boldsymbol{h}+\boldsymbol{l}-\boldsymbol{g}$; at large distances the

leading term contributing to the Fock matrix is a "diagonal" one (that is $\boldsymbol{h}=\boldsymbol{0}$; $\boldsymbol{l}=\boldsymbol{g}$; $\lambda=\mu$; $\rho=\nu$):

$$X^{\boldsymbol{g}}_{\mu\nu} \simeq P^{\boldsymbol{g}}_{\mu\nu}(\mu\boldsymbol{0}\ \mu\boldsymbol{0}|\nu\boldsymbol{g}\ \nu\boldsymbol{g}). \tag{7}$$

The convergence of the $X^{\boldsymbol{g}}$ series is then depending on the long range behavior of the density matrix, which can be shown [19, 20] to decay exponentially in insulators and semiconductors; for metallic systems an oscillatory smoothed behavior is observed [20, 21].

When a simple truncation criterion is adopted, corresponding to including in the calculation only those $\boldsymbol{g}$ and $\boldsymbol{l}$ vectors such that $|\boldsymbol{g}|, |\boldsymbol{l}| \leqslant R^{\mathrm{ex}}$, a relatively rapid convergence of the series is obtained. For example, in the case of magnesium oxide (ionic, large gap), silicon (semiconductor, small gap) and aluminum (metallic) the truncation error on total energy is 0.00004, 0.0002 and 0.001 a.u./cell for $R^{\mathrm{ex}} = 10$ a.u.

(iv) On the storage of the integrals

According to points *ii* and *iii*, we must evaluate the bielectronc integrals of a sort of cluster (the "exact cluster"), whose shape and size depend on the involved distributions and on the required precision. When evaluating those integrals, we do not need to store them separately, but we can group together into symmetrized sums all those integrals that will be eventually multiplied by P factors that are translationally or rotationally equivalent:

$$B^{\boldsymbol{G}}_{\mu\nu} = \sum_{\lambda\rho} \sum_{\boldsymbol{L}} P^{\boldsymbol{L}}_{\lambda\rho} D^{\boldsymbol{GL}}_{\mu\nu\lambda\rho} \tag{8}$$

$$D^{\boldsymbol{GL}}_{\mu\nu\lambda\rho} = \sum_{\lambda'\rho'} \sum_{\boldsymbol{l}} T^{v}_{\lambda\lambda'} T^{v}_{\rho\rho'} d^{\boldsymbol{Gl}}_{\mu\nu\lambda'\rho'} \tag{9}$$

$$d^{\boldsymbol{Gl}}_{\mu\nu\lambda\rho} = \sum_{\boldsymbol{h}}^{\{a\}} (\mu\boldsymbol{0}\ \nu\boldsymbol{g}|\lambda\boldsymbol{h}\ \rho\boldsymbol{h}+\boldsymbol{l}) - 0.5 \sum_{\boldsymbol{h}}^{\{b\}} (\mu\boldsymbol{0}\ \lambda\boldsymbol{h}|\nu\boldsymbol{g}\ \rho\boldsymbol{h}+\boldsymbol{l}). \tag{10}$$

In Eqs. (8–10), capital bold letters indicate irreducible direct lattice vectors; in Eq. (9), $\boldsymbol{L}$ and $\boldsymbol{l}$, the irreducible and the general vector of a star, are related by the point symmetry operator v, whose representation in the basis of the atomic orbitals is T^{v}. The $\boldsymbol{h}$ summation in Eq. (10) extends to the vectors of the "exact cluster". Only the symmetrized D sums need to be stored and processesd at each SCF cycle. The saving factor in terms of I/O and external storage is a function of the order of the point group, and of the extension of the $\boldsymbol{h}$ sums in Eq. (10); in the most favourable cases (three-dimensional high-symmetry small unit cell systems) it can be as large as two orders of magnitude.

(v) Reordering the Coulomb sums

In the McMurchie-Davidson scheme, the two interacting distributions $(\mu\boldsymbol{0}\ \nu\boldsymbol{g})$ and $(\lambda\boldsymbol{h}\ \rho\boldsymbol{h}+\boldsymbol{l})$ (see eq. (10)) are first expanded in Hermite polynomials times gaussian functions (HGTF) centered in the centroids of the distributions; the integrals between HGTF are then evaluated through recursion relations (14, 15).

Taking advantage of this structure, the first term in Eq. (10) can be reordered as follows:

$$d^{GI}_{\mu\nu\lambda\rho}(\text{coul}) = \sum_{\boldsymbol{h}} \sum_{t,t'}^{\{a\}} E(12\boldsymbol{G}; t) E(34\boldsymbol{l}; t') I(t, t'; \boldsymbol{h})$$

$$= \sum_{t,t'} E(12\boldsymbol{G}; t) E(34\boldsymbol{l}; t') \sum_{\boldsymbol{h}}^{\{a\}} I(t, t'; \boldsymbol{h}), \tag{11}$$

where $E(t)$ is the tth expansion coefficient of the product of the two AOs in HGTF, and I is a Coulomb integral between HGTF. When the sum extends to many $\boldsymbol{h}$ vectors, the saving factor can be high because the expansion of the distributions is one of the rate determining steps of the calculation of the integrals.

(vi) The bipolar expansion

A large fraction of the bielectronic integrals of the "exact cluster" refer to non-overlapping distributions, and can be approximated with a bipolar expansion [22]:

$$(\mu\boldsymbol{0}\ \nu\boldsymbol{g}|\lambda\boldsymbol{h}\ \rho\boldsymbol{h}+\boldsymbol{l}) = \sum_{L,L'} \gamma(\mu\nu\boldsymbol{g}; L; \boldsymbol{r}) V(L, L'; \boldsymbol{r}-\boldsymbol{r}') \gamma(\lambda\rho\boldsymbol{l}; L'; \boldsymbol{r}'), \tag{12}$$

where $L \equiv (l, m)$ are the two quantum numbers characterizing the multipoles $\gamma(L)$ of the distribution $(\mu\boldsymbol{0}\ \nu\boldsymbol{g})$, evaluated with respect to the centroid $\boldsymbol{r}$, and $V(L, L')$ is the coupling operator [in particular, $V(0,0) = |\boldsymbol{r}-\boldsymbol{r}'|^{-1}$]. Multipoles can be evaluated through recursion relations, as shown in Ref. [8]. There are several advantages in using a bipolar expansion:

a) due to the translational and point symmetry, only a small set of multipoles need to be evaluated;

b) the order of the expansion can be chosen according to the importance of the integral and to the overall precision of the calculation;

c) the heavy step is a matrix multiplication, which is easily vectorialized;

d) in the case of the Coulomb sums, the same rearrangement as discussed in point *v* is possible; in this case the additional cost of an extra $\boldsymbol{h}$ vector in the sum is practically zero, because the evaluation of the coupling matrix V is very rapid.

The efficiency of the bipolar expansion technique, which is currently being implemented in the code, is documented in the next section.

As a final comment to the structure of CRYSTAL, we would like to observe that it appears very well suited for operation with parallel supercomputers, with relatively small rearrangements. All time consuming steps (calculation and manipulation of lists of integrals or matrix diagonalizations) require essentially the same computation to be performed independently for different $\boldsymbol{g}$ lattice vectors or $\boldsymbol{k}$ wavevectors.

3. Examples of application: "small" and "large" systems

Table 1 provides examples of "extreme" and "normal" applications of the CRYSTAL program.

The diamond calculation has been recently performed while trying to get as close as possible to the HF limit for subsequent application of correlation corrections; the basis set quality and the numerical accuracy have therefore been carried to the maximum allowed refinement. The number of two-electron integrals that need be calculated is enormous because of the computational conditions; on the other hand, use of the rich point symmetry as described under point *iv* in the preceding section, reduces to manageable proportions the number N' of symmetrized D sums to be stored (the overall saving factor is about 80). The computer time required for the self consistent calculation is negligible with respect to integral evaluation.

The polysulphur nitride calculation refers to current work intended to explore the importance of interchain interactions in crystalline conducting polymers. According to present standards it is again an extreme calculation, but in a different sense with respect to the preceding one: there are many atoms in the unit cell, some of them are second row ones, the point symmetry is poor. Two basis sets

Table 1. Examples of application of the CRYSTAL program. N^c, N^{ex}, and N' are the number of two-electron Coulomb and exchange integrals, and of symmetrized D sums, respectively, in 10^6 units. S^c and S^{ex} are the "overlap threshold" for the Coulomb and exchange series, that is the pseudocharge of the product distributions for the interacting electrons in Eqs. (3) and (4) below which the integral is disregarded. L is the maximum order in the multipolar expansion of the shell charge distribution (section 2, *ii*), R^{ex} is the "exchange radius" (section 2, *iii*). t are the computation times in seconds and refer to a Hitachi M200 scalar computer

	Diamond	3-d $(SN)_x$	MgO
Order of group	48	4	48
Atoms per cell	2	8	2
Basis set	Optimized double-zeta + d AOs	STO-3G [+d]	Optimized split-valence
AOs per cell	30	56 [76]	18
Comput. parameters			
$S^{ex} = S^c$	10^{-6}	10^{-6}	10^{-4}
R^{ex} (A)	6.9	5.8	5.8
L	4	4	4
Calculation data			
Part I (integral eval.)			
N^c	160	18 [60]	9
N^{ex}	150	21 [64]	7
N'	4	9 [30]	0.3
t	5500	700 [3700]	250
Part II (SCF stage)			
N_k	29	20	8
SCF cycles	14	9	10
t	200	370 [1100]	20

have been tried. In the first case, a minimal STO-3G set has been employed; in the second, the minimal set has been supplemented by d-type AOs on sulphur for describing important hypervalent aspects of the chemical bonds. In spite of the use of few sampling $\boldsymbol{k}$ points, the self consistent stage takes an appreciable fraction of overall computer time when the richer set is used, because of the large order (76) of the $F(\boldsymbol{k})$ matrices.

Magnesium oxide corresponds to a "normal" calculation of good quality [23]. An optimized split-valence set is used which gives a total energy near the HF limit; however, the number of two-electron integrals is not very large because the AOs are relatively short ranged. Furthermore, because of the rich point symmetry, there are only 200 000 symmetrized D sums to be stored and manipulated. The whole calculation takes only five minutes on a scalar computer.

Table 1 shows that the cost of the computation not only depends on the number of atoms in the unit cell and on the richness of the basis set, but also on the values adopted for the computational parameters which control the treatment of the Coulomb and exchange series and the reciprocal space integration (see section 2). The total energy error associated with those parameters is about 10^{-4} a.u./atom for diamond and $(SN)_x$ and 10^{-3} a.u./atom for MgO, that is four to six orders of magnitude larger than the "internal" or numerical error of standard molecular programs.

With respect to the situation documented in Table 1, work is in progress in three main directions:

a) extension of the performances and generalization of the program so as to be able to describe, at least with minimal basis sets, large unit cell systems (10–30 atoms). This would allow the *ab initio* description of many interesting silicates (at the moment studied with two body model potentials) and of complex surface phenomena in the slab approach;

b) deeper understanding of basis-set effects and careful analysis of the most critical algorithms, in order to be able to get near the Hartree–Fock limit for simple systems, while avoiding numerical instabilities;

c) improvement and speed-up of the program so as to obtain numerical accuracies in total energy to within 10^{-5}–10^{-6} a.u./atom with no additional costs with respect to the present situation. This should avoid the tiresome and time consuming step of the calibration of all the computational parameters, which is necessary at present for finding a compromise between accuracy and costs.

To give an idea of this work, Table 2 reports a few data on MgO concerning CPU time, number of bielectronic integrals to be computed, and convergence of total and kinetic energy toward the "exact" result. The main difference with respect to the procedure used for the results reported in Table 1 is that, whenever it is possible, the bielectronic integrals of the "exact" or "cluster" zone have been approximated using a bipolar expansion whose L_1 and L_2 orders are chosen according to the weighted distance between the two centroids, the amount of pseudocharge and the AO's quantum numbers involved. It appears that, with

Table 2. Dependence of total and kinetic energy of MgO bulk on the "severity" of the computational parameters. To simplify the discussion, only the overlap threshold S^c adopted for the Coulomb integrals is reported; the other parameters of the calculation have been set to values insuring about the same convergence as S^c. Basis set, N^c, N^{ex}, N' and t as in Table 1; $N^f \cdot 10^3$ is the number of matrix elements $F(\mu\nu\boldsymbol{G})$ to be defined. Energies in a.u. Numbers in parentheses are the bielectronic integrals approximated with the bipolar expansion. Last line refers to a calculation in which all the bielectronic integrals have been evaluated exactly

S^c	N^f	N^c	N^{ex}	N'	time	Total E.	Kinetic E.
10^{-3}	1.7	1.6	2.2	0.2	46	−274.643920	274.676342
		(0.6)	(0.9)				
10^{-4}	3.4	9	7	0.3	124	−274.664926	274.653463
		(5)	(4)				
10^{-5}	4.8	46	25	0.8	223	−274.664068	274.655232
		(38)	(21)				
10^{-6}	6.5	121	40	1.5	282	−274.664034	274.655389
		(110)	(35)				
10^{-7}	11.3	246	62	2.4	376	−274.664032	274.655388
		(231)	(56)				
10^{-7}	11.3	246	62	2.4	3740	−274.664034	274.655384

increasing severity of the truncation criteria, the number of integrals that can be approximated with low order expansions also increases, and the additional cost is progressively smaller. At least for this simple high-symmetry system, the target of a 10^{-6} a.u. precision on total energy is at hand.

4. The *ab initio* treatment of defective crystals

A number of embedding schemes have been proposed in past and recent years for the study of local defects in crystals (see for instance [24–27]). However, their application in conjunction with *ab initio* hamiltonians for describing the chemical properties of the defective region [26] still represent exceptional events rather than standard practice. In order to document the intrinsic complexity of this problem, we shall present here some aspects of an HF *ab initio* embedding program that is currently being implemented at our Institute. A "perturbed cluster" approach to the problem [27] is used, and the work equations [28] are a generalization of those previously used in conjunction with empirical or semi-empirical hamiltonians [29]. One starts from the definition of the system Green operator $\mathcal{G}(z)$:

$$\mathcal{Q}(z)\mathcal{G}(z) = \mathcal{I} \qquad [\mathcal{Q}(z) = z\mathcal{I} - \mathcal{F}] \tag{13}$$

($\mathcal{F}$ is the Fock operator), and represents this equation in an AO basis set, partitioned into a subset C that defines a cluster around the defect, and a complementary set D that defines the indented infinite crystal (that is, excluding the cluster region). Simple algebraic manipulations give the exact equations:

$$G^{CC} = \bar{G}^{CC} + \bar{G}^{CC} Q_{CD} G^{DD} Q_{DC} \bar{G}^{CC} \tag{14}$$

$$G^{CD} = -\bar{G}^{CC} Q_{CD} G^{DD}. \tag{15}$$

Here $G^{CC}(z)=[Q_{CC}(z)]^{-1}$ is the Green function for the "pseudo-isolated" cluster, that is, referring to a C system with no coupling elements with the indented crystal, but with a local hamiltonian F_{CC} which reflects the presence of the surrounding crystal. After introducing the sole assumption that the Green function in D is essentially unaltered by the presence of the defect:

$$G^{DD}(z)=G^{DDf}(z) \tag{16}$$

(the superscript f refers to unperturbed-host-crystal quantities), and integrating along a contour γ in the complex z plane encompassing all poles on the left of the Fermi energy ε_F, the following equations are obtained for the density matrix P:

$$\begin{aligned} P^{CC} &= \bar{P}^{CC}+\sum_j [A^C_{jD}[d/d\varepsilon M^{DD}(\bar{\varepsilon}_j)]\tilde{A}^C_{jD}+A^C_{jD}M^{DD}(\bar{\varepsilon}_j)\tilde{T}^C_{jD} \\ &\quad + T^C_{jD}M^{DD}(\bar{\varepsilon}_j)\tilde{A}^C_{jD}]+B^C_D P^{DDf}\tilde{B}^C_D \\ &\equiv P^{CC}_{\text{clust}}+P^{CC}_{\text{coupl}}+P^{CC}_{\text{ovlp}} \end{aligned} \tag{17}$$

$$P^{CD}=-\sum_j A^C_{jD}M^{DD}(\bar{\varepsilon}_j)-B^C_D P^{DDf}\equiv P^{CD}_{\text{coupl}}+P^{CD}_{\text{ovlp}}. \tag{18}$$

In these equations:

- $\bar{P}^{CC}\equiv P^{CC}_{\text{clust}}$ is the density matrix of the pseudo-isolated cluster, obtained using the eigenvectors $|\bar{v}_j\rangle$ and eigenvalues $\bar{\varepsilon}_j$ of the cluster hamiltonian F_{CC}, and limiting the occupied manifold to ε_F;

- the matrices A_j and T_j are defined in terms of F_{CC}, F_{CD}, S_{CC}, S_{CD} (Fock and overlap submatrices), and of $|\bar{v}_j\rangle$ and $\bar{\varepsilon}_j$ [27]; - the energy dependent "coupling matrix" $M^{DD}(e)$ is obtained from a knowledge of the projected density of states $\rho^f(\varepsilon)$ of the host defect-free crystal (this information comes from the solution for the perfect host crystal);

- $B^C_D=(S_{CC})^{-1}S_{CD}$ is zero if an orthogonal basis set is used.

Note (Eq. (17)) that the cluster density matrix P^{CC} is obtained by adding to the pseudo-isolated-cluster contribution P^{CC}_{clust} energy dependent terms (one per each eigenvalue $\bar{\varepsilon}_j$), and a constant term P^{CC}_{ovlp}; the latter is the only one that is left when a closed shell system is considered, treated with a minimal basis set, and disappears if an orthogonal basis set is adopted. Expression (18) for P^{CD} has a similar structure, except of course for the fact that no "zero" term may exist.

Figure 2 reproduces the general scheme of the computation. Comparison with Fig. 1 shows that in the present case all computations are performed in direct space. However, an "embedding step" (step δ) is now present, where information from the host crystal solution is used for realizing the coupling between the cluster and the indented crystal.

Let us consider again in some detail step α of the general SC cycle, where the submatrices F_{CC} and F_{CD} are reconstructed (the latter is needed for the evaluation of P^{CC} and P^{CD}, see above).

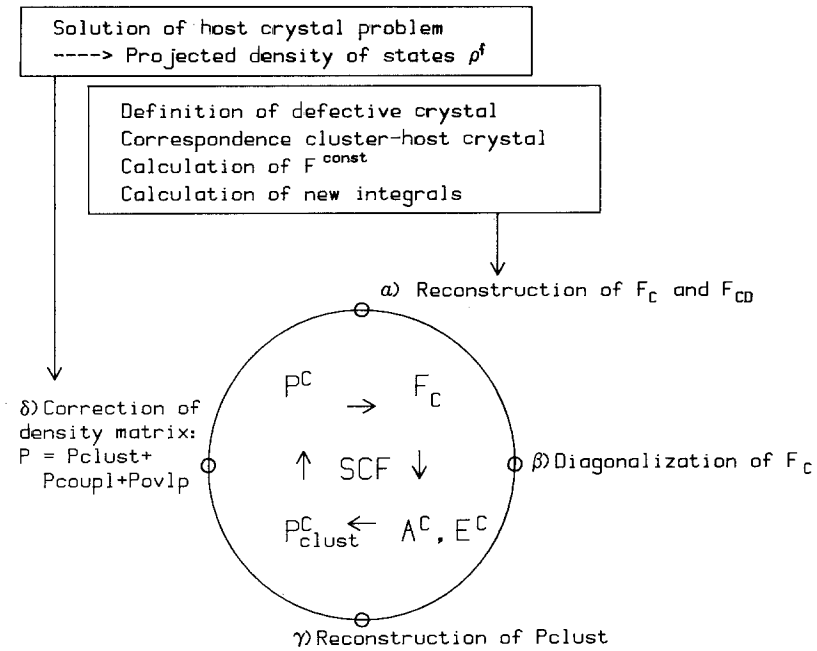

Fig. 2. Scheme of the EMBED program for defective crystals. The computational steps α, β, γ, δ of the SC stage are discussed in the text

The general element $F_{\gamma\mu}(\gamma \in C,\ \mu \in C$ or $D)$ may be expressed as a sum of a constant part and of a variable one that depends on the local part of the density matrix (P^{CC}, P^{CD}, and P^{DC}) and is therefore redefined during the SC procedure:

$$F_{\gamma\mu} = F^{\text{const}}_{\gamma\mu} + F^{\text{var}}_{\gamma\mu} \tag{19}$$

(note that the infinite submatrix P^{DD}, because of the fundamental assumption, Eq. (16), is fixed at the host-crystal value P^{DDf}). $F^{\text{const}}_{\gamma\mu}$ contains the kinetic term, the nuclear attraction terms, and the Coulomb and exchange interaction of the $(\gamma\mu)$ distribution with all $(\delta\delta')$ electron charge distributions associated with the indented crystal:

$$F^{\text{const}}_{\gamma\mu} = T_{\gamma\mu} + Z_{\gamma\mu} + \sum_{\delta\delta'} P^f_{\delta\delta'}[(\gamma\mu|\delta\delta') - \tfrac{1}{2}(\gamma\delta|\delta'\mu)] \tag{20}$$

The two-electron repulsion terms and the corresponding nuclear attraction terms may be combined together and dealt with using the same strategy as previously outlined for the perfect host crystal. Translational symmetry is however lost in the present case, partially at least: if either γ or μ refer to an impurity atom or to a crystal atom displaced with respect to its ideal site, the corresponding integrals must be evaluated afresh. Computation times may become very long if the proper defect region is relatively large. This difficulty is there, no matter which embedding scheme is adopted, if the field created in the defect region by the surrounding host crystal must accurately be evaluated.

As far as the SC-dependent part F^{var} is concerned, it must be observed that its F^{var}_{CD} part comprises in principle an infinite number of elements because of the infinite size of the D set. Again, efficient truncation criteria must be adopted, similar to those used in CRYSTAL. In spite of that, preliminary tests concerning

impurities in graphite show that thousands of matrix elements of F_{CD} must be self-consistently redefined (the tolerances adopted correspond to a numerical precision of a few millihartree/atom). It is evident that schemes relying on the hypothesis that the perturbative potential is localized in a strict vicinity of the defect are rather unrealistic.

Of course, throughout the computation advantage can be taken of the point symmetry that is left in the system after the defect is introduced: the problem may in fact be factored out into separate ones concerning the different irreducible representations of the point group [24].

The *ab initio* study of point defects in crystals is still in its infancy, and enormous progresses can be expected by skilful application of suitable algorithms, of the kind already operative in the treatment of perfect crystals. In any event, because of its intrinsic complexity, the problem will require huge computational effort: it can therefore be expected to represent a typical application of supercomputers in the years to come.

References

1. Dacarogna MM, Cohen ML (1986) Phys Rev B 34:4996
2. Louis SG, Froyen S, Cohen ML (1982) Phys Rev B 26:1738
3. Blaha P, Schwarz K (1987) J Phys F:899
4. Wimmer E, Krakauer H, Weinert M, Freeman AJ (1981) Phys Rev B 24:864
5. Andersen OK (1975) Phys Rev B 12:3060
6. Lambrecht WRL, Andersen OK (1986) Phys Rev B 34:2439
7. Pisani C, Dovesi R (1980) Int J Quantum Chem 17:501
8. Dovesi R, Pisani C, Roetti C, Saunders VR (1983) Phys Rev B 28:5781
9. Causà M, Dovesi R, Pisani C, Roetti C (1986) Surf Sci 175:551
10. Dovesi R, Orlando R, Ricca F, Roetti C (1987) Surf Sci 186:267
11. Kohn W, Sham LJ (1965) Phys Rev 140:A1133
12. Binkley JS, Whiteside RA, Krishnan R, Seeger R, DeFrees IJ, Schlegel HB, Topiol S, Kahn LR, Pople JA (1981) QCPE 13:406
13. Angonoa G, Dovesi R, Pisani C, Roetti C (1984) Phys Stat Sol 122:211
14. McMurchie LE, Davidson ER (1978) J Comput Phys 26:218; McMurchie LE, Davidson ER (1981) 44:289
15. Saunders VR (1983) In: Diercksen GHF, Wilson S (eds) Methods in computational molecular physics. Reidel, Dordrecht
16. Dovesi R (1986) Int J Quantum Chem 29:1755
17. Ewald P (1921) Ann Phys 64:253
18. Causà M, Dovesi R, Pisani C, Roetti C (1986) Phys Rev B 33:1308
19. Piela L, André JM, Fripiat JG, Delhalle J (1981) Chem Phys Lett 77:143
20. Causà M, Dovesi R, Orlando R, Pisani C, Saunders VR (1987) J Phys Chem, in press
21. Surjan PR, Kertész M, Karpfen A, Koller J (1983) Phys Rev B 27:7583
22. Buehler RJ, Hirschfelder JO (1952) Phys Rev 83:628
23. Causà M, Dovesi R, Pisani C, Roetti C (1986) Phys Rev B 33:1308
24. Bernholc J, Lipari NO, Pantelides ST (1980) Phys Rev B 21:3545
25. Inglesfield JE (1981) J Phys C 14:3795
26. Baraff GA, Schlüter M (1983) Phys Rev B 28:2296
27. Pisani C (1986) Philos Mag 51:89
28. Pisani C, Dovesi R, Kantorovich L: unpublished
29. Pisani C (1978) Phys Rev B 17:3143; Pisani C, Dovesi R, Carosso P (1979) Phys Rev B 20:5345; Pisani C, Dovesi R, Ugliengo P (1983) Phys Status Solidi B 116:249; Phys Status Solidi (1983) 116:547