CRYSTAL06

1.0 - v1_0_2

User's Manual

January 31, 2008

Main authors of CRYSTAL06:

R. Dovesi[1], V.R. Saunders[1,2], C. Roetti[1], R. Orlando[1,3],
C. M. Zicovich-Wilson [1,4], F. Pascale[5], B. Civalleri[1], K. Doll[6],
N.M. Harrison[2,7], I.J. Bush[2], Ph. D'Arco[8], M. Llunell[9]

[1] Theoretical Chemistry Group - University of Turin
Dipartimento di Chimica IFM
Via Giuria 5 - I 10125 Torino - Italy

[2] Computational Science & Engineering Department - CCLRC Daresbury
Daresbury, Warrington, Cheshire, UK WA4 4AD

[3] Università del Piemonte Orientale
Department of Science and Advanced Technologies
Via Bellini 25/6 - I 15100 Alessandria - Italy

[4] Departamento de Física, Universidad Autónoma del Estado de Morelos,
Av. Universidad 1001, Col. Chamilpa, 62210 Cuernavaca (Morelos) Mexico

[5] LCM3B - UMR 7036-CNRS
Université Henri Poincaré
Nancy BP 239, F54506 Vandœuvre-les-Nancy, Cedex, France

[6] Max-Planck-Institut für Festkörperforschung
Heisenbergstrasse 1
D-70569 Stuttgart, Germany

[7] Chemistry, Imperial College
South Kensington Campus
London, U.K.

[8] Laboratoire de Pétrologie, Modélisation des Matériaux et Processus
Université Pierre et Marie Curie,
4 Place Jussieu, 75232 Paris CEDEX 05, France

[9] Departament de Química Física, Universitat de Barcelona
Diagonal 647, Barcelona, Spain

Contents

Introduction

The CRYSTAL package performs *ab initio* calculations of the ground state energy, energy gradient, electronic wave function and properties of periodic systems. Hartree-Fock or Kohn-Sham Hamiltonians (that adopt an Exchange- Correlation potential following the postulates of Density-Functional theory) can be used. Systems periodic in 0 (molecules, 0D), 1 (polymers, 1D), 2 (slabs, 2D), and 3 dimensions (crystals, 3D) are treated on an equal footing. In each case the fundamental approximation made is the expansion of the single particle wave functions ('Crystalline Orbital', CO) as a linear combination of Bloch functions (BF) defined in terms of local functions (hereafter indicated as 'Atomic Orbitals', AOs). See Chapter 8.

The local functions are, in turn, linear combinations of Gaussian type functions (GTF) whose exponents and coefficients are defined by input (section 1.2). Functions of symmetry *s*, *p*, *d* and *f* can be used (see page 16). Also available are sp shells (s and p shells, sharing the same set of exponents). The use of sp shells can give rise to considerable savings in CPU time.

The program can automatically handle space symmetry: 230 space groups, 80 layer groups, 99 rod groups, 45 point groups are available (Appendix A). In the case of polymers it cannot treat helical structures (translation followed by a rotation around the periodic axis). However, when commensurate rotations are involved, a suitably large unit cell can be adopted.

Point symmetries *compatible with translation symmetry* are provided for molecules.

Input tools allow the generation of slabs (2D system) or clusters (0D system) from a 3D crystalline structure, the elastic distortion of the lattice, the creation of a super-cell with a defect and a large variety of structure editing. See Section 2.1

Previous releases of the software in 1988 (CRYSTAL88, [1]), 1992 (CRYSTAL92, [2]), 1996 (CRYSTAL95, [3]), 1998 (CRYSTAL98, [4]), and 2003 (CRYSTAL03, [5]) have been used in a wide variety of research with notable applications in studies of stability of minerals, oxide surface chemistry, and defects in ionic materials. See "Applications" in

http://www.crystal.unito.it

The CRYSTAL package has been developed over a number of years. For basic theory and algorithms see "Theory" in:

http://www.crystal.unito.it/theory.html

The required citation for this work is:

R. Dovesi, V.R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N.M. Harrison, I.J. Bush, Ph. D'Arco, M. Llunell
CRYSTAL06 User's Manual, University of Torino, Torino, 2006

CRYSTAL06 output will display the references relevant to the property computed, when necessary.
Updated information on the CRYSTAL code as well as tutorials to learn basic and advanced CRYSTAL usage are in:

http://www.crystal.unito.it

Functionality

The basic functionality of the code is outlined below.

The single particle potential	
Restricted Hartree-Fock Theory 2.3	
Unrestricted Open Shell Hartree-Fock Theory 2.3	
Density Functional Theory for Exchange and Correlation 2.3	
Spin Density Functional Theory 2.3	
Hybrids HF-DFT (B3LYP-B3PW) 2.3	
Effective Core Pseudo potentials 2.2	
Finite field perturbation added to the Hamiltonian 2.1	
Algorithms	
Parallel processing (replicated data) - See http://www.crystal.unito.it/Manuals/crystal03_P.pdf	
Massive Parallel Processing (distributed data)	
Traditional SCF	
Full Direct SCF 2.3	
Structural Editing	
Use of space, layer, rod and point group symmetry	
Deformation of the crystallographic cell 2.1	
Removal 2.1, insertion 2.1, substitution 2.1 of atoms	
Displacement of atoms 2.1	
Rotation of groups of atoms 2.1	
Extraction of surface models from a 3D crystal structure 2.1	
Cluster generation from a 3D crystal 2.1	
Cluster of molecules from molecular crystal 2.1	
Properties	
Band structure 5.2, Density of states 5.2	
Electronic charge density maps (2D, 3D grid) 5.2, 5.2	
Mulliken population analysis 2.3	
Spherical harmonic atom and shell multipoles 5.2	
X-ray structure factors 5.2	
Electron momentum distributions 5.2, 5.2	
Compton profiles 8.27	
First order density matrix	
Reciprocal form factors	
Electrostatic potential, field and field gradients 5.2, 5.2	
Spin polarized generalization of properties	
Hyperfine electron-nuclear spin tensor 5.2	
A posteriori Density Functional correlation energy 5.2	
Localization of Crystalline Orbitals 5.2	
Spontaneous polarization through Berry phase approach 5.3	
Spontaneous polarization through localized orbitals approach 5.3	
Piezoelectricity through Berry phase approach 5.3	
Piezoelectricity through localized orbitals approach 5.3	
Optical dielectric constant 5.2	
Analytic (nuclear coordinates and cell parameters) gradient of the energy	
Harmonic frequencies at Γ point	4
Geometry optimizer	(in cartesian and redundant internal coordinates) 3

Conventions

In the description of the input data which follows, the following notation is adopted:

-	•	new record
-	∗	free format record
-	*An*	alphanumeric datum (first *n* characters meaningful)
-	atom label	sequence number of a given atom in the primitive cell, as printed in the output file after reading of the geometry input
-	symmops	symmetry operators
-	▭, [ ]	default values.
-	*italic*	optional input

- ———————— *optional input records follow* ———————— II
- ———————— additional input records follow ———————— II

Part of the code is written in fortran 77. The name of the variables is associated with the type of data, following the fortran 77 convention: if the first letter of the name is I, J, K, L, M or N, the type is integer. Otherwise the type is real.

Arrays are read in with a simplified implied DO loop instruction of Fortran 77:
(dslist, i=m1,m2)
where: *dslist* is an input list; *i* is the name of an integer variable, whose value ranges from *m1* to *m2*.

Example (page 27): `LB(L),L=1,NL`
`NL` integer data are read in and stored in the first NL position of the array LB.

All the keywords are entered with an A format (case insensitive); the keywords must be typed left-justified, with no leading blanks.

---

conventional atomic number (usually called NAT) is used to associate a given basis set with an atom. The real atomic number is the remainder of the division NAT/100.

Output files may have a name assigned by the OPEN instruction, or the default name assigned to a fortran unit by the system. Mosts of the operating system assign the name *fort.fortran_unit_number*.

Acknowledgements

Embodied in the present code are elements of programs distributed by other groups.
In particular: the atomic SCF package of Roos et al. [6], the GAUSS70 gaussian integral package and STO-nG basis set due to Hehre et al. [7], the code of Burzlaff and Hountas for space group analysis [8] and Saunders' ATMOL gaussian integral package [9].
We take this opportunity to thank these authors. Our modifications of their programs have sometimes been considerable. Responsibility for any erroneous use of these programs therefore remains with the present authors.

It is our pleasure to thank Piero Ugliengo for continuous help, useful suggestions, rigorous testing.

We acknowledge Furio Corà careful testing of the development versions of CRYSTAL

Michele Catti significantly contributed to the implementation of the geometry optimizer with discussion, suggestions, contributions to the coding.

We are also in debt with Cesare Pisani, for his constant support of and interest in the development of the new version of the CRYSTAL program.

Financial support for this research has been provided by the italian MURST (Ministero della Università e della Ricerca Scientifica e Tecnologica), and the United Kingdom CCLRC (Council for the Central Laboratories of the Research Council).

Getting Started

Instructions to download, install, and run the code are available in the web site:
http://www.crystal.unito.it → documentation → README.

Program errors

A very large number of tests have been performed by researchers of a few laboratories, that had access to a test copy of CRYSTAL06. We tried to check as many options as possible, but not all the possible combinations of options have been checked. We have no doubts that errors remain.
The authors would greatly appreciate comments, suggestions and criticisms by the users of CRYSTAL; in case of *errors* the user is kindly requested to contact the authors, sending a copy of both input and output by E-mail to the Torino group (crystal@unito.it).

Chapter 1

Wave function calculation - Basic input route

1.1 Geometry and symmetry information

The first record of the geometry definition must contain one of the keywords:

CRYSTAL	3D system
SLAB	2D system
POLYMER	1D system
MOLECULE	0D system
EXTERNAL	geometry from external file
DLVINPUT	geometry from DLV [10] Graphical User Interface.

Three input schemes are used. The first is for crystalline systems, and is specified by the keyword **CRYSTAL**. The second is for slabs, polymers and molecules as specified by the keywords **SLAB**, **POLYMER** or **MOLECULE** respectively. In the third scheme, with keyword **EXTERNAL** or **DLVINPUT**, the unit cell, atomic positions and symmetry operators may be provided directly from an external file (see Appendix E, page 217). Such an input file can be prepared by the keyword **EXTPRT** (***crystal*** input block 1, page 32; ***properties***). Sample input decks for a number of structures are provided in section 6.1, page 147.

Geometry input for crystalline compounds

rec		variable	value	meaning
•	*	IFLAG		convention for space group identification (Appendix A.1, page 189):
			0	space group sequential number(1-230)
			1	Hermann-Mauguin alphanumeric code
		IFHR		type of cell for rhombohedral groups (meaningless for non-rhombohedral crystals):
			0	hexagonal cell
			1	rhombohedral cell
		IFSO		setting for the origin of the crystal reference frame:
			0	origin derived from the symbol of the space group: where there are two settings, the second setting of the International Tables is chosen.
			1	standard shift of the origin: when two settings are allowed, the first setting is chosen
			>1	non-standard shift of the origin given as input (see test22)
•	*			space group identification code (following IFLAG value):
		IGR		space group sequence number (IFLAG=0)
		or		
	A	AGR		space group alphanumeric symbol (IFLAG=1)
				if IFSO > 1 insert — II
•	*	IX,IY,IZ		non-standard shift of the origin coordinates (x,y,z) in fractions of the crystallographic cell lattice vectors times 24 (to obtain integer values).
•	*	a,[b],[c], [α],[β] [γ]		**minimal** set of crystallographic cell parameters: translation vector[s] length [Ångstrom], crystallographic angle[s] (degrees)
•	*	NATR		number of atoms in the asymmetric unit.
				insert NATR records — II
•	*	NAT		"conventional" atomic number. The conventional atomic number,NAT, is used to associate a given basis set with an atom. The real atomic number is the remainder of the division NAT100
		X,Y,Z		atom coordinates in fractional units of crystallographic lattice vectors
				optional keywords terminated by **END/ENDGEOM or STOP** — **II**

Geometry input for molecules, polymers and slabs

When the geometrical structure of 2D, 1D and 0D systems has to be defined, attention should be paid in the input of the atom coordinates, that are expressed in different units, fractionary (direction with translational symmetry) or Ångstrom (non periodic direction).

translational	unit of measure of coordinates		
symmetry	X	Y	Z
3D	fraction	fraction	fraction
2D	fraction	fraction	Ångstrom
1D	fraction	Ångstrom	Ångstrom
0D	Ångstrom	Ångstrom	Ångstrom

rec	variable	meaning
• ∗	IGR	point, rod or layer group of the system: 0D - molecules (Appendix A.4, page 196) 1D - polymers (Appendix A.3, page 193) 2D - slabs (Appendix A.2, page 192)
		if polymer or slab, insert II
• ∗	a,[b],	minimal set of lattice vector(s)- length in Ångstrom (b for rectangular lattices only)
	[γ]	$\widehat{AB}$ angle (degrees) - triclinic lattices only
• ∗	NATR	number of non-equivalent atoms in the asymmetric unit
		insert NATR records II
• ∗	NAT	conventional atomic number 1.1
	X,Y,Z	atoms coordinates. Unit of measure: 0D - molecules: *x,y,z* in Ångstrom 1D - polymers : *y,z* in Ångstrom, *x* in fractional units of crystallographic cell translation vector 2D - slabs : *z* in Ångstrom, *x, y* in fractional units of crystallographic cell translation vectors
		optional keywords terminated by **END or STOP** **II**

Geometry input from external geometry editor

The keywords **EXTERNAL** and **DLVINPUT** select the third input scheme. They work for molecules, polymers, slabs and crystals. The input data are read from file fort.34. The unit cell, atomic positions and symmetry operators are provided directly according to the format described in Appendix E, page 217. Coordinates in Ångstrom. Such an input file is written when **OPTGEOM** route for geometry optimization is chosen, and can be prepared by the keyword **EXTPRT** (program ***crystal***, input block 1, page 32; program ***properties***).
The geometry defined by **EXTERNAL** can be modified by inserting any geometry editing keyword (page 21) into the input stream after EXTERNAL.

Comments on geometry input

1. All coordinates in Ångstrom. In geometry editing, after the basic geometry definition, the unit of measure of coordinates may be modified by entering the keywords **FRACTION** (page 35) or **BOHR** (page 28).

2. The geometry of a system is defined by the crystal structure ([11], Chapter 1 of ref. [12]). Reference is made to the International Tables for Crystallography [13] for all definitions. The crystal structure is determined by the space group, by the shape and size of the unit cell and by the relative positions of the atoms in the asymmetric unit.

3. The lattice parameters represent the length of the edges of the cell (*a,b,c*) and the angles between the edges ($\alpha = \widehat{b\,c}$; $\beta = \widehat{a\,c}$; $\gamma = \widehat{a\,b}$). They determine the cell volume and shape.

4. Minimal set of lattice parameters to be defined in input:

cubic		a
hexagonal		a,c
rhombohedral	hexagonal cell	a,c
	rhombohedral cell	a, α
tetragonal		a,c
orthorhombic		a,b,c
monoclinic		a,b,c, β (b unique)
		a,b,c, γ (c unique)

	a,b,c, α (a unique - non standard)
triclinic	a,b,c, α, β, γ

5. The asymmetric unit is the largest subset of atoms contained in the unit-cell, where no atom pair related by a symmetry operator can be found. Usually several equivalent subsets of this kind may be chosen so that the asymmetric unit needs not be unique. The asymmetric unit of a space group is a part of space from which, by application of all symmetry operations of the space group, the whole of space is filled exactly.

6. The *crystallographic*, or *conventional cell*, is used as the standard option in input. It may be non-primitive, which means it may not coincide with the cell of minimum volume (*primitive cell*), which contains just one lattice point. The matrices which transform the conventional (as given in input) to the primitive cell (used by **CRYSTAL**) are given in Appendix A.5, page 197, and are taken from Table 5.1 of the International Tables for Crystallography [13].

   Examples. A cell belonging to the face-centred cubic Bravais lattice has a volume four times larger than that of the corresponding primitive cell, and contains four lattice points (see page 44, keyword **SUPERCEL**). A unit cell belonging to the hexagonal Bravais lattice has a volume three times larger than that of the rhombohedral primitive cell (R Bravais lattice), and contains three lattice points.

7. The use of the International Tables to identify the symmetry groups requires some practice. The examples given below may serve as a guide. The printout of geometry information (equivalent atoms, fractional and Cartesian atomic coordinates etc.) allows a check on the correctness of the group selected. To obtain a complete neighborhood analysis for all the non-equivalent atoms, a complete input deck must be read in (blocks 1-3), and the keyword **TESTPDIM** inserted in block 3, to stop execution after the symmetry analysis.

8. Different settings of the origin may correspond to a different number of symmetry operators with translational components.

   Example: bulk silicon - Space group 227 - 1 irreducible atom per cell.

setting of the origin	Si coordinates	symmops with translational component	multiplicity
2nd (default)	1/8 1/8 1/8	36	2
1st	0. 0. 0.	24	2

   NB With 2nd setting, the position (0., 0., 0.) has multiplicity 4.

   The choice is important when generating a supercell, as the first step is the removal of the symmops with translational component. The keyword **ORIGIN** (input block 1, page 39) translates the origin in order to minimize the number of symmops with translational component.

9. When coordinates are obtained from experimental data or from geometry optimization with semi-classical methods, atoms in special positions, or related by symmetry are not correctly identified, as the number of significative digits is lower that the one used by the program ***crystal*** to recognize equivalence or special positions. In that case the coordinates must be edited by hand (see FAQ at www.crystal.unito.it).

10. The symbol of the space group for crystals (IFLAG=1) is given precisely as it appears in the International Tables, *with the first letter in column one* and a blank separating operators referring to different symmetry directions. The symbols to be used for the groups 221-230 correspond to the convention adopted in editions of the International Tables prior to 1983: the 3 axis is used instead of $\bar{3}$. See Appendix A.1, page 189.

    Examples:

Group number	input symbol	
137 (tetragonal)	`P␣42/N␣M␣C`	
10 (monoclinic)	`P␣1␣2/M␣1`	(unique axis b, standard setting)
	`P␣1␣1␣2/M`	(unique axis c)
	`P␣2/M␣1␣1`	(unique axis a)
25 (orthorhombic)	`P␣M␣M␣2`	(standard setting)
	`P␣2␣M␣M`	
	`P␣M␣2␣M`	

11. In the monoclinic and orthorhombic cases, if the group is identified by its number (3-74), the conventional setting for the unique axis is adopted. The explicit symbol must be used in order to define an alternative setting.

12. For the centred lattices (F, I, C, A, B and R) the input cell parameters refer to the centred conventional cell; the fractional coordinates of the input list of atoms are in a vector basis relative to the centred conventional cell.

13. It is sufficient to supply the coordinates of only *one* of a group of atoms equivalent under centring translations (eg: for space group Fm3m only the parameters of the face-centred cubic cell, and the coordinates of one of the four atoms at (0,0,0), (0, $\frac{1}{2}$,$\frac{1}{2}$), ($\frac{1}{2}$,0,$\frac{1}{2}$) and ($\frac{1}{2}$,$\frac{1}{2}$,0) are required).

    The coordinates of only one atom among the set of atoms linked by centring translations are printed. The vector basis is relative to the centred conventional cell. However when Cartesian components of the direct lattice vectors are printed, they are those of the primitive cell.

14. The conventional atomic number NAT is used to associate a given basis set with an atom (see Basis Set input, Section 1.2, page 14). The real atomic number is given by the remainder of the division of the conventional atomic number by 100 (Example: NAT=237, Z=37; NAT=128, Z=28). Atoms with the same atomic number, but in non-equivalent positions, can be associated with different basis sets, by using different conventional atomic numbers: e.g. 6, 106, 1006 (all electron basis set for carbon atom); 206, 306 (core pseudo-potential for carbon atom, Section 2.2, page 50).

    If the remainder of the division is 0, a "ghost" atom is identified, to which no nuclear charge corresponds (it may have electronic charge). This option may be used for enriching the basis set by adding bond basis function [14], or to allow build up of charge density on a vacancy. A given atom may be transformed into a ghost after the basis set definition (input block 2, keyword **GHOSTS**, page 49).

15. The keyword **SLABCUT** (Geometry editing input, page 42) allows the creation of a slab (2D) of given thickness from the 3D perfect lattice. See for comparison test4-test24; test5-test25; test6-test26; test7- test27.

16. For slabs (2D), when two settings of the origin are indicated in the International Tables for Crystallography, setting number 2 is chosen. The setting can not be modified.

17. Conventional orientation of slabs and polymers: Polymers are oriented along the x axis. Slabs are parallel to the xy plane.

18. The keywords **MOLECULE** (for molecular crystals only; page 37) and **CLUSTER** (for any n-D structure; page 29) allow the creation of a non-periodic system (molecule(s) or cluster) from a periodic one.

1.2 Basis set

rec	variable	value	meaning
• ∗	NAT	n	conventional atomic 1.1 number
		<200;> 1000	all-electron basis set
		>200	valence electron basis set. ECP (Effective Core Pseudopotential) must be defined (page 50)
		=99	end of basis set input section
	NSHL	n	number of shells
		0	end of basis set input (when NAT=99)
			if NAT > 200 insert ECP input (page 50) II
			NSHL sets of records - for each shell
• ∗	ITYB		type of basis set to be used for the specified shell:
		0	general BS, given as input
		1	Pople standard STO-nG (Z=1-54)
		2	Pople standard 3(6)-21G (Z=1-54(18)) Standard polarization functions are included.
	LAT		shell type:
		0	1 s AO (S shell)
		1	1 s + 3 p AOs (SP shell)
		2	3 p AOs (P shell)
		3	5 d AOs (D shell)
		4	7 f AOs (D shell) - polarization only
	NG		Number of primitive Gaussian Type Functions (GTF) in the contraction for the basis functions (AO) in the shell
		1≤NG≤10	for ITYB=0 and LAT ≤ 2
		1≤NG≤6	for ITYB=0 and LAT = 3
		2≤NG≤6	for ITYB=1
		6	6-21G core shell
		3	3-21G core shell
		2	n-21G inner valence shell
		1	n-21G outer valence shell
	CHE		formal electron charge attributed to the shell
	SCAL		scale factor (if ITYB=1 and SCAL=0., the standard Pople scale factor is used for a STO-nG basis set.
			if ITYB=0 (general basis set insert NG records II
• ∗	EXP		exponent of the normalized primitive GTF
	COE1		contraction coefficient of the normalized primitive GTF:
			LAT=0,1 → s function coefficient
			LAT=2 → p function coefficient
			LAT=3 → d function coefficient
			LAT=4 → f function coefficient
	COE2		LAT=1 → p function coefficient
			optional keywords terminated by **END/ENDB** or **STOP** II

The choice of basis set is the most critical step in performing *ab initio* calculations of periodic systems, with Hartree-Fock or Kohn-Sham Hamiltonians. Optimization criteria are discussed in Chapter 5.2. When an effective core pseudo-potential is used, the basis set **must** be optimized with reference to that potential (Section 2.2, page 50).

1. A basis set (BS) must be given for each atom with different conventional atomic number defined in the crystal structure input. If atoms are removed (geometry input, keyword **ATOMREMO**, page 27), the corresponding basis set input can remain in the input stream.

2. The basis set for each atom has NSHL shells, whose constituent AO basis functions are built from a linear combination ('contraction') of individually normalized primitive Gaussian-type functions (GTF) (Chapter 8, page 174).

3. A conventional *atomic number* NAT links the basis set with the atoms defined in the crystal structure. The atomic number Z is given by the remainder of the division of the conventional atomic number by 100 (Example: NAT=108, Z=8, all electron; NAT=228, Z=28, ECP). See point 5 below.

4. A conventional atomic number 0 defines ghost atoms, that is points in space with an associated basis set, but lacking a nuclear charge (vacancy). See test 28.

5. Atoms with equal conventional atomic number are associated with the same basis set.

NAT < 200;>1000:	all electron basis set. A maximum of two different basis sets may be given for the same chemical species in different positions: NAT=Z, NAT=Z+100.
NAT > 200:	valence electron basis set. A maximum of two different BS may be given for the same chemical species in positions not symmetry-related: NAT=Z+200, NAT=Z+300. A core pseudo-potential must be defined. See Section 2.2, page 50, for information on core pseudo-potentials.

   Suppose we have four non-equivalent carbon atoms in the unit cell. Conventional atomic numbers 6 106 206 306 mean that carbon atoms (real atomic number 6) unrelated by symmetry are to be associated with different basis sets: the first two (6, 106) all-electron, the second two (206, 306) valence only, with pseudo-potential.

6. The basis set input ends with the card:
   99 0 conventional atomic number 99, 0 shell.
   The optional keywords may follow.

In summary:

1. *CRYSTAL* can use the following all electrons basis sets:

   a) general basis sets, including *s, p, d, f* functions (given in input);
   b) standard Pople basis sets [15] (internally stored as in Gaussian 94 [16]).
      STOnG, Z=1 to 54
      6-21G, Z=1 to 18
      3-21G, Z=1 to 54

   The standard basis sets b) are stored as internal data in the *CRYSTAL* code. They are all electron basis sets, and can not be combined with ECP.

2. **Warning** The standard scale factor is used for STO-nG basis set when the input datum SCAL is 0.0 in basis set input. *All the atoms of the same row are attributed the same Pople STO-nG basis set when the input scale factor SCAL is 1.*

3. Standard polarization functions can be added to 6(3)-21G basis sets of atoms up to Z=18, by inserting a record describing the polarization shell (ITYB=2, LAT=2, p functions on hydrogen, or LAT=3, d functions on 2-nd row atoms; see test 12).

```
  H          Polarization functions exponents                He
 1.1                                                         1.1
----------                        ------------------------------
  Li    Be                          B    C    N    O    F    Ne
 0.8   0.8                         0.8  0.8  0.8  0.8  0.8   --
-----------                       ------------------------------
  Na    Mg                          Al   Si   P    S    Cl   Ar
0.175 0.175                       0.325 0.45 0.55 0.65 0.75 0.85
----------------------------------------------------------------
```

   The formal electron charge attributed to a polarization function must be zero.

4. The shell types available are :

shell code	shell type	n. AO	order of internal storage
0	S	1	s
1	SP	4	s, x, y, z
2	P	3	x, y, z
3	D	5	$2z^2 - x^2 - y^2, xz, yz, x^2 - y^2, xy$
4	F	7	$(2z^2 - 3x^2 - 3y^2)z, (4z^2 - x^2 - y^2)x, (4z^2 - x^2 - y^2)y,$ $(x^2 - y^2)z, xyz, (x^2 - 3y^2)x, (3x^2 - y^2)y$

F shells can be used as polarization functions only. Wave function for atoms with f orbitals partially occupied can not be computed.

When symmetry adaptation of Bloch functions is active (default; NOSYMADA in block3 to remove), if F functions are used, all lower order functions must be present (D, P, S).

The order of internal storage of the AO basis functions is an information necessary to read certain quantities calculated by the program ***properties***. See Chapter 5: Mulliken population analysis (**PPAN**, page 74), electrostatic multipoles (**POLI**, page 135), projected density of states (**DOSS**,page 118) and to provide an input for some options (**EIGSHIFT**, input block 3, page 66).

5. Spherical harmonics d-shells consisting of 5 AOs are used.

6. Spherical harmonics f-shells consisting of 7 AOs are used.

7. The formal shell charges CHE, the number of electrons attributed to each shell, are assigned to the AO following the rules:

shell code	shell type	max CHE	rule to assign the shell charges
0	S	2.	CHE for S functions
1	SP	8.	if CHE>2, 2 for S and (CHE−2) for P functions, if CHE≤2, CHE for S function
2	P	6.	CHE for P functions
3	D	10.	CHE for D functions
4	F	14.	CHE for F functions - it must be 0. in CRYSTAL06.V1.0

8. A maximum of one open shell for each of the s, p and or d atomic symmetries is allowed in the electronic configuration defined in the input. The atomic energy expression is not correct for all possible double open shell couplings of the form $p^m d^n$. Either m must equal 3 or n must equal 5 for a correct energy expression in such cases. A warning will be printed if this is the case. However, the resultant wave function (which is a superposition of atomic densities) will usually provide a reasonable starting point for the periodic density matrix.

9. When extended basis sets are used, all the functions corresponding to symmetries (angular quantum numbers) occupied in the isolated atom are added to the atomic basis set for atomic wave function calculations, even if the formal charge attributed to that shell is zero. Polarization functions are not included in the atomic basis set; *their input occupation number should be zero.*

10. The formal shell charges are used only to define the electronic configuration of the atoms to compute the atomic wave function. The initial density matrix in the SCF step may be a superposition of atomic (or ionic) density matrices (default option, **GUESSPAT**, page 70). When a different guess is required (**GUESSF** or **GUESSP**), the shell charges are not used, but checked for electron neutrality when the basis set is entered.

11. Each atom in the cell may have an ionic configuration, when the sum of formal shell charges (CHE) is different from the nuclear charge. When the number of electrons in

the cell, that is the sum of the shell charges CHE of all the atoms, is different from the sum of nuclear charges, the reference cell is non-neutral. This is not allowed for periodic systems, and in that case the program stops. In order to remove this constraint, it is necessary to introduce a uniform charged background of opposite sign to neutralize the system [17]. This is obtained by entering the keyword **CHARGED** (page 47) after the standard basis set input. The value of total energy must be carefully checked.

12. It may be useful to allow atoms with the same basis set to have different electronic configurations (e.g, for an oxygen vacancy in MgO one could use the same basis set for all the oxygens, but begin with different electronic configuration for those around the vacancy). The formal shell charges attributed in the basis set input may be modified for selected atoms by inserting the keyword **CHEMOD** (input block 2, page 47).

13. The energies given by an atomic wave function calculation with a crystalline basis set should not be used as a reference to calculate the formation energies of crystals. The external shells should first be re-optimized in the isolated atom by adding a low-exponent Gaussian function, in order to provide and adequate description of the tails of the isolated atom charge density [18] (keyword **ATOMHF**, input block 3, page 56).

Optimized basis sets for periodic systems used in published papers are available in:

http://www.crystal.unito.it

1.3 Computational parameters, hamiltonian, SCF control

Default values are set for all computational parameters. Default choices may be modified through keywords. Default choices:

	default	keyword to modify default
hamiltonian:	RHF	UHF, DFT (SPIN) 2.3
tolerances for coulomb and exchange sums :	6 6 6 6 12	TOLINTEG 2.3
Pole order for multipolar expansion:	4	POLEORDR
Max number of SCF cycles:	50	MAXCYCLE 2.3
Convergence on total energy:	10^{-6}	TOLDEE 2.3

For periodic systems, 1D, 2D, 3D, the only *mandatory input information* is the shrinking factor, IS, to generate a commensurate grid of **k** points in reciprocal space, according to Pack-Monkhorst method [19]. The Hamiltonian matrix computed in direct space, $H_{\mathbf{g}}$, is Fourier transformed for each **k** value, and diagonalized, to obtain eigenvectors and eigenvalues:

$$H_k = \sum_g H_g e^{i\mathbf{gk}}$$

$$H_k A_k = Sk A_k E_k$$

A second shrinking factor, ISP, defines the sampling of k points, "Gilat net" [20, 21], used for the calculation of the density matrix and the determination of Fermi energy in the case of conductors (bands not fully occupied).
The two shrinking factors are entered after the keyword **SHRINK** (page 75).
In 3D crystals, the sampling points belong to a lattice (called the Pack-Monkhorst net), with basis vectors:

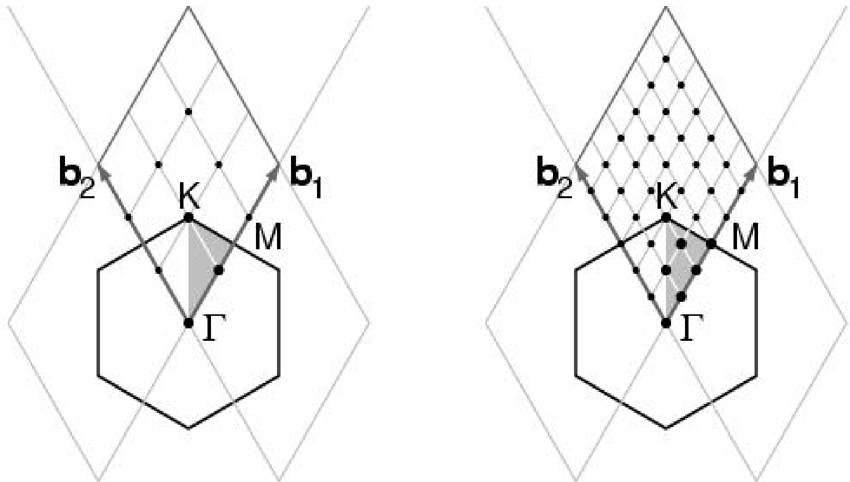
b1/is1, b2/is2, b3/is3 is1=is2=is3=IS, unless otherwise stated

where b1, b2, b3 are the reciprocal lattice vectors, and is1, is2, is3 are integers "shrinking factors".
In 2D crystals, IS3 is set equal to 1; in 1D crystals both IS2 and IS3 are set equal to 1. Only points k_i of the Pack-Monkhorst net belonging to the irreducible part of the Brillouin Zone (IBZ) are considered, with associated a geometrical weight, w_i. The choice of the reciprocal space integration parameters to compute the Fermi energy is a delicate step for metals. See Section 8.7, page 180. Two parameters control the accuracy of reciprocal space integration for Fermi energy calculation and density matrix reconstruction:

IS shrinking factor of reciprocal lattice vectors. The value of IS determines the number of **k** points at which the Fock/KS matrix is diagonalized. Multiples of 2 or 3 should be used, according to the point symmetry of the system (order of principal axes).

In high symmetry systems, it is convenient to assign IS *magic* values such that all low multiplicity (high symmetry) points belong to the Monkhorst lattice. Although this choice does not correspond to maximum efficiency, it gives a safer estimate of the integral.

The **k***-points net is automatically made anisotropic for 1D and 2D systems.*



The figure presents the reciprocal lattice cell of 2D graphite (rhombus), the first Brillouin zone (hexagon), the irreducible part of Brillouin zone (in grey), and the coordinates of the $\mathbf{k}_i$ points according to a Pack-Monkhorst sampling, with shrinking factor 3 and 6.

ISP shrinking factor of reciprocal lattice vectors in the Gilat net (see [22], Chapter II.6). ISP is used in the calculation of the Fermi energy and density matrix. Its value can be equal to IS for insulating systems and equal to 2*IS for conducting systems.

The value assigned to ISP is irrelevant for non-conductors. However, a non-conductor may give rise to a conducting structure at the initial stages of the SCF cycle (very often with DFT hamiltonians), owing, for instance, to a very unbalanced initial guess of the density matrix. The ISP parameter must therefore be defined in all cases.
Note. The value used in the calculation is ISP=IS*NINT(MAX(ISP,IS)/IS)

In the following table the number of sampling points in the IBZ and in BZ is given for a fcc lattice (space group 225, 48 symmetry operators) and hcp lattice (space group 194, 24 symmetry operators). The CRYSTAL code allows 413 k points in the Pack-Monkhorst net, and 2920 in the Gilat net.

IS	points in IBZ fcc	points in IBZ hcp	points BZ
6	16	28	112
8	29	50	260
12	72	133	868
16	145	270	2052
18	195	370	2920
24	413	793	6916
32	897	1734	16388
36	1240	2413	23332
48	2769	5425	55300

1. When an anisotropic net is user defined (IS=0), the ISP input value is taken as ISP1 (shrinking factor of Gilat net along first reciprocal lattice) and ISP2 and ISP3 are set to:
   ISP2=(ISP*IS2)/IS1,
   ISP3=(ISP*IS3)/IS1.

2. User defined anisotropic net is not compatible with SABF (Symmetry Adapted Bloch Functions). See **NOSYMADA**, page 73.

Some tools for accelerating convergence are given through the keywords **LEVSHIFT** (page 71 and tests 29, 30, 31, 32, 38), **FMIXING** (page 69), **SMEAR** (page 77), **BROYDEN** (page 58) and **ANDERSON** (page 56).

At each SCF cycle the total atomic charges, following a Mulliken population analysis scheme, and the total energy are printed.
The default value of the parameters to control the exit from the SCF cycle ($\Delta E < 10^{-6}$ hartree, maximum number of SCF cycles: 50) may be modified entering the keywords:
TOLDEE (tolerance on change in total energy) page 79;
TOLDEP (tolerance on SQM in density matrix elements) page 79;
MAXCYCLE (maximum number of cycles) page 72.

Spin-polarized system

By default the orbital occupancies are controlled according to the 'Aufbau' principle.
To obtain a spin polarized solution an open shell Hamiltonian must be defined (block3, **UHF** or **DFT/SPIN**). A spin-polarized solution may then be computed after definition of the (α-β) electron occupancy. This can be performed by the keywords **SPINLOCK** (page 78) and **BETALOCK** (page 57).

Chapter 2

Wave function calculation - Advanced input route

2.1 Geometry editing

The following keywords allow editing of the crystal structure, printing of extended information, generation of input data for visualization programs. Processing of the input block 1 only (geometry input) is allowed by the keyword **TESTGEOM**.
Each keyword operates on the geometry active when the keyword is entered. For instance, when a 2D structure is generated from a 3D one through the keyword **SLABCUT**, all subsequent geometry editing operates on the 2D structure. When a dimer is extracted from a molecular crystal through the keyword **MOLECULE**, all subsequent editing refers to a system without translational symmetry.
The keywords can be entered in any order: particular attention should be paid to the action of the keywords **KEEPSYMM** 2.1 and **BREAKSYM** 2.1, that allow maintaining or breaking the symmetry while editing the structure. These keywords behave as a switch, and require no further data. Under control of the **BREAKSYM** keyword (the default), subsequent modifications of the geometry are allowed to alter (reduce: the number of symmetry operators cannot be increased) the point-group symmetry. The new group is a subgroup of the original group and is automatically obtained by **CRYSTAL**. However if a **KEEPSYMM** keyword is presented, the program will endeavor to maintain the number of symmetry operators, by requiring that atoms which are symmetry related remain so after a geometry editing (keywords: **ATOMSUBS**, **ATOMINSE**, **ATOMDISP**, **ATOMREMO**).
The space group of the system may be modified after editing. For 3D systems, the file FINDSYM.DAT may be written (keyword **FINDSYM**). This file is input to the program *findsym* (http://physics.byu.edu/ stokesh/isotropy.html), that finds the space-group symmetry of a crystal, given the coordinates of the atoms.

Geometry keywords

Symmetry information			
ATOMSYMM	printing of point symmetry at the atomic positions	28	–
MAKESAED	printing of symmetry allowed elastic distortions (SAED)	36	–
PRSYMDIR	printing of displacement directions allowed by symmetry.	40	–
SYMMDIR	printing of symmetry allowed geom opt directions	45	–
SYMMOPS	printing of point symmetry operators	46	–
TENSOR	tensor of physical properties	46	I

Symmetry information and control

BREAKSYM	allow symmetry reduction following geometry modifications	29	–
KEEPSYMM	maintain symmetry following geometry modifications	36	–
MODISYMM	removal of selected symmetry operators	36	I
PURIFY	cleans atomic positions so that they are fully consistent with the group	40	–
SYMMREMO	removal of all symmetry operators	46	–
TRASREMO	removal of symmetry operators with translational components	46	–

Modifications without reduction of symmetry

ATOMORDE	reordering of atoms in molecular crystals	26	–
NOSHIFT	no shift of the origin to minimize the number of symmops with translational components before generating supercell	39	–
ORIGIN	shift of the origin to minimize the number of symmetry operators with translational components	39	–
PRIMITIV	crystallographic cell forced to be the primitive cell	40	–
SLABINFO	definition of a new cell, with $xy \parallel$ to a given plane	43	I

Atoms and cell manipulation (possible symmetry reduction (BREAKSYMM)

ATOMDISP	displacement of atoms	26	I
ATOMINSE	addition of atoms	26	I
ATOMREMO	removal of atoms	27	I
ATOMROT	rotation of groups of atoms	27	I
ATOMSUBS	substitution of atoms	28	I
ELASTIC	distortion of the lattice	31	I
POINTCHG	point charges input	39	I
USESAED	given symmetry allowed elastic distortions, reads δ	46	I
SUPERCEL	generation of supercell - input refers to primitive cell	44	I
SUPERCON	generation of supercell - input refers to conventional cell	44	I

From crystals to slabs

SLABCUT	generation of a slab parallel to a given plane (3D→2D)	42	I

From periodic structure to clusters

CLUSTER	cutting of a cluster from a periodic structure (3D→0D)	29	I
HYDROSUB	border atoms substituted with hydrogens (0D→0D)	35	I

Molecular crystals

MOLECULE	extraction of a set of molecules from a molecular crystal (3D→0D)	37	I
MOLEXP	variation of lattice parameters at constant symmetry and molecular geometry (3D→3D)	38	I
MOLSPLIT	periodic structure of non interacting molecules (3D→3D)	38	–
RAYCOV	modification of atomic covalent radii	40	I

BSSE correction

MOLEBSSE	counterpoise method for molecules (molecular crystals only) (3D→0D)	36	I
ATOMBSSE	counterpoise method for atoms (3D→0D)	26	I

Auxiliary and control keywords

ANGSTROM	sets inputs unit to Ångstrom	25	–
BOHR	sets input units to bohr	28	–
BOHRANGS	input bohr to Å conversion factor (0.5291772083 default value)	28	I
BOHRCR98	bohr to Å conversion factor is set to 0.529177 (CRYSTAL98 value)	–	
END/ENDG	terminate processing of geometry input		–
FRACTION	sets input unit to fractional	35	–
NEIGHBOR	number of neighbours in geometry analysis	38	I
PARAMPRT	printing of parameters controlling dimensions of static allocation arrays	39	–
PRINTCHG	printing of point charges coordinates in geometry output	39	
PRINTOUT	setting of printing options by keywords	40	–
SETINF	setting of inf array options	42	I
SETPRINT	setting of printing options	42	I
STOP	execution stops immediately	43	–
TESTGEOM	stop after checking the geometry input	46	–

Output of data on external units

COORPRT	coordinates of all the atoms in the cell	30	–
EXTPRT	generation of file as CRYSTAL input	32	–
FINDSYM	generation of file as FINDSYM input	35	–
MOLDRAW	generation of file for the program MOLDRAW	36	–
STRUCPRT	cell parameters and coordinates of all the atoms in the cell	43	–

External electric field - modified Hamiltonian

FIELD	electric field applied along a periodic direction	32	I
FIELDCON	electric field applied along a non periodic direction	34	I

Geometry optimization

OPTGEOM	Geometry optimization	82	I
	Type of optimization (default: atom coordinates)		
FULLOPTG	full geometry optimization		–
CELLONLY	cell parameters optimization		–
INTREDUN	optimization in redundant internal coordinates		–
ITATOCEL	iterative optimization (atom/cell)		–
CVOLOPT	full geometry optimization at constant volume		–
	Initial Hessian		
HESGUESS	initial guess for the Hessian		I
HESSIDEN	initial guess for the Hessian - identity matrix		–
HESSMOD1	initial guess for the Hessian - model 1 (default)		–
HESSMOD2	initial guess for the Hessian - model 2		–
	Convergence criteria modification		
TOLDEG	RMS of the gradient [0.0003]		I
TOLDEX	RMS of the displacement [0.0012]		I
TOLDEE	energy difference between two steps [10^{-7}]		I
MAXCYCLE	max number of optimization steps		I
	Optimization control		
FRAGMENT	partial geometry optimization		I
RESTART	data from previous run		–
FINALRUN	Wf single point with optimized geometry		I
	Gradient calculation control		
NUMGRAD	numerical first derivatives		–
	Printing options		
PRINTFORCES	atomic gradients		–
PRINTHESS	Hessian		–
PRINTOPT	optimization procedure		–
PRINT	verbose printing		–

Frequencies at Γ

FREQCALC	Frequency at Γ - Harmonic calculation 4- [default]	98	I
ANALYSIS			–
[NOANALYSIS]			–
DIELISO			I
DIELTENS			I
FRAGMENT			I
INTENS			–
[NOINTENS]			–
ISOTOPES			I
[MODES]			–
NOMODES			–
NORMBORN			–
NUMDERIV			I
PRESSURE			I
PRINT			–
RESTART			–
SCANMODE			I
STEPSIZE			I
TEMPERAT			I
TESTFREQ			–
[USESYMM]			–
NOUSESYMM			–
END[FREQ]			–

ANHARM	Frequency at Γ - Anharmonic calculation	106	I
	TESTANHA		–
	KEEPSYMM		–
	ISOTOPES		I
	NOGUESS		–
	END[ANHA]		–

ANGLES

This option prints the angle the $\widehat{AXB}$, where X is one of the irreducible (that is, non symmetry equivalent) atoms of the unit cell, and A and B belong to its m-th and n-th stars of neighbors.

rec	variable	meaning
• ∗	NATIR	number of X atoms to be considered; they are the first NATIR in the list of irreducible atoms (flag "T" printed) generated by CRYSTAL
∗	NSHEL	number of stars of neighbors of X to be considered; all the angles $\widehat{AXB}$, where A and B belong to the first NSHEL neighbors of X, are printed out

Though the keyword **ANGLES** can be entered in geometry input, full input deck must be supplied (block 1-2-3), in order to obtain information on bond angles, when neighbors analysis is printed.

Example. Bulk Silicon. There is 1 irreducible atom, and the first star of neighbors contain 4 atoms: (from CRYSTAL output):

```
 COORDINATES OF THE EQUIVALENT ATOMS (FRACTIONARY UNITS)

 N ATOM ATOM    Z          X             Y             Z
   IRR  EQUIV
 1  1     1   14 SI   1.250000E-01  1.250000E-01   1.250000E-01
 2  1     2   14 SI  -1.250000E-01 -1.250000E-01 - 1.250000E-01
    - - - - - - - - - - - - - - - - - - - - - - -
N NUMBER OF NEIGHBORS AT DISTANCE R

STAR ATOM  N R/ANG    R/AU   NEIGHBORS (ATOM LABELS AND CELL INDICES)
 1   1 SI  4 2.3469  4.4351  2 SI   0 0 0   2 SI   1 0 0   2 SI   0 1 0
                             2 SI   0 0 1
```

The number of angles having the irreducible Silicon as vertex is: (4)*(4-1)/2 = 6

```
 ANGLES (DEGREES) ARE INDICATED AS A-X-B(I),I=1,L

 at A  cell   at X at B  cell   angle  at B  cell   angle  at B  cell   angle
                                 AXB                 AXB                 AXB
 2 SI( 0 0 0) 1 SI 2 SI( 1 0 0) 109.47 2 SI( 0 1 0) 109.47 2 SI( 0 0 1) 109.47
 2 SI( 1 0 0) 1 SI 2 SI( 0 1 0) 109.47 2 SI( 0 0 1) 109.47
 2 SI( 0 1 0) 1 SI 2 SI( 0 0 1) 109.47
```

If it is required to consider 6 stars of neighbors to compute all the angles having the irreducible Silicon as vertex, the number of angles computed will be:
(4+12+12+6+12+24)*(4+12+12+6+12+24-1)/2 = 2415

ANGSTROM - unit of measure

The unit of length in geometry editing is set to Ångstrom, (default value).

ATOMBSSE - counterpoise for closed shell atoms and ions

rec	variable	meaning
• ∗	IAT	*label* of the atom in the reference cell
	NSTAR	maximum number of stars of neighbors included in the calculation.
	RMAX	maximum distance explored searching the neighbors of the atom.

A cluster is defined including the selected atom and the basis functions belonging to the NSTAR sets of neighbours, when their distance R from the central atom is smaller than RMAX. The atomic wave function is not computed by the atomic package, but by the standard CRYSTAL route for 0D, 1 atom system. **UHF** and **SPINLOCK** must be used to define a reasonable orbital occupancy. It is suggested to compute the atomic wave function using a program properly handling the electronic configuration of open shell atoms.
Warning. The system is 0D. No reciprocal lattice information is required in the **scf** input (Section 1.3, page 18).

ATOMDISP

rec	variable	meaning
• ∗	NDISP	number of atoms to be displaced
		insert NDISP records II
• ∗	LB	*label* of the atom to be moved
	DX,DY,DZ	increments of the coordinates in the primitive cell [Å].

Selected atoms are displaced in the primitive cell. The point symmetry of the system may be altered (default value **BREAKSYM**, page 29). To displace all the atoms symmetry related, **KEEPSYMM** must be inserted before **ATOMDISP**.
Increments are in Ångstrom, unless otherwise requested (keyword **BOHR**, **FRACTION**, page 25). See tests 17, 20, 37.

ATOMINSE

rec	variable	meaning
• ∗	NINS	number of atoms to be added
		insert NINS records II
• ∗	NA	conventional atomic number
	X,Y,Z	coordinates [Å] of the inserted atom. Coordinates refer to the primitive cell.

New atoms are added to the primitive cell. Coordinates are in Ångstrom, unless otherwise requested (keyword **BOHR**, **FRACTION**, page 25). Remember that the original symmetry of the system is maintained, applying the symmetry operators to the added atoms if the keyword **KEEPSYMM** (page 29) was previously entered. The default is **BREAKSYM** (page 29). Attention should be paid to the neutrality of the cell (see **CHARGED**, page 47). See tests 16, 35, 36.

ATOMORDE

After processing the standard geometry input, the symmetry equivalent atoms in the reference cell are grouped. They may be reordered, following a chemical bond criterion. This simplifies the interpretation of the output when the results of bulk molecular crystals are compared with those of the isolated molecule. See option **MOLECULE** (page 37) and **MOLSPLIT** (page 38). No input data are required.
For molecular crystals only.

ATOMREMO

rec	variable	meaning
• ∗	NL	number of atoms to remove
• ∗	LB(L),L=1,NL	*label* of the atoms to remove

Selected atoms, and related basis set, are removed from the primitive cell (see test 16). A vacancy is created in the lattice. The symmetry can be maintained (**KEEPSYMM**), by removing all the atoms symmetry-related to the selected one, or reduced (**BREAKSYM**, default). Attention should be paid to the neutrality of the cell (see **CHARGED**, page 47).
NB. The keyword **GHOSTS** (basis set input, page 49) allows removal of selected atoms, leaving the related basis set.

ATOMROT

rec	variable	value	meaning
• ∗	NA	0	all the atoms of the cell are rotated and/or translated
		>0	only NA selected atoms are rotated and/or translated.
		<0	the atom with *label* |NA| belongs to the molecule to be rotated. The program selects all the atoms of the molecule on the base of the sum of their atomic radii (Table on page 40).
			if NA > 0, insert NA data II
• ∗	LB(I),I=1,NA		*label* of the atoms to be rotated and/or translated.
• ∗	ITR	>0	translation performed. The selected NA atoms are translated by -**r**, where **r** is the position of the ITR-th atom. ITR is at the origin after the translation.
		≤ 0	a general translation is performed. See below.
		=999	no translation.
	IRO	> 0	a rotation around a given axis is performed. See below.
		< 0	a general rotation is performed. See below.
		=999	no rotation.
			if ITR<0 insert II
• ∗	X,Y,Z		Cartesian components of the translation vector [Å]
			if ITR=0 insert II
• ∗	N1,N2		*label* of the atoms defining the axis.
	DR		translation along the axis defined by the atoms N1 and N2, in the direction N1 → N2 [Å].
			if IRO<0 insert II
• ∗	A,B,G		Euler rotation angles (degree).
	IPAR		defines the origin of the Cartesian system for the rotation
		0	the origin is the barycentre of the NAT atoms
		>0	the origin is the atom of *label* IPAR
			if IRO>0 insert II
• ∗	N1,N2		*label* of the atoms that define the axis for the rotation
	ALPHA	≠ 0.	rotation angle around the N1–N2 axis (degrees)
		0.	the selected atoms are rotated anti-clockwise in order to orientate the N1–N2 axis parallel to the z axis.

This option allows to rotate and/or translate the specified atoms. When the rotation of a molecule is required (NA < 0), the value of the atomic radii must be checked, in order to obtain a correct definition of the molecule. It is useful to study the conformation of a molecule in a zeolite cavity, or the interaction of a molecule (methane) with a surface (MgO).
The translation of the selected group of atoms can be defined in three different ways:

1. Cartesian components of the translation vector (ITR < 0);
2. modulus of the translation vector along an axis defined by two atoms (ITR = 0);

3. sequence number of the atom to be translated to the origin. All the selected atoms are subjected to the same translation (ITR > 0).

The rotation can be performed in three different ways:

1. by defining the Euler rotation angles α, β, γ and the origin of the rotating system (IRO < 0). The axes of the rotating system are parallel to the axes of the Cartesian reference system. (The rotation is given by: $R^{\alpha z} R^{\beta x} R^{\gamma z}$, where R are the rotation matrices).
2. by defining the rotation angle α around an axis defined by two atoms A and B. The origin is at A, the positive direction A→B.
3. by defining a *z'* axis (identified by two atoms A and B). The selected atoms are rotated, in such a way that the A–B *z'* axis becomes parallel to the *z* Cartesian axis. The origin is at A and the positive rotation anti clockwise (IRO>0, α=0).

The selected atoms are rotated according to the defined rules, the cell orientation and the cartesian reference frame are not modified. The symmetry of the system is checked after the rotation, as the new geometry may have a different symmetry.
See tests 15, rotation of the NH_3 molecule in a zeolite cavity, and 16, rotation of the H_2O molecule in the zeolite cavity.

ATOMSUBS

rec	variable	meaning
• *	NSOST	number of atoms to be substituted
		insert NSOST records II
• *	LB	*label* of the atom to substitute
	NA(LB)	conventional atomic number of the new atom

Selected atoms are substituted in the primitive cell (see test 17, 34, 37). The symmetry can be maintained (**KEEPSYMM**), by substituting all the atoms symmetry-related to the selected one, or reduced (**BREAKSYM**, default). Attention should be paid to the neutrality of the cell: a non-neutral cell will cause an error message, unless allowed by entering the keyword **CHARGED**, page 47.

ATOMSYMM

The point group associated with each atomic position and the set of symmetry related atoms are printed. No input data are required. This option is useful to find the internal coordinates to be relaxed when the unit cell is deformed (see **ELASTIC**, page 31).

BOHR

The keyword **BOHR** sets the unit of distance to bohr. When the unit of measure is modified, the new convention is active for all subsequent geometry editing.
The conversion factor Ångstrom/bohr is 0.5291772083 (CODATA 1998). This value can be modified by entering the keyword **BOHRANGS** and the desired value in the record following.
The keyword **BOHRCR98** sets the conversion factor to 0.529177, as in the program CRYSTAL98.
CRYSTAL88 default value was 0.529167).

BOHRANGS

rec	variable	meaning
• *	BOHR	conversion factor Ångstrom/bohr

The conversion factor Ångstrom/bohr can be user-defined.
In CRYSTAL88 the default value was 0.529167.
In CRYSTAL98 the default value was 0.529177.

BOHRCR98

The conversion factor Ångstrom/bohr is set to 0.529177, as in CRYSTAL98. No input data required.

BREAKSYM

Under control of the **BREAKSYM** keyword (the default), subsequent modifications of the geometry are allowed to alter (reduce: the number of symmetry operators cannot be increased) the point-group symmetry. The new group is a subgroup of the original group and is automatically obtained by **CRYSTAL**.
The symmetry may be broken by attributing different spin (**ATOMSPI**, block34, page 57) to atoms symmetry related by geometry.
Example: When a CO molecule is vertically adsorbed on a (001) 3-layer MgO slab, (D_{4h} symmetry), the symmetry is reduced to C_{4v}, if the **BREAKSYM** keyword is active. The symmetry operators related to the σ_h plane are removed. However, if **KEEPSYMM** is active, then additional atoms will be added to the underside of the slab so as to maintain the σ_h plane (see page 26, keyword **ATOMINSE**).

CLUSTER - a cluster (0D) from a periodic system

The **CLUSTER** option allows one to cut a finite molecular cluster of atoms from a periodic lattice. The size of the cluster (which is centred on a specified 'seed point' A) can be controlled either by including all atoms within a sphere of a given radius centred on A, or by specifying a maximum number of symmetry-related stars of atoms to be included.
The cluster includes the atoms B (belonging to different cells of the direct lattice) satisfying the following criteria:

1. those which belong to one of the first N (input data) stars of neighbours of the *seed* point of the cluster.

   and

2. those at a distance R_{AB} from the seed point which is smaller then RMAX (input datum).

The resulting cluster may not reproduce exactly the desired arrangement of atoms, particularly in crystals with complex structures such as zeolites, and so it is possible to specify border modifications to be made after definition of the core cluster.
Specification of the core cluster:

rec	variable	value	meaning
• ∗	X, Y, Z		coordinates of the centre of the cluster [Å] (the seed point)
	NST		maximum number of stars of neighbours explored in defining the core cluster
	RMAX		radius of a sphere centred at X,Y,Z containing the atoms of the core cluster
• ∗	NNA	$\neq 0$	print nearest neighbour analysis of cluster atoms (according to a radius criterion)
	NCN	0	testing of coordination number during hydrogen saturation carried out only for Si (coordination number 4), Al (4) and O(2)
		N	N user-defined coordination numbers are to be defined
			if NNA $\neq$ 0 insert 1 record II
• ∗	RNNA		radius of sphere in which to search for neighbours of a given atom in order to print the nearest neighbour analysis
			if NCN $\neq$ 0 insert NCN records II
• ∗	L		conventional atomic number of atom
	MCONN(L)		coordination number of the atom with conventional atomic number L. MCONN=0, coordination not tested

Border modification:

rec	variable	value	meaning
• ∗	NMO		number of border atoms to be modified
			if NMO > 0 insert NMO records II
• ∗	IPAD		*label* of the atom to be modified *(cluster sequence)*
	NVIC		number of stars of neighbours of atom IPAD to be added to the cluster
	IPAR	= 0	no hydrogen saturation
		≠ 0	cluster border saturated with hydrogen atoms
	BOND		bond length Hydrogen-IPAD atom (direction unchanged).
			if NMO < 0 insert II
• ∗	IMIN		*label* of the first atom to be saturated *(cluster sequence)*
	IMAX		*label* of the last atom to be saturated *(cluster sequence)*
	NVIC		number of stars of neighbours of each atom to be added to the cluster
	IPAR	= 0	no hydrogen saturation
		≠ 0	cluster border saturated with hydrogen atoms
	BOND		H-cluster atom bond length (direction unchanged).

The two kinds of possible modification of the core cluster are (a) addition of further stars of neighbours to specified border atoms, and (b) saturation of the border atoms with hydrogen. This latter option can be essential in minimizing border electric field effects in calculations for covalently-bonded systems.
(Substitution of atoms with hydrogen is obtained by **HYDROSUB**).
The hydrogen saturation procedure is carried out in the following way. First, a coordination number for each atom is assumed (by default 4 for Si, 4 for Al and 2 for O, but these may be modified in the input deck for any atomic number). The actual number of neighbours of each specified border atom is then determined (according to a covalent radius criterion) and compared with the assumed connectivity. If these two numbers differ, additional neighbours are added. If these atoms are not neighbours of any other existing cluster atoms, they are converted to hydrogen, otherwise further atoms are added until the connectivity allows complete hydrogen saturation whilst maintaining correct coordination numbers.
The *label* of the IPAD atoms refers to the generated cluster, *not* to the original unit cell. The preparation of the input thus requires two runs:

1. run using the **CLUSTER** option with NMO=0, in order to generate the sequence number of the atoms in the core cluster. The keyword **TESTGEOM** should be inserted in the geometry input block. Setting NNA ≠ 0 in the input will print a coordination analysis of all core cluster atoms, including all neighbours within a distance RNNA (which should be set slightly greater than the maximum nearest neighbour bond length). This can be useful in deciding what border modifications are necessary.

2. run using the **CLUSTER** option with NMO ≠ 0, to perform desired border modifications.

Note that the standard CRYSTAL geometry editing options may also be used to modify the cluster (for example by adding or deleting atoms) placing these keywords after the specification of the **CLUSTER** input.
Warning. The system is 0D. No reciprocal lattice information is required in the **scf** input (Section 1.3, page 18). See test 16. subsection***COORPRT** Geometry information is printed: cell parameters, fractionary coordinates of all atoms in the reference cell, symmetry operators.
A formatted file, "fort.33" , is written. See Appendix E, page 215. No input data are required. The file "fort.33" has the right format for the program **MOLDEN** [23] which can be downloaded from:
`www.cmbi.ru.nl/molden/molden.html`

ELASTIC

An elastic deformation of the lattice may be defined in terms of the Z or ϵ strain tensors defined in section 8.9, page 182.

rec		variable	value	meaning
•	*	IDEF	±1	deformation through equation 8.36, Z matrix.
			±2	deformation through equation 8.35: ϵ matrix.
			>0	volume conserving deformation (equation 8.37).
			<0	not volume conserving (equation 8.36 or 8.35).
•	*	D11 D12 D13		first row of the matrix.
•	*	D21 D22 D23		second row of the matrix.
•	*	D31 D32 D33		third row of the matrix.

The elastic constant is $V^{-1}\frac{\partial^2 E}{\partial \epsilon_i^2}\mid_{\epsilon_i=0}$, where V is the volume of the primitive unit cell.
The symmetry of the system is defined by the symmetry operators in the new crystallographic cell. The keyword **MAKESAED** gives information on symmetry allowed elastic distortions.
The calculation of the elastic constants with **CRYSTAL** requires the following sequence of steps:

1. select the ϵ_{ij} matrix elements to be changed (for example, $\epsilon_4 \equiv \epsilon_{23} + \epsilon_{32}$), and set the others ϵ_j to zero;
2. perform calculations with different values of the selected matrix element(s) ϵ_i: 0.02, 0.01, 0.001, -0.001, -0.01, -0.02, for example, and for each value compute the total energy E;
3. perform a polynomial fit of E as a function of ϵ_i.

ϵ is adimensional, Z in Å(default) or in bohr (page 25). The suggested value for IDEF is -2 (deformation through equation 8.35, *not* volume conserving). The examples refer to this setting.

Example
Geometry input deck to compute one of the energy points used for the evaluation of the C_{44} (page 185) elastic constants of Li_2O [24].

CRYSTAL	
0 0 0	3D code
225	3D space group number
4.5733	lattice parameter (Å)
2	2 non equivalent atoms in the primitive cell
8 0.0 0.0 0.0	Z=8, Oxygen; x, y, z
3 .25 .25 .25	Z=3, Lithium; x, y, z
ATOMSYMM	printing of the point group at the atomic positions
ELASTIC	
-2	deformation not volume conserving through equation 8.35
0. 0.03 0.03	ϵ matrix input by rows
0.03 0. 0.03	
0.03 0.03 0.	
ATOMSYMM	printing of the point group at the atomic positions after the deformation
. . . .	

A rhombohedral deformation is obtained, through the ϵ matrix. The printout gives information on the crystallographic and the primitive cell, before and after the deformation:

```
LATTICE PARAMETERS (ANGSTROMS AND DEGREES) OF
(1) ORIGINAL PRIMITIVE CELL
(2) ORIGINAL CRYSTALLOGRAPHIC CELL
(3) DEFORMED PRIMITIVE CELL
(4) DEFORMED CRYSTALLOGRAPHIC CELL
```

```
       A          B          C         ALPHA       BETA        GAMMA      VOLUME
(1)  3.233811  3.233811  3.233811  60.000000  60.000000  60.000000  23.912726
(2)  4.573300  4.573300  4.573300  90.000000  90.000000  90.000000  95.650903
(3)  3.333650  3.333650  3.333650  56.130247  56.130247  56.130247  23.849453
(4)  4.577414  4.577414  4.577414  86.514808  86.514808  86.514808  95.397811
```

After the deformation of the lattice, the point symmetry of the Li atoms is C_{3v}, where the C_3 axis is along the (*x,x,x*) direction. The Li atoms can be shifted along the principal diagonal, direction (*x,x,x*) of the primitive cell without altering the point symmetry, as shown by the printing of the point group symmetry obtained by the keyword **ATOMSYMM** (page 28).
See test20 for complete input deck, including shift of the Li atoms.
See test38 ($KCoF_3$).

END

Terminate processing of block 1, geometry definition, input. Execution continues. Subsequent input records are processed, if required.
Processing of geometry input block stops when the first three characters of the string are "END". Any character can follow: ENDGEOM, ENDGINP, etc etc.

EXTPRT

A formatted input deck with explicit structural/symmetry information is written in file "fort.34". If the keyword is entered many times, the data are overwritten. The last geometry is recorded. The deck may be used as crystal geometry input to CRYSTAL through the **EXTERNAL** keyword.
For instance, to enter the final optimized geometry, or a geometry obtained by editing operations who modified the original space group or periodicity.
When geometry optimization is performed, the name of the file is "optcxxx", being xxx the number of the cycle, and it is automatically written at each cycle.
See Appendix E, page 211. No input data are required.

FIELD - Electric field along a periodic direction

rec	variable	value	meaning
• ∗	E0MAX		electric field intensity E_0 (in atomic units)
• ∗	DIRE(I),I=1,3		crystallographic (Miller) indices of the plane perpendicular to the electric field
• ∗	SMFACT		supercell expansion factor
∗	IORTO	0	non-orthogonal supercell
		1	orthogonal supercell
• ∗	MUL		number of term in Fourier expansion for triangular electric potential
∗	ISYM	+1	triangular potential is symmetric with respect to the $z = 0$ plane
		−1	triangular potential is anti-symmetric with respect to the $z = 0$ plane

This option can be used with polymers, slabs and crystals and permits to apply an electric field along a periodic direction of the system.
The effect of a periodic electric field ($\vec{E}$) is taken into account according to a perturbation scheme. The Hamiltonian (Fock or Kohn-Sham) can be written as::

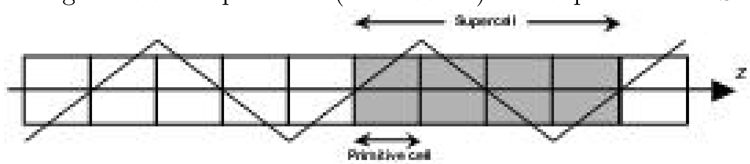
$$\hat{H} = \hat{H}_0 + \hat{H}_1(\vec{E}) \tag{2.1}$$

where $\hat{H}_0$ is the unperturbed Hamiltonian and $\hat{H}_1(\vec{E})$ the electric potential term.

During the SCF procedure crystalline orbitals are relaxed under the effect of the field, leading to a perturbed wave function and charge density.

The applied electric field has a square-wave form, that corresponds to a triangular (”sawtooth”) electric potential.
Due to the form of the potential, a single unit cell must contain both positive and negative part of the square wave electric field. Then, in order to maintain translational invariance of the system a new, expanded, unit cell is automatically created by adopting a supercell approach (see keywords **SUPERCEL/SUPERCON**, page 44).
This procedure consists in two automatic steps: the re-orientation of the c lattice parameter along the chosen field direction and the multiplication of this lattice vector according to the supercell expansion factor ($\vec{C}$=SMFACT$\cdot\vec{c}$, see fig. 2.1). By varying this parameter is possible to control the period of the electric potential and therefore the length of the constant region of the electric field.

Figure 2.1: Triangular electric potential (”sawtooth”) in a supercell with $SMFACT = 4$.



Then, for computational reasons, an automatic rotation of the crystal in the cartesian reference system is performed by aligning $\vec{C}$ (and therefore $\vec{E}$) along the z cartesian direction (see keyword **ROTCRY**, page 41). After thiese transormations the field is along the z direction, and the perturbation $\hat{H}_1(\vec{E})$ takes the form:

$$\hat{H}_1^{\pm}(E_z) = V(z) = -qE_0 \cdot f^{\pm}(z) \tag{2.2}$$

where the f^+ (f^-) function is expanded as a Fourier series and is chosen according to the symmetry of the supercell in the direction of the applied field as follows:

$$f^+(z) = \frac{2C}{\pi^2}\sum_{k=0}^{+\infty}\frac{1}{(2k+1)^2}\cos\left(\frac{2\pi(2k+1)z}{C}\right) \tag{2.3}$$

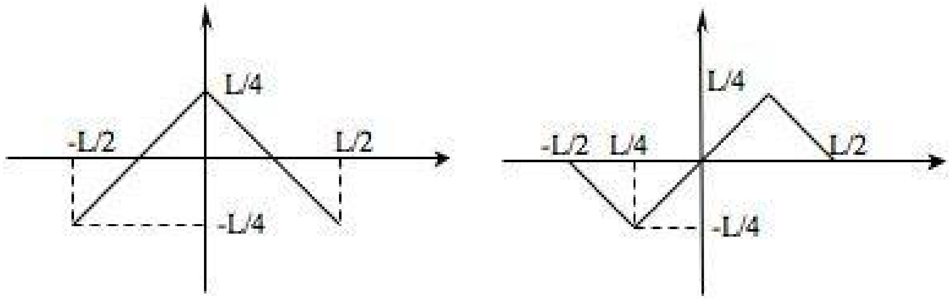
$$f^-(z) = \frac{2C}{\pi^2}\sum_{k=0}^{+\infty}\frac{(-1)^k}{(2k+1)^2}\sin\left(\frac{2\pi(2k+1)z}{C}\right) \tag{2.4}$$

1. In order to evaluate the dielectric constant of a material in the direction of the applied field it is necessary to run a PROPERTIES calculation with the keyword **DIEL** (see page 116). In this way the perturbed wave function is used for the calculation of ϵ, following a macroscopic average scheme, as described in references [25], [26].

2. The field is along the z axis for 3D-crystal calculations; it is along the x for 1D-polymer and 2D-slab calculations.

3. In calculations of the dielectric constant, more accurate results can be achieved by increasing the SMFACT value. This will lead to systems characterized by a high number of atoms with large computational costs. The option IORTO = 0 allows to consider non-orthogonal supercells, characterized by the same dielectric properties of orthogonal cells, but with a lower number of atoms.

Figure 2.2: Left: symmetric triangular electric potential (ISIM = 1). Right anti-symmetric triangular electric potential (ISYM=-1).



4. In 3D-crystals, the electric potential takes a triangular form to maintain translational symmetry and electric neutrality of cell. The symmetry of triangular potential has two options:

   a) ISYM=+1, triangular potential is symmetric with respect to the center of the supercell, along the z axis. Use this option if there is a symmetry plane orthogonal to the z axis.

   b) ISYM=−1, triangular potential is anti-symmetric. This option can be used when the supercell does not have a symmetry plane orthogonal to z axis.

5. MUL, the number of terms in the Fourier expansion, can take values between 1 and 60.

   MUL=40 is sufficient to adequately reproduce the triangular shape of the potential.

6. High E0MAX values are inconsistent with perturbation method, the choice of E0MAX depends on the dielectric susceptibility of the system and on the gap width. For small gap cases, use of eigenvalue level shifting technique is recommended (keyword **LEVSHIFT**, page 71).

7. When an external field is applied, the system can become conducting during the SCF procedure. In order to avoid convergence problems, it is advisable to set the shrinking factor of the Gilat net ISP equal to 2× IS, where IS is the Monkhorst net shrinking factor (see SCF input, page 75).

Conversion factors for electric field:
1 AU = 1.71527E+07 ESU·CM^{-2} = 5.72152E+01 C ·M^{-2} = 5.14226E+11 V·M^{-1}

FIELDCON - Electric field along non-periodic direction

rec	variable	meaning
• ∗	E(I),I=N,3	field components along x,y,z directions

For a brief theoretical introduction see keyword FIELD.
This option can be used with molecules, polymers, slabs and permits to apply an electric field along a non-periodic direction of the system.

1. For molecules (N=1) three components of the field must be supplied, as the field can be directed along any direction.

2. For polymers (N=2) two components (y,z) of the field must be defined; the x component of the field must be zero because the default orientation of polymers is along the x axis.

3. For slabs (N=3) just one component (z) of the field have to be defined; the x,y components must be zero because the default orientation of slabs in is in x-y plan.

Conversion factors for electric field:
1 AU = 1.71527E+07 ESU·CM^{-2} = 5.72152E+01 C ·M^{-2} = 5.14226E+11 V·M^{-1}
This option can evaluate the dielectric response of the molecule, polymer or slab in a direction of non periodicity (see option FIELD for a field along a periodicity direction).
Consider the following expansion of the total energy of the system as a function of the applied field:

$$E(F_0) = E_0 - \mu F_0 - \frac{1}{2!}\alpha F_0^2 - \frac{(}{1})3!\beta F_0^3 - \frac{1}{4!}\gamma F_0^4 - \cdots \quad (2.5)$$

By fitting the E vs F_0 data the μ, α, β and γ values can be derived. See http://www.crystal.unito.it → tutorials → Static dielectric constants..

FINDSYM

Geometry information is written in file FINDSYM.DAT, according to the input format of the program FINDSYM.

http://stokes.byu.edu/findsym.html

FINDSYM: Identify the space group of a crystal, given the positions of the atoms in a unit cell. When geometry editing modifies the basic input space group, the symmetry of the system is identified by the symmetry operators only. The program *FINDSYM* allows identification of the space group.

FRACTION

The keyword **FRACTION** means input coordinates given as fraction of the lattice parameter in subsequent input, along the direction of translational symmetry:

x,y,z	crystals (3D)
x,y	slabs (2D; *z* in Ångstrom or bohr)
x	polymers (1D; *y,z* in Ångstrom or bohr)

no action for 0D. When the unit of measure is modified, the new convention is active for all subsequent geometry editing.

FREQCALC - Harmonic frequencies at Γ

See Chapter 4, page 98.

HYDROSUB - substitution with hydrogen atoms

rec	variable	meaning
• ∗	NSOST	number of atoms to be substituted with hydrogen
		insert NSOST records II
• ∗	LA	*label* of the atom to substitute
	LB	*label* of the atom linked to LA
	BH	bond length B-Hydrogen

Selected atoms are substituted with hydrogens, and the bond length is modified. To be used after **CLUSTER**.

KEEPSYMM

In any subsequent editing of the geometry, the program will endeavour to maintain the number of symmetry operators, by requiring that atoms which are symmetry related remain so after geometry editing (keywords: **ATOMSUBS**, **ATOMINSE**, **ATOMDISP**, **ATOMREMO**) or the basis set (keywords **CHEMOD**, **GHOSTS**).
Example: When a CO molecule is vertically adsorbed on a (001) 3-layer MgO slab, (D_{4h} symmetry) (see page 26, keyword **ATOMINSE**), the symmetry is reduced to C_{4v}, if the **BREAKSYM** keyword is active. The symmetry operators related to the σ_h plane are removed. However, if **KEEPSYMM** is active, then additional atoms will be added to the underside of the slab so as to maintain the σ_h plane.

MAKESAED

Symmetry allowed elastic distortions are printed. No input data required.

MODISYMM

rec	variable	meaning
• ∗	N	number of atoms to be attached a flag
• ∗	LA,LF(LA),L=1,N	atom *labels* and flags (n couples of integers in 1 record).

The point symmetry of the lattice is lowered by attributing a different "flag" to atoms related by geometrical symmetry. The symmetry operators linking the two atoms are removed and the new symmetry of the system is analyzed. For instance, when studying spin-polarized systems, it may be necessary to apply different spins to atoms which are related by geometrical symmetry.

MOLDRAW

A formatted input deck for the visualization program ***MOLDRAW*** [27] is written in file MOLDRAW.DAT . If the keyword is entered many times, the data are overwritten. The last geometry can be visualized.
The last version of the program ***MOLDRAW*** reads ***crystal*** standard output, and can generate a movie from an optimization run. No input data are required. See:

http://www.moldraw.unito.it.

MOLEBSSE - counterpoise for molecular crystals

rec	variable	meaning
• ∗	NMOL	number of molecules to be isolated
		insert NMOL records
• ∗	ISEED	*label* of one atom in the n-th molecule
	J,K,L	integer coordinates (direct lattice) of the primitive cell containing the ISEED atom
• ∗	NSTAR	maximum number of stars of neighbours included in the calculation
	RMAX	maximum distance explored searching the neighbours of the atoms belonging to the molecule(s)

The counterpoise method [28] is applied to correct the Basis Set Superposition Error in molecular crystals. A molecular calculation is performed, with a basis set including the basis functions of the selected molecules and the neighbouring atoms. The program automatically finds all the atoms of the molecule(s) containing atom(s) ISEED (keyword **MOLECULE**, page 37). The molecule is reconstructed on the basis of the covalent radii reported in Table on page 40. They can be modified by running the option **RAYCOV**, if the reconstruction of the molecule fails. The radius of the hydrogen atom is very critical when intermolecular hydrogen bonds are present.
All the functions of the neighbouring atoms in the crystal are added to the basis set of the selected molecule(s) such that both the following criteria are obeyed:

1. the atom is within a distance R lower than RMAX from at least one atom in the molecule

   and

2. the atom is within the NSTAR-th nearest neighbours of at least one atom in the molecule.

For molecular crystals only.

Warning Do not use with ECP

Warning. The system obtained is 0D. No reciprocal lattice information is required in the **scf** input (Section 1.3, page 18). See test 19.

MOLECULE - Extraction of n molecules from a molecular crystal

rec	variable	meaning
• ∗	NMOL	number of molecules to be isolated
		insert NMOL records II
• ∗	ISEED	*label* of one atom in the nth molecule
	J,K,L	integer coordinates (direct lattice) of the primitive cell containing the ISEED atom

The option **MOLECULE** isolates one (or more) molecules from a molecular crystal on the basis of chemical connectivity, defined by the sum of the covalent radii (Table on page 40).
The covalent radii can be modified by running the option **RAYCOV**, if the reconstruction of the molecule fails. The covalent radius of the hydrogen atom is very critical when intermolecular hydrogen bonds are present.
The input order of the atoms (atoms symmetry related are grouped) is modified, according to the chemical connectivity. The same order of the atoms in the bulk crystal is obtained by entering the keyword **ATOMORDE** (see Section 2.1, page 26). The total number of electrons attributed to the molecule is the sum of the shell charges attributed in the basis set input (input block 2, Section 1.2, page 14) to the atoms selected for the molecule.
The keyword **GAUSS98**, entered in input block 2 (basis set input), writes an input deck to run Gaussian 98 (see page 49)

For molecular crystals only.

Warning. The system is 0D. No reciprocal lattice information is required in the **scf** input (Section 1.3, page 18).
Test 18 - Oxalic acid. In the 3D unit cell there are four water and two oxalic acid molecules. The input of test 18 refers to a cluster containing a central oxalic acid molecule surrounded by four water molecules.

MOLEXP - Variation of lattice parameters at constant symmetry and molecular geometry

rec	variable	meaning
• ∗	δa,[δb],[δc],	increments of the minimal set of crystallographic cell parameters:
	[$\delta\alpha$],[$\delta\beta$]	translation vectors length [Ångstrom],
	[$\delta\gamma$]	crystallographic angles (degrees)

The cell parameters (the minimum set, see page 11) are modified, according to the increments given in input. The volume of the cell is then modified. The symmetry of the lattice and the geometry (bond lengths and bond angles) of the molecules within the cell is kept. The fractional coordinates of the barycentre of the molecules are kept constant, the cartesian coordinates redefined according to the modification of the lattice parameters. Optimization of the geometry with reference to the compactness of the lattice is allowed, keeping constant the geometry of the molecules. When there are very short hydrogen bonds linking the molecules in the lattice, it may be necessary a modification of the atomic radii to allow proper identification of the molecules (see option **RAYCOV**, page 40)

MOLSPLIT - Periodic lattice of non-interacting molecules

In order to compare bulk and molecular properties, it can be useful to build a density matrix as a superposition of the density matrices of the isolated molecules, arranged in the same geometry as in the crystal. The keyword **MOLSPLIT** (no additional input required) performs an expansion of the lattice, in such a way that the molecules of the crystal are at an "infinite" distance from each other. The crystal coordinates are scaled so that the distances inside the molecule are fixed, and the distances among the molecules are expanded by a factor 100, to avoid molecule-molecule interactions. *The 3D translational symmetry is not changed.* Reciprocal lattice information is required in the **scf** input (Section 1.3, page 18).
A standard wave function calculation of the expanded crystal is performed. The density matrix refers to the non-interacting subsystems. Before running *properties*, the lattice is automatically contracted to the bulk situation given in input. If a charge density or electrostatic potential map is computed (**ECHG**, **POTM** options), it corresponds to the superposition of the charge densities of the isolated molecules in the bulk geometry.

This option must be used only for molecular crystals only (no charged fragments).

Warning: the DFT grid is not designed for the expanded lattice yet. Large memory allocation may be necessary.

See test 21.

NEIGHBOR/NEIGHPRT

rec	variable	meaning
• ∗	INEIGH	number of neighbours of each non-equivalent atom to be printed

The option is active when analyzing the crystal structure (bond lengths and bond angles) and when printing the bond populations following Mulliken analysis. Full input deck must be given (block 1-2-3),in order to obtain neighbors analysis of all the non-equivalent atoms
For each non-equivalent atom information on the first INEIGH neighbours is printed: number, type, distance, position (indices of the direct lattice cell).

Warning: the neighbors analysis is performed after the symmetry analysis and the screening of the integrals. If very soft tolerances for the integrals screening are given in input, it may happen that the information is not given for all the neighbors requested, as their are not taken into account when truncation criteria are applied.

NOSHIFT

It may be used before **SUPERCEL** keyword. It avoids shift of the origin in order to minimize the number of symmetry operators with finite translation component. No input data are required.

OPTGEOM - Full geometry optimization

See Chapter 3, page 82.

ORIGIN

The origin is moved to minimize the number of symmetry operators with finite translation components. Suggested before cutting a slab from a 3D structure (option **SLABCUT**, page 42). No input data are required.

PARAMPRT - printing of parametrized dimensions

The parameters controlling the dimensions of the static allocation arrays of the program are printed. No input data are required.

POINTCHG

rec	variable	meaning
• ∗	NCH	number of point charges to be added
		insert NCH records II
• ∗	X,Y,Z,QC	cartesian coordinates [Å], charge(au). Coordinates refer to the primitive cell.

Dummy atoms with formal atomic number 93, mass zero, nuclear charge as given in input (file POINTCHG.INP), are added to the primitive cell. Data are read in free format.

```
record        type of data     content
  1           1 integer       N, number of point charges
 2..2+N-1     4 real          x   y   z    charge
```

Coordinates are in Ångstrom, unless otherwise requested (keyword **BOHR**, page 25). Charges are net charges (1 electron = -1). The symmetry of the system must be removed by the keyword **SYMMREMO**.
As point charges are formally considered as "atoms", *they must be the last addition of centres to the system*.
No electron charge should be attributed to those atoms in basis set input (no atomic wave function calculation is possible). The default basis set defined by the program is a single s gaussian, with exponent 100000.
Attention should be paid to the neutrality of the cell. If the absolute value of the sum of the charges is less than 10^{-3}, the value of the charges is "normalized" to obtain 0.
The data given in input are printed. To obtain printing of coordinates and neighbour analysis of the dummy atoms in geometry output, insert the keyword **PRINTCHG**.

Not compatible with **OPTGEOM**, **FREQCALC**, **ANHARM**, **FIELD**, **FIELDCON**, **NOBIPOLA**.

PRIMITIV

Some properties (**XFAC**, **EMDL**, **EMDP**, **PROF**) input the oblique coordinates of the **k** points in the reciprocal lattice with reference to the conventional cell, though the computation refers to the primitive one. This option allows entering directly the data with reference to the primitive cell. The transformation matrix from primitive to crystallographic (Appendix A.5, page 197) is set to the identity. No effect on the CPU time: CRYSTAL always refers to the primitive cell. No input data are required.

PRINTCHG

Coordinates of the dummy atoms inserted after the keyword **POINTCHG** are printed in geometry output, basis set output, neighbor analysis. No input data required.

PRINTOUT - Setting of printing environment

Extended printout can be obtained by entering selected keywords in a printing environment beginning with the keyword **PRINTOUT** and ending with the keyword **END**. The possible keywords are found in the fifth column of the table on page 209.
Extended printing request can be entered in any input block. Printing requests are not transferred from wave function to properties calculation.
See Appendix D, page 207.

PRSYMDIR

Printing of displacement directions allowed by symmetry. The printing is done after the neighbor analysis, before computing the wave function. Full input must be supplied (3 blocks). Test run allowed with the keyword **TESTPDIM**.
No input data required.

PURIFY

This cleans up the atomic positions so that they are fully consistent with the group (to within machine rounding error). Atomic position are automatically redefined after basic geometry input. No input data are required.

RAYCOV - covalent radii modification

rec	variable	meaning
• ∗	NCOV	number of atoms for which the covalent radius is redefined
		insert NCOV records II
• ∗	NAT	atomic number ($0 \leq$NAT ≤ 92)
	RAY	covalent radius of the atom with atomic number NAT ([Å], default, or bohr, if the keyword **BOHR** precedes in the deck)

The option **RAYCOV** allows modification of the covalent radius default value for a given atom.

```
                              Table of covalent radii (Angstrom)
 H                                                                                     He
0.68                                                                                  1.47
---------                                                         -----------------------------
 Li   Be                                                           B     C     N     O     F     Ne
1.65 1.18                                                         0.93 0.81 0.78 0.78 0.76 1.68
---------                                                         -----------------------------
 Na   Mg                                                           Al    Si    P     S     Cl    Ar
2.01 1.57                                                         1.50 1.23 1.15 1.09 1.05 0.97
-----------------------------------------------------------------------------------------------
 K    Ca   Sc   Ti   V    Cr   Mn   Fe   Co   Ni   Cu   Zn   Ga   Ge   As   Se   Br   Kr
2.31 2.07 1.68 1.47 1.41 1.47 1.47 1.47 1.41 1.41 1.41 1.41 1.36 1.31 1.21 1.21 1.21 2.10
-----------------------------------------------------------------------------------------------
 Rb   Sr   Y    Zr   Ni   Mo   Tc   Ru   Rh   Pd   Ag   Cd   In   Sn   Sb   Te   I    Xe
2.31 2.10 1.94 1.60 1.52 1.52 1.42 1.36 1.42 1.47 1.68 1.62 1.62 1.52 1.52 1.47 1.47 2.66
-----------------------------------------------------------------------------------------------
 Cs   Ba   La   Hf   Ta   W    Re   Os   Ir   Pt   Au   Hg   Tl   Pb   Bi   Po   At   Rn
2.73 2.10 1.94 1.63 1.63 1.63 1.63 1.63 1.63 1.63 1.63 1.63 1.99 1.89 1.68 1.42 1.42 1.62
-----------------------------------------------------------------------------------------------
```

The choice of the covalent radius of hydrogen may be very critical when extracting a molecule from a hydrogen bonded molecular crystal. See test 15.

ROTCRY - Rotation of the crystal with respect to the reference system - developers only

This option allows to rotate the crystal with respect to the original orthonormal Cartesian reference system. The SCF procedure, both for HF and DFT calculations, is performed in the rotated geometry.
The rotation can be performed in three different ways:

1. By defining the Euler rotation angles α, β, γ and the origin of the rotating system. (The rotation is given by: $\mathrm{R}_z^\alpha \mathrm{R}_x^\beta \mathrm{R}_z^\gamma$, where R_t^θ are the rotation matrices about t by angle θ).
2. By explicitly defining the rotation matrix.
3. An automatic procedure that reorient the crystal aligning $\vec{c}$ along z Cartesian axis.

		ANGROT	Rotation defined by Euler angles α, β, γ
rec		variable	meaning
•	∗	ALPHA,BETA,GAMMA	α, β, γ rotation Euler angles (dgrees)

or

		MATROT	Rotation matrix by input
rec		variable	meaning
•	∗	R11 R12 R13	first row of the matrix.
•	∗	R21 R22 R23	second row of the matrix.
•	∗	R31 R32 R33	third row of the matrix.

or

AUTO	Automatically align c along z

The rotation involves: direct and reciprocal lattice parameters, coordinates of atoms and symmetry operators. When a DFT calculation is performed also the points of the numerical integration grid are rotated in order to preserve numerical accuracy.
Note that this keyword is different from **ATOMROT** (see pag. 27) that rotates a group of atoms without affecting the reference system.

SETINF - Setting of INF values

rec	variable	meaning
• ∗	NUM	number of INF vector positions to set
• ∗	J,INF(J),I=1,NUM	position in the vector and corresponding value

The keyword **SETINF** allows setting of a value in the INF array. It can be entered in any input section.

SETPRINT - Setting of printing options

rec	variable	meaning
• ∗	NPR	number of LPRINT vector positions to set
• ∗	J,LPRINT(J),I=1,NPR	**prtrec**; position in the vector and corresponding value

The keyword **SETPRINT** allows setting of a value in the LPRINT array, according to the information given in Appendix D, page 209. It can be entered in any input section.

SLABCUT (SLAB)

rec	variable	meaning
• ∗	h, k, l	crystallographic (Miller) indices of the plane parallel to the surface
• ∗	ISUP	label of the surface layer
	NL	number of atomic layers in the slab

The **SLABCUT** option is used to create a slab of given thickness, parallel to the given plane of the 3D lattice.
A new Cartesian frame, with the z axis orthogonal to the (hkl) plane, is defined. A *layer* is defined by a set of atoms with same z coordinate, with reference to the new Cartesian frame. The thickness of the slab, the 2D system, is defined by the number of layers. No reference is made to the chemical units in the slab. The neutrality of the slab is checked by the program.

1. The crystallographic (Miller) indices of the plane refer to the conventional cell (cubic and hexagonal systems).

2. A two-sided layer group is derived from the 3D symmetry group of the original crystal structure: the origin may be shifted to maximize the order of the layer group (keyword **ORIGIN**, page 39).

3. The unit cell is selected with upper and lower surface parallel to the (hkl) plane.

4. The 2D translation vectors $\mathbf{a_1}$ and $\mathbf{a_2}$ are chosen according to the following criteria:

   (a) minimal cell area;

   (b) shortest translation vectors;

   (c) minimum $|cos(\gamma)|$, where γ is the angle between $\mathbf{a_1}$ and $\mathbf{a_2}$.

5. The surface layer ISUP may be found from an analysis of the information printed by the **SLABINFO** (page 43) option. This information can be obtained by a test run, inserting in the geometry input block the keyword **TESTGEOM** (page 46). Only the geometry input block is processed, then the program stops.

Two separate runs are required in order to get the information to prepare the input for a full **SLABCUT** option run:

1. keyword **SLABINFO**: Rotation of the 3D cell, to have the *z* axis perpendicular to the (hkl) place, with numbering of the atomic layers in the rotated reference cell, according to the z coordinate of the atoms (insert **STOP** after **SLABINFO** to avoid further processing).
2. keyword **SLAB**: Definition of the 2D system, a slab of given thickness (NL, number of atomic layers) parallel to the (hkl) crystallographic plane, with the ISUP-th atom on the surface layer

The **SLABCUT** option, combined with **ATOMINSE** (page 26), **ATOMDISP** (page 26), etc. can be used to create a slab of given thickness, with an atom (or group of atoms) adsorbed at given position. This is achieved by adding new atoms to the 2D structure, obtained after executing the **SLAB** option.

Test cases 5-6-7 refer to a 2D system; test cases 25-26-27 refer to the same system, but generated from the related 3D one. See also tests 35, 36, 37.

SLABINFO - 3D cell with z axis orthogonal to a given plane

rec	variable	meaning
• *	h,k,l	Crystallographic (Miller) indices of the basal layer of the new 3D unit cell

1. A new unit cell is defined, with two lattice vectors perpendicular to the [hkl] direction. The indices refer to the Bravais lattice of the crystal; the hexagonal lattice is used for the rhombohedral systems, the cubic lattice for cubic systems (non primitive).
2. A new Cartesian reference system is defined, with the *xy* plane parallel to the (hkl) plane.
3. The atoms in the reference cell are re-ordered according to their *z* coordinate, in order to recognize the layered structure, parallel to the (hkl) plane.
4. The layers of atoms are numbered. This information is necessary for generating the input data for the **SLABCUT** option.
5. After neighboring analysis, the program stops. If the keyword **ROTATE** was entered, execution continues. The shape of the new cell may be very different, computational parameters must be carefully checked.
6. the keyword **ORIGIN** can be used to shift the origin after the rotation of the cell, and minimize the number of symmetry operators with translational component. Useful to maximize the point group of the 2D system that can be generated from 3D using the keyword **SLABCUT** (page 42).

STOP

Execution stops immediately. Subsequent input records are not processed.

STRUCPRT

A formatted deck with cell parameters and atoms coordinates (bohr) in cartesian reference is written in the file STRUC.INCOOR . See appendix E, page 216.

SUPERCEL

rec	variable	meaning
• ∗	E	expansion matrix E (IDIMxIDIM elements, input by rows: 9 reals (3D); 4 reals (2D); 1 real (1D)

A supercell is obtained by defining the new unit cell vectors as linear combinations of the primitive cell unit vectors (use **SUPERCON** for conventional cell vectors reference). The point symmetry is defined by the number of symmetry operators in the new cell. It may be reduced, not increased.
The new translation vectors $\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3$ are defined in terms of the old vectors $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ and of the matrix E, read in input by rows, as follows:

$$\begin{aligned}
\mathbf{b'_1} &= e_{11} \cdot \mathbf{b}_1 + e_{12} \cdot \mathbf{b}_2 + e_{13} \cdot \mathbf{b}_3 \\
\mathbf{b'_2} &= e_{21} \cdot \mathbf{b}_1 + e_{22} \cdot \mathbf{b}_2 + e_{23} \cdot \mathbf{b}_3 \\
\mathbf{b'_3} &= e_{31} \cdot \mathbf{b}_1 + e_{32} \cdot \mathbf{b}_2 + e_{33} \cdot \mathbf{b}_3
\end{aligned}$$

The symmetry is automatically reduced to the point symmetry operators without translational components and a further reduction of the symmetry is also possible.
Before building the supercell, the origin is shifted in order to minimize the number of symmetry operators with translational components (see page 12). To avoid this operation, insert **NOSHIFT** before **SUPERCEL**
Atoms that are related by translational symmetry in the unit cell are considered inequivalent in a supercell.
The supercell option is a useful starting point for the study of defective systems, of chemisorption and anti ferromagnetism, by combining the **SUPERCEL**option with the options described in this chapter: **ATOMREMO** (page 27), **ATOMSUBS** (page 28), **ATOMINSE** (page 26), **ATOMDISP** (page 26), **SLAB** (page 42).
To study anti ferromagnetic (AFM) states, it may be necessary to generate a supercell, and then attribute different spin to atoms related by translational symmetry (**ATOMSPIN**, input block 3, page 57). See tests 17, 30, 31, 34, 37, 43, 47.

Example. Construction of supercells of face-centred cubic 3D system ($a = 5.42$ Å).
The crystallographic cell is non-primitive, the expansion matrix refers to primitive cell vectors. The E matrix has 9 elements:

```
 PRIMITIVE CELL
 DIRECT LATTICE VECTORS COMPONENTS
       X         Y         Z
B1     .000      2.710     2.710
B2     2.710     .000      2.710
B3     2.710     2.710     .000

   2 UNITS SUPERCELL (a)
       EXPANSION MATRIX                         DIRECT LATTICE VECTORS
E1      .000     1.000     1.000     B1     5.420     2.710     2.710
E2     1.000      .000     1.000     B2     2.710     5.420     2.710
E3     1.000     1.000      .000     B3     2.710     2.710     5.420

   2 UNITS SUPERCELL (b)
       EXPANSION MATRIX                         DIRECT LATTICE VECTORS
E1     1.000     1.000    -1.000     B1      .000      .000     5.420
E2      .000      .000     1.000     B2     2.710     2.710      .000
E3     1.000    -1.000      .000     B3    -2.710     2.710      .000

   4 UNITS SUPERCELL (c) crystallographic cell
       EXPANSION MATRIX                         DIRECT LATTICE VECTORS
E1    -1.000     1.000     1.000     B1     5.420      .000      .000
E2     1.000    -1.000     1.000     B2      .000     5.420      .000
E3     1.000     1.000    -1.000     B3      .000      .000     5.420
```

```
   8 UNITS SUPERCELL
       EXPANSION MATRIX                       DIRECT LATTICE VECTORS
 E1      2.000      .000      .000       B1      .000     5.420     5.420
 E2       .000     2.000      .000       B2     5.420      .000     5.420
 E3       .000      .000     2.000       B3     5.420     5.420      .000

  16 UNITS SUPERCELL
       EXPANSION MATRIX                       DIRECT LATTICE VECTORS
 E1      3.000    -1.000    -1.000       B1    -5.420     5.420     5.420
 E2     -1.000     3.000    -1.000       B2     5.420    -5.420     5.420
 E3     -1.000    -1.000     3.000       B3     5.420     5.420    -5.420

   27 UNITS SUPERCELL
       EXPANSION MATRIX                       DIRECT LATTICE VECTORS
 E1      3.000      .000      .000       B1      .000     8.130     8.130
 E2       .000     3.000      .000       B2     8.130      .000     8.130
 E3       .000      .000     3.000       B3     8.130     8.130      .000

  32 UNITS SUPERCELL
       EXPANSION MATRIX                       DIRECT LATTICE VECTORS
 E1     -2.000     2.000     2.000       B1    10.840      .000      .000
 E2      2.000    -2.000     2.000       B2      .000    10.840      .000
 E3      2.000     2.000    -2.000       B3      .000      .000    10.840
```

a), b) Different double cells
c) quadruple cell. It corresponds to the crystallographic, non-primitive cell, whose parameters are given in input (page 12).

Example. Construction of supercells of hexagonal $R\bar{3}$ (corundum lattice) cubic 3D system. The crystallographic cell is non-primitive: CRYSTAL refer to the primitive cell, with volume 1/3 of the conventional one. The E matrix has 9 elements:

```
GEOMETRY INPUT DATA:
LATTICE PARAMETERS  (ANGSTROMS AND DEGREES) - CONVENTIONAL CELL
    A         B         C        ALPHA        BETA       GAMMA
 4.76020   4.76020  12.99330    90.00000    90.00000   120.00000

TRANSFORMATION WITHIN CRYSTAL CODE FROM CONVENTIONAL TO PRIMITIVE CELL:

 LATTICE PARAMETERS  (ANGSTROMS AND DEGREES) - PRIMITIVE CELL
    A         B         C        ALPHA        BETA       GAMMA       VOLUME
 5.12948   5.12948   5.12948    55.29155    55.29155    55.29155    84.99223

   3 UNITS SUPERCELL crystallographic cell
       EXPANSION MATRIX                       DIRECT LATTICE VECTORS
 E1      1.000    -1.000      .000      B1     4.122    -2.380      .000
 E2       .000     1.000    -1.000      B2      .000     4.760      .000
 E3      1.000     1.000     1.000      B3      .000      .000    12.993

 LATTICE PARAMETERS (ANGSTROM AND DEGREES)
    A        B        C      ALPHA     BETA      GAMMA     VOLUME
 4.76020  4.76020 12.99330   90.000    90.000   120.000   254.97670
```

SUPERCON

When the crystallographic cell is non-primitive, a supercell is obtained by defining the new unit cell vectors as linear combinations of the *conventional cell* vectors.
See **SUPERCEL**, page 44 for input instructions and information.

SYMMDIR

The symmetry allowed directions, corresponding to internal degrees of freedom are printed.
No input data are required.

SYMMOPS

Point symmetry operator matrices are printed in the Cartesian representation. No input data are required.

SYMMREMO

All the point group symmetry operators are removed. Only the identity operator is left. The wave function can be computed. No input data are required.
Warning: the CPU time may increase by a factor MVF (order of point-group), both in the integral calculation and in the **scf** step. The size of the bielectronic integral file may increase by a factor MVF^2.

TENSOR

	rec	variable	meaning
•	$*$	IORD	order of the tensor (≤ 4)

This option evaluates and prints the non zero elements of the tensor of physical properties up to order 4.

TESTGEOM

Execution stops after reading the geometry input block and printing the coordinates of the atoms in the conventional cell. Neighbours analysis, as requested by the keyword **NEIGHBOR**, is not executed. The geometry input block must end with the keyword **END** or **ENDG**. No other input blocks (basis set etc) are required.

TRASREMO

Point symmetry operators with fractional translation components are removed. It is suggested to previously add the keyword **ORIGIN** (page 39), in order to minimize the number of symmetry operators with finite translation component. No input data are required.

USESAED

	rec	variable	meaning
•	$*$	δ(i),i=1,nsaed	δ for each distortion

Given the symmetry allowed elastic distortion (SAED), (printed by the keyword **MAKESAED**, page 36) δ for the allowed distortion are given in input.

2.2 Basis set input

Symmetry control			
ATOMSYMM	printing of point symmetry at the atomic positions	28	–
Basis set modification			
CHEMOD	modification of the electronic configuration	47	I
GHOSTS	eliminates nuclei and electrons, leaving BS	49	I
Auxiliary and control keywords			

CHARGED	allows non-neutral cell	47	–
NOPRINT	printing of basis set removed	49	–
PARAMPRT	printing of parameters controlling code dimensions	39	–
PRINTOUT	setting of printing options	40	I
SETINF	setting of inf array options	42	I
SETPRINT	setting of printing options	42	I
STOP	execution stops immediately	43	–
SYMMOPS	printing of point symmetry operators	46	–
END/ENDB	terminate processing of basis set definition keywords		–
Output of data on external units			
GAUSS98	printing of an input file for the GAUSS94/98 package	49	–

ATOMSYMM

See input block 1, page 28

CHARGED - charged reference cell

The unit cell of a periodic system must be neutral. This option forces the overall system to be neutral even when the number of electrons in the reference cell is different from the sum of nuclear charges, by adding a uniform background charge density to neutralize the charge in the reference cell.

Warning - Do not use for total energy comparison.

CHEMOD - modification of electronic configuration

rec	variable	meaning
•	∗ NC	number of configurations to modify
•	∗ LA	*label* of the atom with new configuration
	∗ CH(L),L=1,NS	shell charges of the LA-th atom. The number NS of shells must coincide with that defined in the basis set input.

The **CHEMOD** keyword allows modifications of the shell charges given in the basis set input, which are used in the atomic wave function routines. The original geometric symmetry is checked, taking the new electronic configuration of the atoms into account. If the number of symmetry operators should be reduced, information on the new symmetry is printed, and the program stops. No automatic reduction of the symmetry is allowed. Using the information printed, the symmetry must be reduced by the keyword **MODISYMM** (input block 1, page 36).
See test 37. MgO supercell, with a Li defect. The electronic configuration of the oxygen nearest to Li corresponds to O^-, while the electronic configuration of those in bulk MgO is O^{2-}. The basis set of oxygen is unique, while the contribution of the two types of oxygen to the initial density matrix is different.

END

Terminate processing of block 2, basis set, input. Execution continues. Subsequent input records are processed, if required.

GAUSS98 - Printing of input file for GAUSS98 package

The keyword **GAUSS98** writes in file GAUSSIAN.DAT an input deck to run Gaussian 94 (or Gaussian 98) [16, 29]. The deck can be prepared without the calculation of the wave function by entering the keyword **TESTPDIM** in input block 3 (page 79). For periodic systems, coordinates and basis set for all the atoms in the reference cell only are written (no information on translational symmetry).
If the keyword is entered many times, the data are overwritten. The file GAUSSIAN.DAT contains the data corresponding to the last call.
The utility program *gautocry* reads basis set input in Gaussian format (as prepared by ***http://www.emsl.pnl.gov/forms/basisform.html***) and writes it in CRYSTAL format. No input data required.

1. The route card specifies:

method	HF
basis set	GEN 5D 7F
type of job	SP
geometry	UNITS=AU GEOM=COORD

2. The title card is the same as in *CRYSTAL* input.

3. The molecule specification defines the molecular charge as the net charge in the reference cell. If the system is not closed shell, the spin multiplicity is indicated with a string "??", and must be defined by the user.

4. Input for effective core pseudopotentials is not written. In the route card PSEUDO = CARDS is specified; the pseudopotential parameters used for the crystal calculation are printed in the ***crystal*** output.

5. The scale factors of the exponents are all set to 1., as the exponents are already scaled.

6. the input must be edited when different basis sets are used for atoms with the same atomic number (e.g., CO on MgO, when the Oxygen basis set is different in CO and in MgO)

Warning: Only for 0D systems! The programs does not stop when the keyword **GAUSS94** is entered for 1-2-3D systems. Coordinates and basis set of all the atoms in the primitive cell are written, formatted, in file GAUSSIAN.DAT, following Gaussian 94 scheme.

Warning If you run Gaussian 98 using the input generated by CRYSTAL with the keyword GAUSS98 you do not obtain the same energy. There are 3 main differences between a standard CRYSTAL run and a GAUSSIAN run.

1. CRYSTAL adopts by default bypolar expansion to compute coulomb integrals when the two distributions do not overlap. To compute all 2 electron integrals exactly, insert keyword NOBIPOLA in input block 3;

2. CRYSTAL adopts truncation criteria for Coulomb and exchange sums: to remove them, in input block 3 insert:

   ```
   TOLINTEG
   20 20 20 20 20
   ```

3. CRYSTAL adopts the NIST conversion factor bohr/Angstrom CODATA98:
   1 Å= 0.5291772083 bohr
   To modify the value, in input block 1 insert:

   ```
   BOHRANGS
   value_of_new_conversion_factor
   ```

GHOSTS

rec	variable	meaning
• ∗	NA	number of atoms to be transformed into ghosts
• ∗	LA(L),L=1,NA	*label* of the atoms to be transformed.

Selected atoms may be transformed into *ghosts*, by deleting the nuclear charge and the shell electron charges, but leaving the basis set centred at the atomic position. The conventional atomic number is set to zero.
If the system is forced to maintain the original symmetry (**KEEPSYMM**), all the atoms symmetry related to the given one are transformed into ghosts.
Useful to create a vacancy (Test 37), leaving the variational freedom to the defective region and to evaluate the basis set superposition error (BSSE), in a periodic system. The periodic structure is maintained, and the energy of the isolated components computed, leaving the basis set of the other one(s) unaltered. For instance, the energy of a mono-layer of CO molecules on top of a MgO surface can be evaluated including the basis functions of the first layer of MgO, or, vice-versa, the energy of the MgO slab including the CO ad-layer basis functions.
See test36 and test37.

Warning Do not use with ECP.

Warning The keyword **ATOMREMO** (input block 1, page 27) creates a vacancy, removing nuclear charge, electron charge, and basis functions. The keyword **GHOSTS** creates a vacancy, but leaves the basis functions at the site, so allowing better description of the electron density in the vacancy.
Warning - Removal of nuclear and electron charge of the atoms selected is done after complete processing of the input. They look still as "atoms" in the printed output before that operation.

NOPRINT

Printing of basis set is removed. No input data required.

PARAMPRT - Printing of parametrized dimensions

See input block 1, page 39.

PRINTOUT - Setting of printing environment

See input block 1, page 40.

SETINF - Setting of INF values

See input block 1, page 42.

SETPRINT - Setting of printing options

See input block 1, page 42.

STOP

Execution stops immediately. Subsequent input records are not processed.

SYMMOPS

See input block 1, page 46

Effective core pseudo-potentials - ECP

rec	variable	value	meaning
• A	PSN		pseudo-potential keyword:
		HAYWLC	Hay and Wadt large core ECP.
		HAYWSC	Hay and Wadt small core ECP.
		BARTHE	Durand and Barthelat ECP.
		DURAND	Durand and Barthelat ECP.
		INPUT	free ECP - input follows.
		if PSN = INPUT insert	II
• *	ZNUC		effective core charge (ZN in eq. 2.7).
	M		Number of terms in eq. 2.8
	M0		Number of terms in eq. 2.9 for ℓ =0.
	M1		Number of terms in eq. 2.9 for ℓ =1.
	M2		Number of terms in eq. 2.9 for ℓ =2.
	M3		Number of terms in eq. 2.9 for ℓ =3.
		insert M+M0+M1+M2+M3 records	II
• *	ALFKL		Exponents of the Gaussians: $\alpha_{k\ell}$.
	CGKL		Coefficient of the Gaussians: $C_{k\ell}$.
	NKL		Exponent of the r factors: $n_{k\ell}$.

Valence-electron only calculations can be performed with the aid of effective core pseudo-potentials (ECP). The ECP input must be inserted into the basis set input of the atoms with conventional atomic number > 200.

The form of pseudo-potential W_{ps} implemented in *CRYSTAL* is a sum of three terms: a Coulomb term (C), a local term (W0) and a semi-local term (SL):

$$W_{ps} = C + W0 + SL \tag{2.6}$$

where:

$$C = -Z_N/r \tag{2.7}$$

$$W0 = \sum_{k=1}^{M} r^{n_k} C_k e^{-\alpha_k r^2} \tag{2.8}$$

$$SL = \sum_{\ell=0}^{3} [\sum_{k=1}^{M_\ell} r^{n_{k\ell}} C_{k\ell} e^{-\alpha_{k\ell} r^2}] P_\ell \tag{2.9}$$

Z_N is the effective nuclear charge, equal to total nuclear charge minus the number of electrons represented by the ECP, P_ℓ is the projection operator related to the ℓ angular quantum number, and M, n_k, α_k, M_ℓ, $n_{k\ell}$, $C_{k\ell}$, $\alpha_{k\ell}$ are atomic pseudo-potential parameters.

1. Hay and Wadt (HW) ECP ([30, 31]) are of the general form 2.6. In this case, the NKL value given in the tables of ref. [30, 31] must be decreased by 2 (2 → 0, 1 → -1, 0 → -2).

2. Durand and Barthelat (DB) ([32] - [33], [34], [35]), and Stuttgart-Dresden [36] ECPs contain only the Coulomb term C and the semi-local SL term.

3. In Durand and Barthelat ECP the exponential coefficient α in SL depends only on ℓ (i.e. it is the same for all the M_k terms).

$$SL = \sum_{\ell=0}^{3} e^{-\alpha_\ell r^2} [\sum_{k=1}^{M_\ell} r^{n_{k\ell}} C_{k\ell}] P_\ell \qquad (2.10)$$

The core orbitals replaced by Hay and Wadt *large core* and Durand-Barthelat ECPs are as follows:

Li-Ne	= [He]
Na-Ar	= [Ne]
first series	= [Ar]
second series	= [Kr]
third series	= [Xe]$4f^{14}$.

The core orbitals replaced by Hay and Wadt *small core* ECPs are as follows:

K-Cu	= [Ne]
Rb-Ag	= [Ar] $3d^{10}$
Cs-Au	= [Kr] $4d^{10}$.

The program evaluates only those integrals for which the overlap between the charge distribution $\varphi^0_\mu \ \varphi^g_\nu$ (page 174) and the most diffuse Gaussian defining the pseudopotential is larger than a given threshold T_{ps} (the default value is 10^{-5}). See also **TOLPSEUD** (Section 1.3).

Pseudopotential libraries

The following periodic tables show the effective core pseudo-potentials included as internal data in the *CRYSTAL* code.

```
HAY AND WADT LARGE CORE ECP. CRYSTAL92 DATA
 -------                                    -----------------
  Na Mg                                      Al Si P  S  Cl Ar
 -----------------------------------------------------------
  K  Ca Sc Ti V  Cr Mn Fe Co Ni Cu Zn Ga Ge As Se Br Kr
 -----------------------------------------------------------
  Rb Sr Y  Zr Nb Mo Tc Ru Rh Pd Ag Cd In Sn Sb Te I  Xe
 -----------------------------------------------------------
  Cs Ba    Hf Ta W  Re Os Ir Pt Au Hg Tl Pb Bi
 -----------------------------------------------------------
```

```
HAY AND WADT SMALL CORE ECP. CRYSTAL92 DATA

 ------------------------------------------------------------
  K  Ca Sc Ti V  Cr Mn Fe Co Ni Cu
 ------------------------------------------------------------
  Rb Sr Y  Zr Nb Mo Tc Ru Rh Pd Ag
 ------------------------------------------------------------
  Cs Ba    Hf Ta W  Re Os Ir Pt Au
 ------------------------------------------------------------
```

```
DURAND AND BARTHELAT'S LARGE CORE ECP - CRYSTAL92 DATA
 ------                                      -----------------
  Li Be                                      B  C  N  O  F  Ne
 ------                                      -----------------
  Na Mg                                      Al Si P  S  Cl Ar
 ------------------------------------------------------------
  K  Ca Sc Ti V  Cr Mn Fe Co Ni Cu Zn Ga Ge As Se Br Kr
 ------------------------------------------------------------
  Rb    Y                       Ag    In Sn Sb    I
 ------------------------------------------------------------
                                      Tl Pb Bi
 ------------------------------------------------------------
```

BARTHE, **HAYWSC** and **HAYWLC** pseudopotential coefficients and exponents are inserted as data in the **CRYSTAL** code. The data defining the pseudo-potentials where included in CRYSTAL92, and never modified. The keyword INPUT allows entering updated pseudo-potentials, when available. An *a posteriori* check has been possible for **HAYWLC** and **HAYWSC** only, as the total energy of the atoms for the suggested configuration and basis set has been published [30, 37]. Agreement with published atomic energies data is satisfactory (checked from Na to Ba) for Hay and Wadt small core and large core pseudo-potentials, when using the suggested basis sets. The largest difference is of the order of 10^{-3} hartree.
For Durand and Barthelat the atomic energies are not published, therefore no check has been performed. The printed data should be carefully compared with those in the original papers. The authors of the ECP should be contacted in doubtful cases.

Valence Basis set and pseudopotentials

Hay and Wadt ([30, 37]) have published basis sets suitable for use with their small and large core pseudopotentials. and in those basis set the s and p gaussian functions with the same quantum number have different exponent. It is common in *CRYSTAL* to use sp shells, where basis functions of *s* and *p* symmetry share the same set of Gaussian exponents, with a consequent considerable decrease in CPU time. The computational advantage of pseudopotentials over all-electron sets may thus be considerably reduced.
Basis set equivalent to those suggested by Hay and Wadt can be optimized by using *CRYSTAL* as an atomic package (page 56), or any atomic package with effective core pseudopotentials. See Chapter 5.2 for general comments on atomic basis function optimization. Bouteiller *et al* [38] have published a series of basis sets optimized for Durand and Barthelat ECPs.

Stuttgart-Dresden ECP (formerly STOLL and PREUSS ECP)

The most recent pseudopotential parameters, optimized basis sets, a list of references and guidelines for the choice of the pseudopotentials can be found at
http://www.theochem.uni-stuttgart.de/pseudopotentials/
These can be used in CRYSTAL via the INPUT keyword (basis set input, block2, page 50).

RCEP Stevens et al.

Conversion of Stevens et al. pseudopotentials An other important family of pseudopotentials for the first-, second-, third-, fourth and fifth-row atoms of the periodic Table (excluding the lanthanide series) is given by Stevens et al. [39, 40]. Analytic Relativistic Compact Effective Potential (RCEP) are generated in order to reproduce the "exact" pseudo-orbitals and eigenvalues as closely as possible. The analytic RCEP expansions are given by:

$$r^2 V_l(r) = \sum_k A_{lk} r^{n_{l,k}} e^{-B_{lk} r^2}$$

An example of data for Ga atom (Table 1, page 616 of the second paper) is:

	A_{lk}	n_{lk}	B_{lk}
V_d	-3.87363	1	26.74302
V_{s-d}	4.12472	0	3.46530
	260.73263	2	9.11130
	-223.96003	2	7.89329
V_{p-d}	4.20033	0	79.99353
	127.99139	2	17.39114

The corresponding Input file for the CRYSTAL program will be as follows:

```
INPUT
    21. 1 3 2 0 0
        26.74302    -3.87363 -1
         3.46530     4.12472 -2
         9.11130  260.73263  0
         7.89329 -223.96003  0
        79.99353     4.20033 -2
        17.39114  127.99139  0
```

Note that for the r-exponent (n_{lk}), -2 has been subtracted to the value given in their papers, as in the case of Hay and Wadt pseudopotentials.

2.3 Computational parameters, hamiltonian, SCF control

Single particle Hamiltonian

RHF	Restricted Closed Shell		74	–
UHF	Unrestricted Open Shell		80	–
DFT	DFT Hamiltonian		80	–
	SPIN	spin-polarized solution	59	–
	Choice of the exchange-correlation functionals			
	EXCHANGE	exchange functional	59	I
	LDA	Dirac-Slater [41] (LDA)		
	VBH	von Barth-Hedin [42] (LDA)		
	BECKE	Becke [43] (GGA)		
	PWGGA	Perdew-Wang 91 (GGA)		
	PBE	Perdew-Becke-Ernzerhof [44] (GGA)		
	CORRELAT	correlation functional	59	I
	VBH	von Barth-Hedin [42] (LDA)		
	PWGGA	Perdew-Wang 91 (GGA)		
	PBE	Perdew-Becke-Ernzerhof [44] (GGA)		
	PZ	Perdew-Zunger [45] (LDA)		
	PWLSD	Perdew-Wang 92 [46, 47, 48] (GGA)		
	VWN	Vosko,-Wilk-Nusair [49] (LDA)		
	P86	Perdew 86 [50] (LDA)		
	LYP	Lee-Yang-Parr [51] (GGA)		
	HYBRID	hybrid mixing	60	I
	NONLOCAL	local term parameterization	60	I
	B3PW	B3PW parameterization	60	–
	B3LYP	B3LYP parameterization	60	–
	Numerical accuracy control			
	[BECKE]	selection of Becke weights (default)		–
	SAVIN	selection of Savin weights		–
	RADIAL	definition of radial grid		I
	ANGULAR	definition of angular grid		I
	LGRID	"large" predefined grid		I
	XLGRID	"extra large" predefined grid		I
	TOLLDENS	density contribution screening 6		I
	TOLLGRID	grid points screening 14		I
	RADSAFE	safety radius for grid point screening		I
	BATCHPNT	grid point grouping for integration		I
	Atomic parameters control			
	RADIUS	customized atomic radius	65	I
	FCHARGE	customized formal atomic charge	65	I
	Auxiliary			
	PRINT	extended printing		
	END	close DFT input block		

Numerical accuracy and computational parameters control

BIPOLAR	Bipolar expansion of bielectronic integrals	58	I
BIPOSIZE	size of coulomb bipolar expansion buffer	58	I
EXCHSIZE	size of exchange bipolar expansion buffer	58	I
INTGPACK	Choice of integrals package [0]	70	I
NOBIPOLA	All bielectronic integrals computed exactly	73	–
POLEORDR	Maximum order of multipolar expansion [4]	73	I
TOLINTEG	Truncation criteria for bielectronic integrals [6 6 6 6 12]	80	I
TOLPSEUD	Pseudopotential tolerance [6]	80	I

Type of run

ATOMHF	Atomic wave functions	56	I
MPP	MPP execution (programmers only)	72	I
SCFDIR	SCF direct (mono+biel int computed)	74	–
NOMONDIR	SCF semidirect (mono on disk, biel computed)	73	–
EIGS	S(**k**) eigenvalues - basis set linear dependence check	65	–
FIXINDEX	Reference geometry to classify integrals	67	–

Integral file distribution

BIESPLIT	writing of bielectronic integrals in n files [$n = 1$],max=10	57	I
MONSPLIT	writing of mono-electronic integrals in n file [$n = 1$], max=10	72	I

Numerical accuracy control and convergence tools

ANDERSON	Fock matrix mixing	56	I
BROYDEN	Fock matrix mixing	58	I
FMIXING	Fock/KS matrix (cycle i and i-1) mixing [0]	69	I
LEVSHIFT	level shifter [no]	71	I
MAXCYCLE	maximum number of cycles [50]	72	I
SMEAR	Finite temperature smearing of the Fermi surface [no]	77	I
TOLDEE	convergence on total energy [5]	79	I
TOLDEP	convergence on density matrix [16]	79	I

Initial guess

EIGSHIFT	alteration of orbital occupation before SCF [no]	66	I
GUESSF	Fock/KS matrix from previous run	69	–
GUESSP	density matrix from a previous run	70	–
[**GUESSPAT**]	superposition of atomic densities	70	–

Spin-polarized system

ATOMSPIN	setting of atomic spin to compute atomic densities	57	I
BETALOCK	beta electrons locking	57	I
SPINLOCK	spin difference locking	78	I
SPINEDIT	editing of the spin density matrix	78	I

Auxiliary and control keywords

END	terminate processing of block3 input		–
KSYMMPRT	printing of Bloch functions symmetry analysis	71	–
NEIGHBOR	number of neighbours to analyse in **PPAN**	38	I
PARAMPRT	output of parameters controlling code dimensions	39	–
PRINTOUT	setting of printing options	40	I
NOSYMADA	No Symmetry Adapted Bloch Functions	73	–
SYMADAPT	Symmetry Adapted Bloch Functions (default)	79	–
SETINF	setting of inf array options	42	I
SETPRINT	setting of printing options	42	I
STOP	execution stops immediately	43	–
TESTPDIM	stop after symmetry analysis	79	–
TESTRUN	stop after integrals classification and disk storage estimate	79	–
Output of data on external units			
NOFMWF	wave function formatted output not written in file fort.98.	73	–
SAVEWF	wave function data written every two SCF cycles	74	–
Post SCF calculations			
POSTSCF	post-scf calculations when convergence criteria not satisfied	74	–
EXCHGENE	exchange energy evaluation (spin polarized only)	67	–
GRADCAL	analytical gradient of the energy	69	–
PPAN	population analysis at the end of the SCF [no]	74	

ANDERSON

Anderson's method [52], as proposed by Hamann [53], is applied. No input data are required. See test49_dft, a metallic Lithium 5 layers slab, PWGGA Hamiltonian.
MPP doesn't support Anderson mixing.

ATOMHF - Atomic wave function calculation

The Hartree-Fock atomic wave functions for the symmetry unique atoms in the cell are computed by the atomic program [6]. Full input (geometry, basis set, general information, SCF) is processed. No input data are required. The density matrix, constructed from a superposition of atomic densities, is computed and written on Fortran unit 9, along with the wave function information. The ***crystal*** program then stops. It is then possible to compute charge density (**ECHG**) and classical electrostatic potential (**CLAS**) maps by running the program ***properties***. This option is an alternative to the keyword **PATO** in the program ***properties*** (page 134), when the calculation of the periodic wave function is not required. The atomic wave function, eigenvalues and eigenvectors, can be printed by setting the printing option 71.

1. The atomic basis set may include diffuse functions, as no periodic calculation is carried out.

2. A maximum of two open shells of different symmetry (s, p, d) are allowed in the electronic configuration. In the electronic configuration given in input the occupation number of the shells must follow the rules given in Section 1.2.

3. For each electronic configuration, the highest multiplicity state is computed. Multiplicity cannot be chosen by the user.

Warning: DFT wave function for isolated atoms can not be computed.

ATOMSPIN - Setting of atomic spin

	rec	variable	meaning
•	∗	NA	number of atoms to attribute a spin
•	∗	LA,LS(LA),L=1,NA	*atom labels* and spin (1, 0, -1)

The setting of the atomic spins is used to compute the density matrix as superposition of atomic densities (**GUESSPAT** must be SCF initial guess); it does not work with **GUESSF** or **GUESSP**). The symmetry of the lattice may be reduced by attributing a different spin to geometrically symmetry related atoms. In such cases a previous symmetry reduction should be performed using the **MODISYMM** keyword. The program checks the symmetry taking the spin of the atoms into account. If the spin pattern does not correspond to the symmetry, the program prints information on the new symmetry, and then stops.
The formal spin values are given as follows:

`1`	atom spin is taken to be alpha;
`0`	atom spin is irrelevant;
`-1`	atom spin is taken to be beta.

In a NiO double-cell (four atoms, Ni1 Ni2 O1 O2) we might use:

atom	`Ni1`	`Ni2`		
spin	`1`	`1`	for starting ferromagnetic solutions:	↑ ↑
spin	`1`	`-1`	for starting anti ferromagnetic solutions:	↑ ↓

SPINLOCK forces a given $n_\alpha - n_\beta$ electrons value: to obtain a correct atomic spin density to start SCF process, the atomic spins must be set even for the ferromagnetic solution.
See test 30 and 31.

BETALOCK - Spin-polarized solutions

	rec	variable	meaning
•	∗	INF97	n_β electrons
	∗	INF98	number of cycles the n_β electrons is maintained

The total number of of β electrons at all **k** points can be locked at the input value. The number of α electrons is locked to $(N$ + INF95$)/2$, where N is the total number of electrons in the unit cell. INF95 must be odd when the number of electrons is odd, even when the number of electrons is even. See **SPINLOCK** for alternative way to define spin setting.

BIESPLIT - Splitting of large bielectronic integral files

	rec	variable	meaning
•	∗	NFILE	number of files to be used [1] (max 10)

Very compact crystalline systems, and/or very diffuse basis functions and/or very tight tolerances can produce billions integrals to be stored. The storage of bielectronic integrals can be avoided by running the direct SCF code **scfdir** rather than the standard SCF, at the expenses of a certain amount of CPU time.
When the standard SCF code is used, distributing the integrals on several disk files can improve performance.

BIPOLAR - Bipolar expansion approximation control

	rec	variable	meaning
•	*	ITCOUL	overlap threshold for Coulomb 14
	*	ITEXCH	overlap threshold for exchange 10

The bipolar approximation is applied in the evaluation of the Coulomb and exchange integrals (page 178). ITCOUL and ITEXCH can be assigned any intermediate value between the default values (14 and 10) (see page 178) and the values switching off the bipolar expansion (20000 and 20000).

BIPOSIZE -Size of buffer for Coulomb integrals bipolar expansion

	rec	variable	meaning
•	*	ISIZE	size of the buffer in words

Size (words) of the buffer for bipolar expansion of Coulomb integrals (default value 100000. The size of the buffer is printed in the message:

```
BIPO BUFFER LENGTH (WORDS) = XXXXXXX
or
COULOMB BIPO BUFFER TOO SMALL - TO AVOID I/O SET BIPOSIZE = XXXXXX
```

BROYDEN

	rec	variable	meaning
•	*	W0	*W0* parameter in Anderson's paper [54]
	*	IMIX	percent of Fock/KS matrices mixing when Broyden method is switched on
	*	ISTART	SCf iteration after which Broyden method is active (minimum 2)

A modified Broyden [55] scheme, following the method proposed by Johnson [54], is applied after the ISTART SCF iteration, with IMIX percent of Fock/KS matrices simple mixing. The value of % mixing given in input after the keyword **FMIXING** is overridden by the new one. Level shifter should be avoided when Broyden method is applied.
Suggested values:

```
FMIXING
80
BROYDEN
0.0001 50 2
```

MPP doesn't support Broyden mixing.
See test50_dft, a metallic Lithium 5 layers slab, PWGGA Hamiltonian.

END

Terminate processing of block 3,(last input block). Execution continues. Subsequent input records are not processed.

DFT

The Kohn-Sham [56, 57] DFT code is controlled by keywords, that must follow the general keyword **DFT**, in any order. These keywords can be classified into four groups:

1 Choice of the exchange-correlation functionals

2 *Integration grid and numerical accuracy control (optional)*

3 *DF energy gradient (optional)*

4 *Atomic parameters (optional)*

The DFT input block ends with the keyword **END** or **ENDDFT**. Default values are supplied for all computational parameters. Choice of exchange and/or correlation potential is mandatory.

1. Choice of the exchange-correlation functionals

EXCHANGE and **CORRELAT** keywords, each followed by an alpha-numeric record, allow the selection of the exchange and correlation functionals.
If the correlation potential is not set (keyword **CORRELAT**), an exchange-only potential is used in the Hamiltonian. If the exchange potential is not set (keyword **EXCHANGE**), the Hartree-Fock potential is used.

CORRELAT Correlation Potential (default: no correlation).

Insert one of the following keywords II

PZ	LSD. Perdew-Zunger parameterization of the Ceperley-Alder free electron gas correlation results [45]
PWLSD	LSD. Perdew-Wang parameterization of the Ceperley-Alder free electron gas correlation results [48]
VWN	LSD. Vosko-Wilk-Nusair parameterization of the Ceperley-Alder free electron gas correlation results [49]
VBH	LSD. von Barth-Hedin [42]
P86	GGA. Perdew 86 [50]
WCGGA	GGA - Wu-Cohen [58]
PWGGA	GGA. Perdew-Wang [59]
LYP	GGA. Lee-Yang-Parr [51]
PBE	GGA. Perdew-Burke-Ernzerhof [44]

EXCHANGE Exchange potential (default: Hartree-Fock exchange).

Insert one of the following keywords II

LDA	LSD. Dirac-Slater [41]
VBH	LSD. von Barth-Hedin [42]
BECKE	GGA. Becke [43]
PWGGA	GGA. Perdew-Wang [59]
PBE	GGA. Perdew-Becke-Ernzerhof [44]

All functionals are formulated in terms of total density and spin density. Default is total density. To use functionals of spin density insert the keyword **SPIN**.

SPIN unrestricted spin DF calculation (default: restricted)

It is also possible to incorporate part of the exact Hartree-Fock exchange into the exchange functional through the keyword **HYBRID**. Any mixing (0-100) of exact Hartree-Fock and DFT exchange can be used.

NONLOCAL allows modifying the relative weight of the local and non-local part both in the exchange and the correlation potential with respect to standard definition of GGA type potentials.

HYBRID	Hybrid method - 1 record follows:
• ∗ A	Fock exchange percentage (default 100.)
NONLOCAL	setting of non-local weighting parameters - 1 record follows:
• ∗ B	exchange weight of the non-local part of exchange
C	weight of the non-local correlation
B3PW	Becke's 3 parameter functional [60] combined with the non-local correlation PWGGA [61, 46, 47, 48]
B3LYP	Becke's 3 parameter functional [60] combined with the non-local correlation LYP
PBE0	[62]

B3PW and **B3LYP** are global keywords, defining hybrid exchange-correlation functionals completely. They replace the following sequences:

B3PW	**B3LYP**
corresponds to the sequence:	corresponds to the sequence:
EXCHANGE	EXCHANGE
BECKE	BECKE
CORRELAT	CORRELAT
PWGGA	LYP
HYBRID	HYBRID
20	20
NONLOCAL	NONLOCAL
0.9 0.81	0.9 0.81

B3LYP in CRYSTAL is based on the 'exact' form of the Vosko-Wilk-Nusair correlation potential (corresponds to a fit to the Ceperley-Alder data). In the original paper [49]) it is reported as functional V, which is used to extract the 'local' part of the LYP correlation potential.
The Becke's 3 parameter functional can be written as:

$$E_{xc} = (1 - A) * (E_x^{LDA} + B * E_x^{BECKE}) + A * E_x^{HF} + (1 - C) * E_c^{VWN} + C * E_c^{LYP}$$

A, B, and C are the input data of **HYBRYD** and **NONLOCAL**.

Examples of possible selection of the correlation and exchange functionals are:

exchange	correlation	
——	PWGGA	Hartree-Fock exchange, non local Perdew-Wang correlation.
LDA	VWN	probably the most popular LDA formulation
VBH	VBH	was the most popular LDA scheme in the early LDA solid state applications (1975-1985).
PWGGA	PWGGA	- - - - - -
BECKE	LYP	- - - - - -

2. Integration grid and numerical accuracy control

No input data are required: Becke weights are chosen by default, as well as a set of safe values for the computational parameters of integration.
The generation of grid points in CRYSTAL is based on an atomic partition method, originally developed by Becke [63] for molecular systems and then extended to periodic systems [64]. Each atomic grid consists of a radial and an angular distribution of points. Grid points are generated through a radial and an angular formula: Gauss-Legendre radial quadrature and Lebedev two-dimensional angular point distribution are used.
Lebedev angular grids are classified according to progressive accuracy levels, as given in the following table:

LEV	CR98	ℓ	N_{ang}		LEV	CR98	ℓ	N_{ang}
1	1	9	38		16		53	974
2	2	11	50		17		59	1202
3		13	74	*	18		65	1454
4		15	86		19		71	1730
5	3	17	110		20		77	2030
6		19	146		21		83	2354
7		21	170		22		89	2702
8	4	23	194		23		95	3074
9		25	230	*	24		101	3470
10	5	27	266	*	25		107	389
11	6	29	302		26		113	4334
12		31	350		27		119	4802
13	7	35	434		28		125	5294
14		41	590		29		131	5810
15		47	770					

Index of Lebedev accuracy levels
- LEV: Lebedev accuracy level
- CR98: corresponding index in CRYSTAL98
- ℓ: maximum quantum number of spherical harmonics used in Lebedev derivation
- N_{ang}: number of angular points generated per radial point
- * : sets with negative weights, to be avoided

If one Lebedev accuracy level is associated with the whole radial range, the atomic grid is called *unpruned*, or *uniform.* In order to reduce the grid size and maintain its effectiveness, the atomic grids of spherical shape can be partitioned into shells, each associated with a different angular grid. This procedure, called grid *pruning*, is based on the assumption that core electron density is usually almost spherically symmetric, and surface to be sampled is small.
Also, points far from the nuclei need lower point density, as associated with relatively small weights, so that more accurate angular grids are mostly needed within the valence region than out of it.
The choice of a suitable grid is crucial both for numerical accuracy and need of computer resources.

Different formulae have been proposed for the definition of grid point weights. In CRYSTAL Becke and Savin weights are available; Becke weights are default, and provide higher accuracy.

[BECKE] Becke weights [65]. Default choice.

SAVIN Savin weights [66]

A default grid is available in CRYSTAL, however the user can redefine it by the following keywords:

RADIAL	Radial integration information
rec variable	meaning
• ∗ NR	number of intervals in the radial integration [default 1]
• ∗ RL(I),I=1,NR	radial integration intervals limits in increasing sequence [default 4.0] (last limit is set to ∞)
• ∗ IL(I),I=1,NR	number of points in the radial quadrature in the I-th interval [default 55].

ANGULAR	Angular integration information
rec variable	meaning
• ∗ NI	number of intervals in the angular integration [default 1]
• ∗ AL(I),I=1,NI	upper limits of the intervals in increasing sequence. The last limit must be 9999.0 [default 9999.0]
• ∗ LEV(I),I=1,NI	accuracy level in the angular integration over the I-th interval; positive for Lebedev level (see Lev in page 61) [default 13]

The *default grid* is a pruned (55,434) grid, having 55 radial points and a maximum number of 434 angular points in regions relevant for chemical bonding. Each atomic grid is split into ten shells with different angular grids.
This grid is good enough for either single-point energy calculations or medium-accuracy geometry optimizations. Due to the large pruning, the cost of the calculation is modest.

Default grid - corresponds to the sequence:

`RADIAL`	Keyword to specify the radial grid
`1`	Number of intervals in the radial part
`4.0`	Radial integration limits of the i-th interval
`55`	Number of radial points in the i-th interval
`ANGULAR`	Keyword to specify the angular grid
`10`	Number of intervals in the angular part
`0.4 0.6 0.8 0.9 1.1 2.3 2.4 2.6 2.8 9999.0`	Angular integration limits of the i-th interval
`1 2 5 8 11 13 11 8 5 1`	Angular grid accuracy level of the i-th interval

Information on the size of the grid, grid thresholds, and radial (angular) grid is reported in the CRYSTAL output with the following format:

```
SIZE OF GRID=      40728
BECKE WEIGHT FUNCTION
RADSAFE =     2.00
TOLERANCES - DENSITY:10**- 6; POTENTIAL:10**- 9; GRID WGT:10**-14

RADIAL INTEGRATION  - INTERVALS (POINTS,UPPER LIMIT):           1( 55,  4.0*R)

ANGULAR INTEGRATION - INTERVALS (ACCURACY LEVEL [N. POINTS] UPPER LIMIT):
  1(  1[  38]   0.4)   2(  2[  50]   0.6)   3(  5[ 110]   0.8)   4(  8[ 194]   0.9)
  5( 11[ 302]   1.1)   6( 13[ 434]   2.3)   7( 11[ 302]   2.4)   8(  8[ 194]   2.6)
  9(  5[ 110]   2.8)  10(  1[  38]9999.0)
```

Two more pre-defined grids are available which can be selected to improve accuracy by inputing the following global keywords:

LGRID Large grid

Global keyword to choose a larger grid than default, corresponding to the sequence:

```
RADIAL
1
4.0
75
ANGULAR
5
0.1667 0.5 0.9 3.05 9999.0
2 6 8 13 8
```

The *large grid* is a pruned (75,434) grid, having 75 radial points and a maximum number of 434 angular points in the region relevant for chemical bonding. Five shells with different angular points are adopted to span the radial range as proposed by Gill et al. [67]. Due to a larger number of radial points and less aggressive pruning, this grid gives more accurate results than the default grid. It is recommended for high-accuracy energy calculations and geometry optimizations. It is also recommended for periodic systems containing second-row and third-row atoms (transition metals).

XLGRID Extra large grid

Global keyword to choose an even larger integration grid, corresponding to the sequence:

```
RADIAL
1
4.0
75
ANGULAR
5
0.1667 0.5 0.9 3.5 9999.0
4 8 12 16 12
```

The *extra-large grid* is a pruned (75,974) grid, consisting of 75 radial points and 974 angular points in the region of chemical interest. This is a very accurate grid and is recommended when numerical derivatives of energy or related properties (i.e. spontaneous polarization) and gradients have to be computed (e.g. bulk modulus, elastic constants, piezoelectric tensor, ferroelectric transitions). It is also recommended for heavy atoms (fourth-row and heavier).

XXLGRID Extra large grid

Very large grid use for benchmark calculations. It corresponds to:

```
RADIAL
1
4.0
99
ANGULAR
5
0.1667 0.5 0.9 3.5 9999.0
6 10 14 18 14
```

DISTGRID developers only

This forces the code to distribute the DFT grid across the available processors. Only parallel jobs are affected.

REPLGRID	developers only

This forces the code to replicate the DFT grid on the available processors. Only parallel jobs are affected.
(This directive at present has no effect - the defaults is a replicated grid. However I am thinking of changing the default for MPP jobs to distgrid, and then replgrid will have a use for certain medium sized jobs where load balancing is an issue).IJB

Unpruned grids

To switch from a pruned grid to the corresponding unpruned grid, only one shell must be defined in the radial part and the same angular accuracy is used everywhere. The use of unpruned grids increases the cost of the calculations by about 50-60% with respect to the pruned grid.
For example, to transform the default grid to the corresponding unpruned grid input the following data:

```
ANGULAR
1
9999.0
13
```

Numerical accuracy and running time are also controlled by the following keywords:

TOLLGRID	
• ∗ IG	DFT grid weight tolerance [default 14]
TOLLDENS	
• ∗ ID	DFT density tolerance [default 6]

The DFT density tolerance ID controls the level of accuracy of the integrated charge density N_{el} (number of electron per cell):

$$N_{el} = \int_{cell} \rho(\mathbf{r}) d\mathbf{r} = \sum_{\mu,\nu,\mathbf{g},\mathbf{l}} P_{\mu,\nu}^{\mathbf{g}+\mathbf{g}'} \sum_i w(\mathbf{r}_i) \varphi_\mu^{\mathbf{g}}(\mathbf{r}_i) \varphi_\nu^{\mathbf{g}'}(\mathbf{r}_i)$$

all contributions where $|\varphi_\mu(\mathbf{r}_i)| < 10^{-ID}$ or $|\varphi_\nu(\mathbf{r}_i)| < 10^{-ID}$ are neglected (see Chapter 8.11 for notation). The default value of ID is 6.
Grid points with integration weights less than 10^{-IG} are dropped. The default value of IG is 14.

RADSAFE	
• ∗ RAD	for developers only [default 2]

BATCHPNT	
• ∗ BATCH	average number of points in a batch for numerical integration [default 100]

Default value of BATCH is 100. In the calculation of the exchange-correlation contribution to the Kohn-Sham hamiltonian matrix elements and energy gradients, the grid is partitioned into batches of points as suggested by Ahlrichs [68]. However, in CRYSTAL the number of points per batch is not constant, as it depends on point density, so that BATCH does not correspond to the maximun number of points in a batch. As a consequence, in special cases, memory requirement may become huge and cause problems in dynamic allocation at running time.

When the program runs out of memory, it stops with the following error message:

```
ERROR *** sub_name *** array_name ALLOCATION
```

where `array_name` is one of the following:

```
DF0 KSXC1 KSXC2 KSXC2Y KSXC2Z DFXX DFYY DFZZ DFXY DFYZ DFXZ
RHO FRHO AXJ,AYJ,AZJ,VGRID GRAZ GRAY GRAZ
```

In these cases it is recommended that the value of BATCH be reduced, although this may result in some degree of inefficiency (minimum value: 1).

CHUNKS

• ∗	NCHU	maximum number of points allowed in a batch for numerical integration [default 10000000]

3. DF energy gradient

[NEWTON]

The current default when computing DFT analytical gradients in CRYSTAL is to include weight derivatives. Weight derivatives are mandatory when low quality grids are adopted.

4. Atomic parameters

The radius attributed to each atom for the integration is computed from the nuclear charge and the net charge. It is possible to enter for selected atoms a given atomic radius or a formal charge.

No check on symmetry requirements is performed. If the selected atoms has other atoms symmetry related, radius (or charge) of those atoms must be defined. The keyword **ATOMSYMM** inserted in the first input block (geometry) prints the set of atoms symmetry related.

• A	**RADIUS**	
• ∗	NUMAT	number of atoms selected
	insert NUMAT records	II
• ∗	LB	*label* of the atom
	RAD(LB)	radius (Å) attributed to the atom

• A	**FCHARGE**	
• ∗	NUMAT	number of atoms selected
	insert NUMAT records	II
• ∗	LB	*label* of the atom
	FCH(LB)	formal charge attributed to the atom

EIGS - Check of basis set linear dependence

In order to check the risk of basis set linear dependence, it is possible to calculate the eigenvalues of the overlap matrix. Full input (geometry, basis set, general information, SCF) is processed. No input data are required. The overlap matrix in reciprocal space is computed at all the **k**-points generated in the irreducible part of the Brillouin zone, and diagonalized. The eigenvalues are printed.

The higher the numerical accuracy obtained by severe computational conditions, the closer to 0 can be the eigenvalues without risk of numerical instabilities. Negative values indicate numerical linear dependence. The ***crystal*** program stops after the check (even if negative eigenvalues are not detected).

The Cholesky reduction scheme [69], adopted in the standard SCF route, requires linearly independent basis functions.

MPP doesn' support EIGS.

EIGSHIFT - Alteration of orbital occupation before SCF

rec	variable	meaning
• ∗	NORB	number of elements to be shifted
		> 0 level shift of diagonal elements only
		< 0 off-diagonal level shift
		insert NORB records
		if NORB > 0
• ∗	IAT	label of the atom
	ISH	sequence number of the shell in the selected atom Basis Set (as given in Basis Set input)
	IORB	sequence number of the AO in the selected shell (see Section 1.2, page 16).
	SHIF1	α (or total, if Restricted) Fock/KS matrix shift
	[SHIF2	β Fock matrix shift - spin polarized only]
		if NORB < 0
• ∗	IAT	label of the atom
	ISH	sequence number of the shell in the selected atom Basis Set
	IORB1	sequence number of the AO in the selected shell
	IORB2	sequence number of the AO in the selected shell
	SHIF1	α (or total, if Restricted) Fock/KS matrix shift
	[SHIF2	β Fock matrix shift - spin polarized only]

Selected diagonal Fock/KS matrix elements can be shifted upwards when computing the initial guess, to force orbital occupation. This option is particularly useful in situations involving d orbital degeneracies which are not broken by the small distortions due to the crystal field, but which are broken by some higher-order effects (e.g. spin-orbit coupling). The **EIGSHIFT** option may be used to artificially remove the degeneracy in order to drive the system to a stable, non-metallic solution. The eigenvalue shift is removed after the first SCF cycle.
If the shift has to be applied to matrix elements of atoms symmetry related, the input data must be repeated as many times as the atoms symmetry related.

Example: $KCoF_3$ (test 38). In the cubic environment, two β electrons would occupy the three-fold degenerate t_{2g} bands. A state with lower energy is obtained if the degeneracy is removed by a tetragonal deformation of the cell (keyword **ELASTIC**), and the d_{xy} orbital (see page 16 for d orbital ordering) is shifted upwards by 0.3 hartree.
Warning EIGSHIFT acts on the atoms as specified in input. If there are atoms symmetry-related to the chosen one, hamiltonian matrix elements shift is not applied to the others. The programs checks the symmetry compatibility, and, if not satisfied, stops execution. The matrix elements of all the atoms symmetry-related must be shifted, if the symmetry of the systems must be kept
The keyword **UTMOST** prints information on the atoms symmetry related in the cell.

EIGSHROT

Consider now the case of CoF_2. The first six neighbors of each Co^{2+} ion form a slightly distorted octahedron (2 axial and 4 equatorial equivalent distances); also in this case, then, we are interested in shifting upwards the d_{xy} orbital, in order to drive the solution towards the following occupation:

α: all five d orbitals
β: d_{xz} and d_{yz}

The principal axis of the CoF_6 octahedron, however, is not aligned along the z direction, but lies in the xy plane, at 45^0 from the x axis. The cartesian reference frame must then be reoriented before the shift of the d_{xy} orbital.
To this aim the option **EIGSHROT** must be used. The reoriented frame can be specified in two ways, selected by a keyword:

rec		variable	meaning
•		**MATRIX**	keyword - the rotation matrix R is provided
•	∗	R11 R12 R13	first row of the matrix.
•	∗	R21 R22 R23	second row of the matrix.
•	∗	R31 R32 R33	third row of the matrix.
			or
•		**ATOMS**	keyword - the rotation is defined by three atoms of the crystal
•	∗	IA	*label* of first atom in the reference cell
		AL,AM,AN	indices (direct lattice, input as reals) of the cell where the first atom is located
•	∗	IB	*label* of second atom in the reference cell
		BL,BM,BN	indices (direct lattice, input as reals) of the cell where the second atom is located
•	∗	IC	*label* of third atom in the reference cell
		CL,CM,CN	indices (direct lattice, input as reals) of the cell where the third atom is located
			insert **EIGSHIFT** input records (Section 2.3, page 66)

When the rotation is defined by three atoms, the new reference frame is defined as follows :
Z-axis from atom 2 to atom 1
X-axis in the plane defined by atoms 1-2-3
Y-axis orthogonal to Z- and X-axis

Notice that the wave function calculation is performed in the original frame: the aim of the rotation is just to permit a shift of a particular orbital. An equivalent rotation of the eigenvectors can be obtained in ***properties*** by entering the keyword **ROTREF**, so allowing AO projected Density of States according to the standard orientation of the octahedron.
Example:

```
CoF2 example
```

EXCHGENE - Exchange energy calculation

In RHF calculations Coulomb and exchange integrals are summed during their calculation, and there is no way to separate the exchange contribution to the total energy. In UHF/ROHF calculations, this option allows the independent calculation and printing of the exchange contribution to the total energy. See equation 8.19, page 177.
No input data are required. See tests 29, 30, 31, 38.

EXCHSIZE - Size of buffer for exchange integrals bipolar expansion

rec		variable	meaning
•	∗	ISIZE	size of the buffer in words

Size (words) of the buffer for bipolar expansion of exchange integrals (default value 100000). The size of the buffer is printed in the message:

```
EXCH. BIPO BUFFER:  WORDS USED = XXXXXXX
or
EXCH. BIPO BUFFER TOO SMALL - TO AVOID I/O SET EXCHSIZE = XXXXXX
```

FIXINDEX - Geometry and basis set optimization tools

No input data required.
When the geometrical and/or the basis set parameters of the system are changed, **maintaining the symmetry and the setting**, the truncation criteria of the Coulomb and exchange series, based on overlap (Chapter 8) can lead to the selection of different numbers of bielectronic integrals. This may be the origin of numerical noise in the optimization curve. When small

changes are made on the lattice parameter or on the Gaussian orbital exponents, the indices of the integrals to be calculated can be selected for a reference geometry (or basis set), "frozen", and used to compute the corresponding integrals with the modified geometry (or basis set). This procedure is recommended only when basis set or geometry modifications are relatively small.

- The corresponding irreducible atoms in the two geometries must be entered in the same order, and their position in the second geometry must be slightly shifted in comparison with the first geometry (reference);
- the reference geometry must correspond to the most compact structure;
- the reference basis set must have the lowest outer exponent.

This guards against the loss of significant contributions after, for example, expansion of the lattice.
If estimate of resource is requested with TESTRUN, the reference geometry is used.
Two sets of input data must be given:

1. complete input (geometry, Section 1.1; basis set, Section 1.2; general information, Section 1.3; SCF, Section 1.3), defining the reference basis set and/or geometry;
2. "restart" option input, selected by one of the following keywords (format A) to be added after the SCF input:

GEOM	restart with new geometrical parameters
	insert geometry input, page 9
	or
BASE	restart with new basis set
	insert basis set input, page 14
	or
GEBA	restart with new basis set and new geometrical parameters
	insert geometry input, page 9
	insert basis set input, page 14

BASE: the only modification of the basis set allowed is the value of the orbital exponent of the GTFs and the contraction coefficient; the number and type of shells and AOs cannot change.

GEOM: geometry variation *must* keep the symmetry and the setting unchanged.

The resulting structure of the input deck is as follows:

0 Title

1 standard geometry input (reference geometry). Section 1.1

1b *geometry editing keywords (optional; Section 2.1)*

END

2 standard basis set input (reference basis set). Section 1.2

2b *basis set related keywords (optional; Section 2.2)*

END

3 **FIXINDEX**

3 *block3 keywords*

END

GEOM	BASE	GEBA
⇓	⇓	⇓
geometry input(p 9)	basis set input(p 14)	geometry input (p 9)
(block 1, *1b*)	(block 2, *2b*)	(block 1, *1b*)
END	**END**	**END**
		basis set input(p 14)
		(block 2, *2b*)
		END

Warning: The reference geometry and/or basis set is overwritten by the new one, after symmetry analysis and classification of the integrals. If the reference geometry is edited through appropriate keywords, *the same editing* **must** be performed through the second input. Same for basis set input.

If the geometry is defined through the keyword EXTERNAL, the reference geometry data should be in file fort.34, the wave function geometry in file fort.35.

See tests 5 and 20.

FMIXING - Fock/KS matrix mixing

rec	variable	meaning
• ∗	IPMIX	percent of Fock/KS matrices mixing

The Fock/KS matrix at cycle i is defined as:

$$F_i' = (1-p)F_i + pF_{i-1}'$$

where p, input datum IPMIX, is the % of mixing. Too high a value of p (>50%) causes higher number of SCF cycles and can force the stabilization of the total energy value, without a real self consistency.

GRADCAL

No input data required.
Analytic calculation of the nuclear coordinates gradient of the HF, UHF, DFT energies after SCF (all electrons and ECP).
If numerical gradient is requested for the geometry optimization (**NUMGRALL**, page 89; **NUMGRATO**, page 89; **NUMGRCEL**, page 89;), analyical gradient is not computed.

GUESSF - Fock/KS matrix from a previous run

The Fock/KS matrix F^0 (direct lattice) is read from disk (from file fort.20), and diagonalized (after Fourier transformation to the reciprocal lattice), to compute the first cycle density matrix. The data set containing F^0 is written in file fort.9 at the end of a previous SCF run. No input data are required. When wave function information are stored formatted in file fort.98, the data must be converted to binary by the keyword **RDFMWF**, page 122 of the ***properties*** program). The two cases, the present one and that used for the restart, must have the same symmetry, and the same number of atoms, basis functions and shells. Atoms and shells must be in the same order. The program does not check the 1:1 old–new correspondence. Different geometrical parameters, computational conditions or exponents of the Gaussian primitives are allowed. In geometry and/or basis set optimization, this technique will significantly reduce the number of SCF cycles. The following scheme shows how to proceed.

1. First run to generate the Fock/KS matrix

Program	inp. block	section	comments
crystal	0	1	Title
	1	1.1	geometry input
	2	1.2	basis set input
	3	1.3	general information
	4	1.3	scf input
	save wf in file fort.9 (binary) or file.98 (formatted)		

2. Second run - the Fock/KS matrix is read in as a guess to start scf

copy file fort.9 to fort.20 (or convert file fort.98 and then copy)			
Program	inp. block	section	comments
crystal	0	1	Title
	1	1.1	geometry input
	2	1.2	basis set input
	3	1.3	general information input **FIXINDEX**
	4	1.3	scf input (**GUESSF**)
	1b	1.1	geometry input present case

Warning The modification of the geometry may result in a different order in the storage of the Fock/KS matrix elements associated to each overlap distribution in the present and the previous run. To avoid the mismatch it is strongly recommended to classify the integrals of the present case using the geometry of the previous case (**FIXINDEX**, page 82).

GUESSP - Density matrix from a previous run

The density matrix P^0 (direct lattice) is read from disk (from file fort.20) to start the SCF cycles. Same procedure as for **GUESSF**. No input data are required.
The density matrix can be edited to modify the spin state. See **SPINEDIT**, page 78.

GUESSPAT - Superposition of atomic densities

The standard initial guess to start the SCF cycle is the superposition of atomic (or ionic) densities. No input data are required. The electronic configuration of the atoms is entered as a shell occupation number in the basis set input (page 14). Different electronic configurations may be assigned to atoms with the same atomic number and basis set (but not symmetry related) through the keyword **CHEMOD** (page 14).

INTGPACK - Choice of bielectronic integrals package

rec	variable	value	meaning
• ∗	IPACK	[0]	*s, sp* shells→ POPLE; *p, d* shells → ATMOL
		1	ATMOL for Coulomb integrals;
			POPLE for exchange integrals
		2	POPLE for Coulomb integrals;
			ATMOL for exchange integrals
		3	ATMOL for Coulomb integrals;
			ATMOL for exchange integrals

By default the bielectronic integrals are computed using a set of routines derived from Pople's GAUSSIAN 70 package [7], if *s* and *sp* shells are involved, and by routines derived from ATMOL [9] for *p* and *d* shells. The value of IPACK allows different choices. Integrals involving *p* or *d* shells are always computed by ATMOL. The ATMOL package can compute integrals over functions of any quantum number, but the symmetry treatment implemented in the CRYSTAL package allows usage of *s, p* and *d* functions only. The use of *sp* shells (*s* and *p* orbitals sharing the same exponent) reduces the time required to compute the integrals considerably.

KSYMMPRT

Symmetry Adapted Bloch Functions [70, 71] (page 79)are used as basis for the Fock matrix diagonalization. The results of the symmetry analysis in reciprocal space are printed. At each **k**-point: number of point symmetry operators, number of active IRs, maximum IR dimension and maximum block dimension in the Fock matrix factorization. TESTRUN stops execution after this information is printed.
No input data required.
Extended information can be obtained by setting the value N of LPRINT(47) (keyword **SET-PRINT**, page 42) before **KSYMMPRT**.

N	information
0	Basic Symmetry Information - At each **k**-point: list of point symmetry operators, IR dimensions and number of Irreducible Sets.
> 0	Symmetry Information - At each **k**-point $\leq$ N: class structure, character table and IR information concerning the K-Little Group. For the rest of the **k**-point the same information as -1 is printed.
< -1	Full Symmetry Information - At each **k**-point: the same information as N > 0, together with the matrix representatives of the point operators.

MPP doesn't support KSYMMPRT.

LEVSHIFT - Eigenvalue level shifting

rec	variable	value	meaning
• ∗	ISHIFT		The level shifter is set to ISHIFT *0.1 hartree.
	ILOCK	0	no locking
		1	causes a lock in a particular state (eg non-conducting) even if the solution during the SCF cycles would normally pass through or even converge to a conducting state.

The eigenvalue level shifting technique is well known in molecular studies [72, 73], and may also be used for periodic systems. The technique involves the addition of a negative energy shift to the diagonal Fock/KS matrix elements (in the Crystalline Orbital basis) of the *occupied* orbitals and thus reducing their coupling to the "unoccupied" set. This shift may be maintained (ILOCK=1) or removed (ILOCK=0) after diagonalization. The former case causes a lock in a particular state (eg non- conducting) even if the solution during the SCF cycles would normally pass through or even converge to a conducting state. This option provides an alternative damping mechanism to Fock/KS matrix mixing (**FMIXING**, page 69). The locking is effective only if ISHIFT is large enough. If locking is used, the Fermi energy and the eigenvalues are depressed by the value of the level shifter. Suggested values :

1. Normal cases require no mixing of Fock/KS matrices in successive cycles to converge: ISHIFT=0 (default).
2. When 20% to 30% mixing of Fock/KS matrices is necessary, an ISHIFT value of between 1 and 3 (giving a level shift of 0.1 to 0.3 hartree) may produce an equivalent or even superior convergence rate.
3. If serious convergence difficulties are encountered, ISHIFT=10 will normally be adequate, corresponding to a level shift of 1 hartree. But it may happen that the system moves towards an excited state, and no convergence is obtained.

See tests 29, 30, 31, 32, 38.

MAXCYCLE

rec	variable	meaning
• ∗	NMAX	maximum number of SCF cycles [50]

The possibility to modify the maximum number of SCF cycles allows: increasing the number of cycles in case of very slow convergence (metals, magnetic systems, DFT);
The keyword **POSTSCF** forces saving wave function data in file fort.9, even if SCF ends before reaching convergence for "too many cycles".

MONSPLIT - Splitting of large monoelectronic integral files

rec	variable	meaning
• ∗	NFILE	number of files to be used [1] (max 10)

Very large basis sets can produce billions monoelectronic integrals to be stored, as the number of monoelectronic integrals scales as the square of basis set size. The multipolar expansion technique based on the atoms reduces the disk space up to a factor 3, compared to the value printed as estimate. The distribution of the integrals over several disk files may be necessary, if available disk space is limited.

MPP - Massive Parallel Execution - Programmers only

No input data required.
Massive Parallel Libraries are linked, and matrices in K space are distributed over the processors. MPP doesn't support:

keyword	block	
Frequency calculation	**FREQCALC**	1
Geometry optimization	**OPTGEOM**	1
Anderson mixing	**ANDERSON**	3
Broyden mixing	**BROYDEN**	3
Symmetry analysis of Bloch Functions	**KSYMMPRT**	3
Bloch Functions Symmetry Adapted	**SYMADAPT**	3
Printing of eigenvalues of overlap matrix in k space	**EIGS**	3

MYBIPOLA - Bipolar expansion approximation control

rec	variable	meaning
• ∗	ILCOUL	maximum multipole order for Coulomb [4]
∗	ITCOUL	overlap threshold for Coulomb [14]
∗	IFCOUL	reducing factor for Coulomb [90]
• ∗	ILEXCH	maximum multipole order for exchange [2]
∗	ITEXCH	overlap threshold for exchange [10]
∗	IFEXCH	reducing factor for exchange [70]

The bipolar approximation is applied in the evaluation of the Coulomb and exchange integrals (page 178). Maximum values for ILCOUL and ILEXCH are 8 and 4, respectively. ITCOUL and ITEXCH can be assigned any intermediate value between the default values (14 and 10) (see page 178) and the values switching off the bipolar expansion (20000 and 20000). Increasing IFCOUL and IFEXCH the threshold is lightly modified in order to increase the number of approximated integrals, and vice versa.

Warning - for developers only

NEIGHBOR/NEIGHPRT

See input block 1, page 38

NOBIPOLA - Bipolar expansion approximation suppression

All the bielectronic integrals, coulomb and exchange, are evaluated exactly. The overlap threshold both for coulomb and exchange integrals is set to 20000.
No input data required. The CPU time in the **integrals** program may increase up to a factor 3.

NOBIPCOU - Bipolar expansion approximation of coulomb integrals suppression

Coulomb bielectronic integrals are evaluated exactly. The overlap threshold for coulomb integrals is set to 20000.
No input data required.

NOBIPEXC - Bipolar expansion approximation of exchange integrals suppression

Exchange bielectronic integrals are evaluated exactly. The overlap threshold for exchange integrals is set to 20000. No input data required.

NOFMWF - Wave function formatted output

CRYSTAL writes the formatted wave function in file fort.98 at the end of SCF by default. This keyword deletes this feature.

NOMONDIR - Monoelectronic integrals on disk

No input data required.
In the SCF step bielectronic integrals are computed at each cycle, while monoelectronic integrals are computed once and read from disk at each cycle.

NOSYMADA

The Symmetry Adapted Functions are not used in the Fock matrix diagonalization. No input data are required. This choice increases the diagonalization CPU time when the system has symmetry operators.

PARAMPRT - - printing of parametrized dimensions

See input block 1, page 39.

POLEORDR - Maximum order of multipolar expansion

	rec	variable	meaning
•	∗	IDIPO	maximum order of pole [4]

Maximum order of shell multipoles in the long-range zone for the electron-electron Coulomb interaction. Maximum value = 6. See Section 8.3, page 176.

POSTSCF

Calculation to be done after scf (gradient, population analysis) are performed even if convergence is not reached. It may be useful when convergence is very slow, and scf ends for "TOO MANY CYCLES" very close to the convergence criteria required.
No input data are required.

PPAN/MULPOPAN - Mulliken Population Analysis

Mulliken population analysis is performed at the end of SCF process.
No input data are required.
Bond populations are analysed for the first n neighbours (n default value 3; see **NEIGHBOR**, page 38, to modify the value).
Computed data:

1. $a_\mu = \sum_\nu \sum_g P^g_{\mu\nu} S^g_{\mu\nu}$ orbital charges
2. $s_l = \sum_{\mu \in l} a_\mu$ shell charges
3. $q_A = \sum_{l \in A} s_l$ atomic charges
4. $b(A^0, B^g) = \sum_{\mu \in A} \sum_{\nu \in B} P^g_{\mu\nu} S^g_{\mu\nu}$ bond populations between the non-equivalent atoms in the unit cell (A^0) and their first NVI neighbours (B^g in cell g). The printed values must be multiplied by 2 when B≠A to compare with standard molecular calculations.

Formatted data are written in file PPAN.DAT (opened in forrtran unit 24).
See Appendix E, page 216.

PRINTOUT - Setting of printing environment

See input block 1, page 40.

RHF [default]

A restricted closed-shell hamiltonian calculation is performed ([74, 22], Chapter 8 of ref. [12]).
Default choice.

ROHF

Obsolete. See **UHF**

SCFDIR

No input data required.
In the SCF step monoelectronic and bielectronic integrals are evaluated at each cycle. No screening of the integrals is performed.

SAVEWF

The wave function is written in file fort.79 every two cycles. The format is the same as in file fort.9, written at the end of SCF.
No input data required.

SETINF - Setting of INF values

See input block 1, page 42

SETPRINT - Setting of printing options

See input block 1, page 42.

SHRINK - Pack-Monkhorst/Gilat shrinking factors

rec	variable	value	meaning
			if the system is periodic insert II
• ∗	IS		Shrinking factor in reciprocal space (Section 8.7, page 180)
	ISP		Shrinking factor for a denser k point net (Gilat net) in the evaluation of the Fermi energy and density matrix.
			if IS = 0 insert II
• ∗	IS1,IS2,IS3		Shrinking factors along B1,B2,B3 (reciprocal lattice vectors); to be used when the unit cell is highly anisotropic
			optional keywords terminated by **END** or **STOP** II

For periodic systems, 1D, 2D, 3D, the mandatory input information is the shrinking factor, IS, to generate a commensurate grid of **k** points in reciprocal space, according to Pack-Monkhorst method. The Hamiltonian matrix computed in direct space, $H_{\mathbf{g}}$, is Fourier transformed for each **k** value, and diagonalized, to obtain eigenvectors and eigenvalues:

$$H_k = \sum_g H_g e^{i\mathbf{gk}}$$

$$H_k A_k = Sk A_k E_k$$

A second shrinking factor, ISP, defines the sampling of k points, "Gilat net" [20, 21], used for the calculation of the density matrix and the determination of Fermi energy in the case of conductors (bands not fully occupied).
In 3D crystals, the sampling points belong to a lattice (called the Pack-Monkhorst net), with basis vectors:

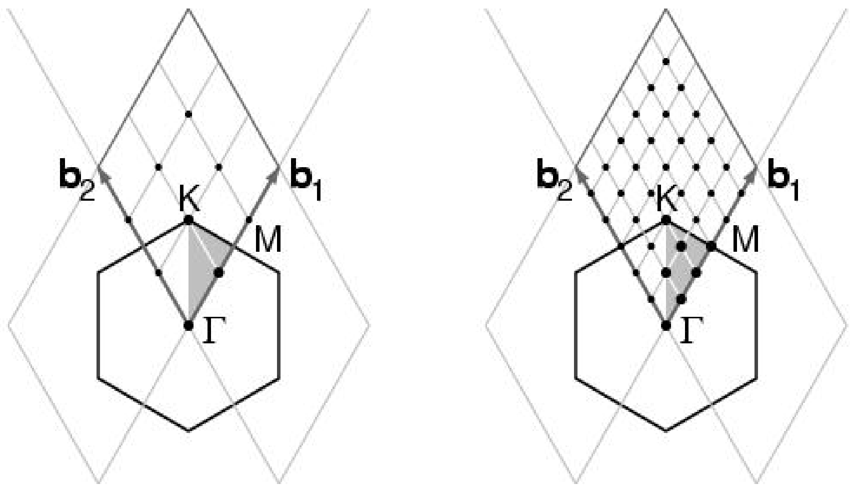
b1/is1, b2/is2, b3/is3 is1=is2=is3=IS, unless otherwise stated

where b1, b2, b3 are the reciprocal lattice vectors, and is1, is2, is3 are integers "shrinking factors".
In 2D crystals, IS3 is set equal to 1; in 1D crystals both IS2 and IS3 are set equal to 1. Only points k_i of the Pack-Monkhorst net belonging to the irreducible part of the Brillouin Zone (IBZ) are considered, with associated a geometrical weight, w_i. The choice of the reciprocal space integration parameters to compute the Fermi energy is a delicate step for metals. See Section 8.7, page 180. Two parameters control the accuracy of reciprocal space integration for Fermi energy calculation and density matrix reconstruction:

IS shrinking factor of reciprocal lattice vectors. The value of IS determines the number of **k** points at which the Fock/KS matrix is diagonalized. Multiples of 2 or 3 should be used, according to the point symmetry of the system (order of principal axes).

In high symmetry systems, it is convenient to assign IS *magic* values such that all low multiplicity (high symmetry) points belong to the Monkhorst lattice. Although this choice does not correspond to maximum efficiency, it gives a safer estimate of the integral.

The **k**-points net is automatically made anisotropic for 1D and 2D systems.



The figure presents the reciprocal lattice cell of 2D graphite (rhombus), the first Brillouin zone (hexagon), the irreducible part of Brillouin zone (in grey), and the coordinates of the $\mathbf{k}_i$ points according to a Pack-Monkhorst sampling, with shrinking factor 3 and 6.

ISP shrinking factor of reciprocal lattice vectors in the Gilat net (see [22], Chapter II.6). ISP is used in the calculation of the Fermi energy and density matrix. Its value can be equal to IS for insulating systems and equal to 2*IS for conducting systems.

The value assigned to ISP is irrelevant for non-conductors. However, a non-conductor may give rise to a conducting structure at the initial stages of the SCF cycle, owing, for instance, to a very unbalanced initial guess of the density matrix. The ISP parameter must therefore be defined in all cases.
Note. The value used in the calculation is ISP=IS*NINT(MAX(ISP,IS)/IS), a multiple integer of IS. For instance:

```
input data  IS    ISP   ISP for wf calculation
            3      4     3
            3      6     6
            3      8     6
```

In the following table the number of sampling points in the IBZ and in BZ is given for a fcc lattice (space group 225, 48 symmetry operators) and hcp lattice (space group 194, 24 symmetry operators). The CRYSTAL code allows 413 k points in the Pack-Monkhorst net, and 2920 in the Gilat net.

IS	points in IBZ fcc	points in IBZ hcp	points BZ
6	16	28	112
8	29	50	260
12	72	133	868
16	145	270	2052
18	195	370	2920
24	413	793	6916
32	897	1734	16388
36	1240	2413	23332
48	2769	5425	55300

1. When an anisotropic net is user defined (IS=0), the ISP input value is taken as ISP1 (shrinking factor of Gilat net along first reciprocal lattice) and ISP2 and ISP3 are set to:
ISP2=(ISP*IS2)/IS1,
ISP3=(ISP*IS3)/IS1.

2. User defined anisotropic net is not compatible with SABF (Symmetry Adapted Bloch Functions). See **NOSYMADA**, page 73.

Some tools for accelerating convergence are given through the keywords **LEVSHIFT** (page 71 and tests 29, 30, 31, 32, 38), **FMIXING** (page 69), **SMEAR** (page 77), **BROYDEN** (page 58) and **ANDERSON** (page 56).
At each SCF cycle the total atomic charges, following a Mulliken population analysis scheme, and the total energy are printed.
The default value of the parameters to control the exit from the SCF cycle ($\Delta E < 10^{-6}$ hartree, maximum number of SCF cycles: 50) may be modified entering the keywords:
TOLDEE (tolerance on change in total energy) page 79;
TOLDEP (tolerance on SQM in density matrix elements) page 79;
MAXCYCLE (maximum number of cycles) page 72.

SMEAR

rec	variable	meaning
• ∗	WIDTH	temperature smearing of Fermi surface

Modifies the occupancy of the eigenvalues (f_j) used in reconstructing the density matrix from the step function, (equation 8.9, page 175) to the Fermi function;

$$f_j = (1 + e^{\frac{(\epsilon_j - \epsilon_F)}{k_b T}})^{-1} \tag{2.11}$$

where ϵ_F is the Fermi energy and $k_b T$ is input as WIDTH in hartree.
The smearing of the Fermi surface surface may be useful when studying metallic systems in which the sharp cut-off in occupancy at ϵ_F can cause unphysical oscillations in the charge density. It may also result in faster convergence of the total energy with respect to k-point sampling.
In density functional theory the use of Fermi surface smearing finds a formal justification in the finite temperature DFT approach of Mermin [75]. In this case the "free energy" of the system may be computed as;

$$\begin{aligned} F &= E(T) - k_b T S(T) \\ &= E - k_b T \sum_i^{N_{states}} f_i \ln f_i + (1 - f_i) \ln(1 - f_i) \end{aligned} \tag{2.12}$$

where S is the electronic entropy. Often we wish to compute properties for the athermal limit (T=0). For the free electron gas the dependencies of the energy and entropy on temperature are;

$$\begin{aligned} E(T) &= E(0) + \alpha T^2 \\ S(T) &= 2\alpha T \end{aligned} \tag{2.13}$$

and so the quantity

$$E0 = \frac{F(T) + E(T)}{2} = E(0) + O(T^3) \tag{2.14}$$

may be used as an estimate of E(0).
Figure 2.3 shows the effect of WIDTH on the convergence of the Li(100) surface energy. Despite the dense k-space sampling (IS=24, ISP=48) the surface energy is rather unstable at low temperature (0.001H). There is a significant improvement in the stability of the solution for higher values of WIDTH (0.02H) but use of E(T) results in a surface energy of 0.643 J/M^2 significantly above that obtained by extrapolating E(T) to the T=0 limit (0.573 J/M^2). The use of E0 at WIDTH=0.02H results in an excellent estimate of the surface energy - 0.576 J/M^2.

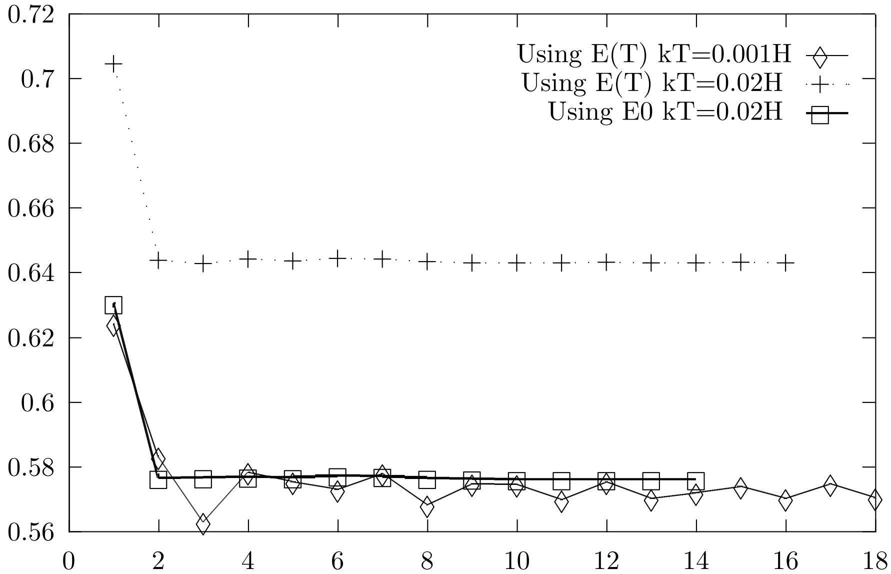


Figure 2.3: The surface energy (J/M^2) of Li(100) for various numbers of layers in a slab model showing the effects of WIDTH (0.02H and 0.001H) and the use of E(T) or E0

SPINEDIT - Editing of the spin density matrix

rec	variable	meaning
• ∗	N	number of atoms for which spin must be reversed
• ∗	LB, L=1,N	*atom labels*

The spin density matrix from a previous run is edited to generate an approximate guess for a new spin configuration. The sign of the elements of the spin density matrix of selected atoms is reversed. The keyword **SPINEDIT** must be combined with **UHF** (input block 3, page 80) and **GUESSP**.
Example: the anti ferromagnetic solution for the spinel $MnCr_2O_4$ can be obtained by calculating the ferro magnetic solution, and using as guess to start the SCF process the density matrix of the ferromagnetic solution with reversed signs on selected atoms.

SPINLOCK - Spin-polarized solutions

rec	variable	meaning
• ∗	NSPIN	n_α-n_β electrons
∗	NCYC	number of cycles the difference is maintained

The difference between the number of α and β electrons at all **k** points can be locked at the input value. The number of α electrons is locked to (N + NSPIN)/2, where N is the total number of electrons in the unit cell; the number of β electrons is locked to (N - NSPIN)/2. NSPIN must be odd when the number of electrons is odd, even when the number of electrons is even.
Example. Bulk NiO. If a anti ferromagnetic solution is required, a double cell containing 2 NiO units must be considered (test 30). The two Ni atoms, related by translational symmetry, are considered inequivalent. The number of electron is 72, each Ni ion is expected to have two unpaired electrons.

INF95	type of solution	corresponding to the spin setting
0	anti ferromagnetic	↑ ↓ ↑ ↓
4	ferromagnetic	↑ ↑ ↑ ↑

See tests 29, 30, 32, 33, 37, 38.

STOP

Execution stops immediately. Subsequent input records are not processed.

SYMADAPT

A computational procedure for generating space-symmetry-adapted Bloch functions, when BF are built from a basis of local functions (AO), is implemented. The method, that applies to any space group and AOs of any quantum number, is based on the diagonalization of Dirac characters [70, 71].
The Symmetry Adapted Functions are used in the Fock matrix diagonalization. No input data are required. This choice reduces the diagonalization CPU time when the system has symmetry operators. Default choice.
Not supported by MPP execution.

TESTPDIM

The program stops after processing of the full input (all four input blocks) and performing symmetry analysis. The size of the Fock/KS and density matrices in direct space is printed. No input data are required.
It may be useful to obtain information on the neighbourhood of the non equivalent atoms (up to 3, default value; redefined through the keyword **NEIGHBOR**, input block 1, page 38).

TESTRUN - Integrals classification and selection

The symmetry analysis is performed, and the monoelectronic and bielectronic integrals classified and selected, according to the the truncation criteria adopted. The size of the Fock/KS and density matrices (direct lattice) and the disk space required to store the bielectronic are printed. The value printed as "disk space for monoelectronic integrals" is an upper limit. The new technique of *atomic* multipolar expansion (not *shell* multipolar expansion as in CRYSTAL95) reduces the required space to about 1/3 of the printed value.
Full input (geometry, basis set, general information, SCF) is processed. No input data after the keyword are required. This type of run is fast, and allows an estimate of the resources to allocate for the traditional SCF wave function calculation.

TOLDEE - SCF convergence threshold on total energy

	rec	variable	meaning
•	*	ITOL	10^{-ITOL} threshold for convergence on total energy

The default value for single point calculation is 6, but 7 in geometry optimization process.

TOLDEP - SCF convergence threshold on density matrix

	rec	variable	meaning
•	*	ITOL	10^{-ITOL} threshold for convergence on ΔP

For developers only.

TOLINTEG - Truncation criteria for bielectronic integrals (Coulomb and HF exchange series)

rec	variable	meaning
• ∗	ITOL1	overlap threshold for Coulomb integrals- page 176 [6]
	ITOL2	penetration threshold for Coulomb integrals-page 177 [6]
	ITOL3	overlap threshold for HF exchange integrals-page 177 [6]
	ITOL4	pseudo-overlap (HF exchange series-page 177) [6]
	ITOL5	pseudo-overlap (HF exchange series-page 177) [12]

The five ITOL parameters control the accuracy of the calculation of the bielectronic Coulomb and exchange series. Selection is performed according to overlap-like criteria: when the overlap between two Atomic Orbitals is smaller than 10^{-ITOL}, the corresponding integral is disregarded or evaluated in a less precise way. Criteria for choosing the five tolerances are discussed in Chapter 8.

TOLPSEUD - Truncation criteria for integrals involving ECPs

rec	variable	meaning
• ∗	ITPSE	overlap threshold for ECP integrals [6]

The program evaluates only those integrals for which the overlap between the charge distribution $\varphi_\mu^0\ \varphi_\nu^g$ (page 174) and the most diffuse Gaussian defining the pseudopotential is larger than a given threshold $T_{ps}=10^{-ITPSE}$ (default value 10^{-6}; it was 5 in CRYSTAL98).

UHF/ROHF - Hamiltonian for Open Shell Systems

For the description of systems containing unpaired electrons (such as molecules with an odd number of electrons, radicals, ferromagnetic and anti ferromagnetic solids) a single determinant is not an appropriate wave-function; in order to get the correct spin eigenfunction of these systems, it is necessary to choose a linear combination of Slater determinants (whereas, in closed shell systems, a single determinant gives always the appropriate spin eigenfunction) ([22, 76], Chapter 6 of ref. [12]).
In the Restricted Open Shell (**ROHF**) [74] Hamiltonian, the same set of molecular (i.e. crystalline) orbitals describes alpha and beta electrons; levels can be doubly occupied (by one alpha and one beta electron, as in the RHF closed shell approach), singly occupied or left vacant. The wave-function is multi-determinantal; in the special case of *half-closed shell* systems, where we can define a set of orbitals occupied by paired electrons and a second set occupied by electrons with parallel spins, the wave-function is formed by a single determinant. This particular mono-determinantal approach can be used in the open-shell part of CRYSTAL. The correct spin state must be defined by the keyword **SPINLOCK**.
Another mono-determinantal approach for the study of open-shell systems is the UHF method [77]. In this theory, the constraint of double occupancy is absent and α electrons are allowed to populate orbitals other than those occupied by the β electrons. Energy levels corresponding to a ROHF and UHF description are plotted in fig. 2.4.
The double occupancy constraint allows the ROHF approach to obtain solutions that are eigenfunctions of the spin operator, $\widehat{S^2}$, whereas UHF solutions are formed by a mixture of spin states. The greater variational freedom allows the UHF method to produce wave-functions that are energetically more stable than the corresponding ROHF ones; another advantage of the UHF method is that it allows solutions with locally negative spin density (i.e. anti ferromagnetic systems), a feature that ROHF solutions can never exhibit.
ROHF solution is not supported by CRYSTAL any more.
Related keywords
SPINLOCK definition of (n_α - n_β electrons)
BETALOCK definition of n_β electrons.

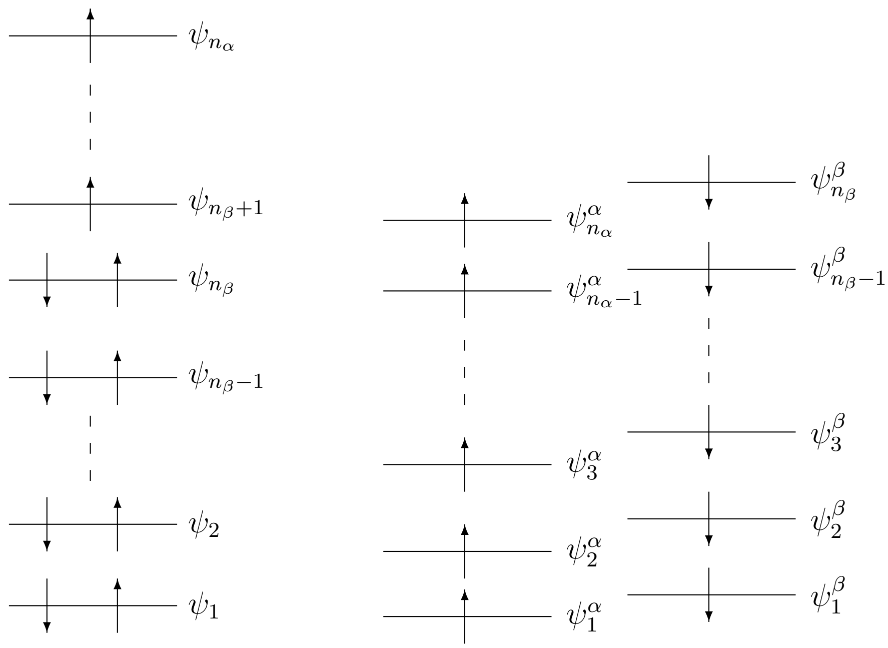


Figure 2.4: Molecular Orbitals diagram for the Restricted Open Shell method (ROHF, left) and for the Unrestricted Open Shell method (UHF, right)

Developers only

FULLTIME - Detailed timing report

A more detailed report of the timing data is generated:

```
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT SHELXG  TELAPSE   19.68 TCPU   18.42
WWWWWWWWWWWWWWW SHELXG      MX    1.07 MN    1.07  MD   1.07
QQQQQQQQQQQQQQQ SHELXG      MX    1.07 MN    0.92  MD   0.98
```

The first line is the standard data. The second line reports the minimum, maximum and mean wall time since the last report. The last line reports the minimum, maximum and mean cpu time since the last report. The minimum, maximum and mean operations are across processors, and so this directive is most useful for parallel job.

DCDIAG - ”divide and conquer” diagonalization

This directive is ONLY for MPP jobs. It instructs the code to use the divide and conquer algorithm for the diagonalization stage. This algorithm can be up to four times quicker than the standard, but it has been found very, very, occasionally to generate incorrect results.

CMPLXFAC - Detailed timing report

This directive is ONLY for MPP jobs. For load balancing reasons the MPP code must know how many times more expensive a calculation on a complex k point is relative to a real one. This allows the user to specify a value for this. The default value is 2.333333.

QVRSGDIM - limitimg size switch for multipole moments gradients

	rec	variable	meaning
•	*	NFILE	limiting size of multipole moment gradients to switch from generation by pairs to generation by shells. Default 90000000.

Chapter 3

Geometry optimization

crystal allows geometry optimization of systems with any periodicity: molecules, polymers, slabs, and crystals. Unconstrained relaxation of the structure and different optimizations with constraints can be carried out. The full symmetry of the system is preserved.
Geometry optimization can be performed in symmetrized fractional coordinates or redundant internal coordinates.

The geometry optimization run is controlled by keywords, that must follow the general keyword **OPTGEOM**, in any order. These keywords can be classified into three groups:

1. *General keywords*:
   (a) Initial Hessian
   (b) Hessian updating
   (c) Step control
   (d) Convergence criteria
   (e) Coordinate system related options
   (f) Optimization procedure control
   (g) Printing options
   (h) Numerical derivatives
2. *Geometry optimization in redundant coordinates*
3. *Constrained geometry optimization*:
   (a) Constant volume optimization
   (b) Fixing lattice parameters
   (c) Linear constraints between atomic coordinates
   (d) Partial optimization of atomic positions
   (e) Fixing internal coordinates

The **OPTGEOM** input block ends with the keyword **END** (or ENDOPT, ENDOPTGEOM, as the first three characters only are processed) and must be specified as the last keyword in geometry input section.
Default values are supplied for all computational parameters.

By default an unconstrained geometry optimization of the atomic positions at fixed cell is performed.

Users can find supplementary information and input examples in the CRYSTAL Tutorials Project web page at the CRYSTAL web site (http://www.crystal.unito.it/tutorials).

Geometry optimization strategy

A Quasi-Newton optimization scheme is implemented. Gradients are evaluated every time the total energy ‘is computed; the second derivative matrix (i.e. Hessian matrix) is built from the gradients. The initial Hessian matrix is obtained from a model Hessian as proposed by Schlegel and updated by using the BFGS algorithm[78, 79, 80, 81]. By default the direction of the step at each cycle is computed by means of a Newton-like scheme, while the length is determined by linear minimization along an extrapolated quadratic polynomial. Optionally, the step considered may be the Newton step (direction and length) controlled by the Trust Radius scheme (see **ALLOWTRUSTR** pag. 86)

HF and DFT (pure and hybrid functionals) analytical gradients are used for insulators and conductors, both for all-electron and ECP calculations.
Note that for conducting systems analytic first derivatives are not fully implemented when the keyword **SMEAR** is used. In that case, *numerical* first derivatives must be computed (see keywords **NUMGRALL**, **NUMGRATO** and **NUMGRCEL**).

For atomic positions, geometry optimization is performed in symmetrized fractional coordinates, in order to exploit the point group symmetry of the lattice. The keyword **PRSYMDIR** (input block 1, page 40) may be used to print the so-called *symmetry allowed directions* adopted in the geometry optimization. If there are no symmetry allowed directions, the program prints a warning message and stops, unless **FULLOPTG** or **CELLONLY** is requested (see below).
To optimize the lattice parameters a set of symmetry preserving cell deformations (see Symmetry Allowed Elastic Distortions, **USESAED**, pag. 46) is defined that are related to changes of isotropic volume and of axial ratios.
By default, the *symmetry allowed deformations* are printed in the output file.
When a full optimization of atom positions and cell parameters is carried out, a normalized combined set of symmetrized directions and deformations is adopted.
Optional choice (keyword **INTREDUN**) is the geometry optimization in redundant internal coordinates. In such a case, atomic displacements and cell deformations are implicitly determined by the internal coordinate system.

Default choices

Type of optimization:

The default geometry optimization type is the relaxation of the nuclear coordinates at fixed lattice parameters in symmetrized fractional coordinates.
Optional choices:

FULLOPTG	full optimization of atomic positions and cell parameters in symmetrized fractional coordinates;
CELLONLY	optimization of cell parameters only;
ITATOCEL	full optimization, iterative procedure: atoms-cell-atoms-cell-......
INTREDUN	full optimization of atomic positions and cell parameters in redundant internal coordinates;

Convergence criteria

A stationary point on the potential energy surface is found when the forces acting on atoms are numerically zero. Geometry optimization is usually completed when the gradients are below a given threshold.
In ***crystal***, the optimization convergence is checked on the root-mean-square (RMS) and the absolute value of the largest component of both the gradients and the estimated displacements. When these four conditions are all satisfied at a time, optimization is considered complete. In some cases (see pag. 96), the optimization process stops with a warning message controlled by the threshold in the energy change between consecutive optimization steps.

Default values are set for all computational parameters, and they may be modified through keywords. Default choices:

	default	keyword
RMS on gradient	0.000300 a.u.	TOLDEG
largest component of gradient	1.5 * 0.000300	1.5 * TOLDEG
RMS on estimated displacements	0.0012 a.u.	TOLDEX
absolute value of largest displacement	1.5 * 0.0012	1.5 * TOLDEX
max number of optimization cycles	100	MAXCYCLE
energy change between optimization steps threshold	10^{-7}a.u.	TOLDEE

Initial Hessian guess

The initial Hessian is generated by means of a classical model as proposed by Schlegel.

H.B. Schlegel, Theoret. Chim. Acta 66 (1984) 333
J.M. Wittbrodt and H.B. Schlegel, J. Mol. Struct. (Theochem) 398-399 (1997) 55

It adopts a simple valence force field. Empirical rules are used to estimate the diagonal force constants for a set of redundant internal coordinates (stretches, bends and torsions). Parameters are available from H to At.

Warning - To define bonds the sum of covalent radii (see page 40) is used. For ionic systems it may be necessary to modify the default values (see **RAYCOV**, page 40).

Hessian updating technique

BFGS Broyden-Fletcher-Goldfarb-Shanno scheme [78, 79, 80, 81].

Optional choices:

1. Schlegel's updating scheme [82], (**OLDCG**), optimization scheme as in CRYSTAL03
2. Powell's updating scheme (**POWELL**)

SCF convergence and guess

The default value for SCF convergence criterion on total energy is set to 10^{-7} (TOLDEE in block3 input to modify it).

After the first step, for the SCF cycle, the density matrix is recovered from the previous geometry optimization step. This can be skipped by inserting the keyword **NOGUESS**. A superposition of atomic densities is then adopted on each step as SCF initial guess.

Output files

The following files (formatted) are written during geometry optimization, and may be saved for further processing.

fort.33 Cartesian coordinates of the atoms in the unit cell and total energy for each geometry optimization step are written in file fort.33 in a simple *xyz* format (see Appendix E, page 215).

This file is suitable to be read by molecular graphics programs (e.g. Molden...) to display the animation of the geometry optimization run.

fort.34 If optimization is successful, the last geometry in written in file fort.34 (format described in Appendix E, page 217).

The file can be read to define the basic geometry input. See **EXTERNAL**, page 11

optaxxx At each xxx optimization step, the geometry is written in file optaxxx (optimization of atoms coordinates only), or optcxxx (optimization of cell parameters or full optimization) in the format of "fort.34" file (see Appendix E, page E). The file must be renamed "fort.34" if used to enter geometry (keyword **EXTERNAL**).

The "history" of the optimization allows restarting from a given step with different parameters, when the procedure did not converge.

OPTINFO.DAT contains information to restart optimization. (see keyword **RESTART** in **OPTGEOM** input block).

OPTHESS.DAT The hessian matrix is written, and can be read to define the initial guess for the Hessian (keyword **HESSINP**)

SCFOUT.LOG SCF and optimization process printout is routed to file SCFOUT.LOG after the first cycle.

General keywords

Initial Hessian

By default an estimated model Hessian is adopted. The Hessian matrix is stored in file OPTHESS.DAT at each optimization step. This may be useful to restart the optimization from a previous run performed at a lower level of theory (e.g. a smaller basis set). An initial Hessian can also be obtained as numerical first-derivative, but this process can be very expensive.
The most general way to select the initial Hessian matrix is through the keyword **HESGUESS**

HESGUESS	defines the initial guess for the Hessian
• * ICODE	Initial guess code:
-1	identity matrix (HESSIDEN)
0	numerical estimate (two-points formula) (HESSNUM)
1	estimated Hessian - model 1 (HESSMOD1)
2	estimated Hessian - model 2 (HESSMOD2 - default)
3	external guess (read from file OPTHESS.DAT: HESSINP)

To help users, guess-related keywords have been adopted:

HESSFREQ	initial guess for the hessian - input from file FREQINFO.DAT obtained from frequencies calculations (Under development)

HESSIDEN	initial guess: identity matrix

HESSINP	external guess (read from file OPTHESS.DAT)

HESSMOD1	initial guess: Lindh's model Hessian [83]

A model Hessian based on a simple 15-parameter function of the nuclear positions as proposed by Lindh et al. is used as initial Hessian. Parameters are available for the first three rows of the periodic table.

R. Lindh, A. Bernhardsson, G. Karlstrom and P.-A. Malmqvist, Chem. Phys. Lett. 241 (1996) 423

HESSMOD2	initial guess: Schlegel's model Hessian [84, 85] [default]

The initial Hessian is generated by means of a classical model as proposed by Schlegel.

H.B. Schlegel, Theoret. Chim. Acta 66 (1984) 333

J.M. Wittbrodt and H.B. Schlegel, J. Mol. Struct. (Theochem) 398-399 (1997) 55

It adopts a simple valence force field. Empirical rules are used to estimate the diagonal force constants for a set of redundant internal coordinates (stretches, bends and torsions). Parameters are available from H to At.

Warning - To define bonds the sum of covalent radii (see page 40) is used. For ionic systems it may be necessary to modify the default values (see **RAYCOV**, page 40).

HESSNUM	initial guess: numerical estimate

Hessian updating

Different Hessian updating schemes are available for minimization:

BFGS	Hessian update - Broyden-Fletcher-Goldfarb-Shanno scheme [78, 79, 80, 81] - [default]

OLDCG	Hessian updating - old Schlegel updating scheme[82] (CRYSTAL03)

POWELL	Hessian update - symmetric Powell scheme [86]

Convergence criteria

These options are available to modify the default values:

TOLDEE	threshold on the energy change between optimization steps
• ∗ IG	$\|\Delta E\| < 10^{-IG}$ (default: 7)

The value of IG must be larger or equal to the threshold adopted for the SCF convergence. The value is checked when input block 3, defining the SCF convergence criteria, is processed.

TOLDEG	convergence criterion on the RMS of the gradient
• ∗ TG	max RMS of the gradient (default: 0.0003)

TOLDEX	convergence criterion on the RMS of the displacement
• ∗ TX	max RMS of the displacement (default: 0.0012)

Step control

To avoid the predicted step size being too large, two options are available:

Simple scaling
A simple scaling of the displacement vector is the default option. Each component is scaled by a factor that makes the largest component of the displacement vector equal to 0.5 a.u.

Trust Radius
A more sophisticated and accurate technique to control the step size is the trust radius region scheme. The trust radius limits the step length of the displacement at each cycle, according to the quadratic form of the surface in the actual region. The default maximum value for minimization is 0.5.
To make this option active, the keyword **ALLOWTRUSTR** must be specified along with **BFGS**.

Related keywords are discussed below:

ALLOWTRUSTR	activate the trust radius technique to control the step size

MAXTRADIUS

• ∗ TRMAX	maximum value allowed for the trust radius - default [∞]

This is useful in minimizations along flat potential surfaces in order to avoid too large displacements from one point to the next one. Default value: minimization: ∞

NOTRUSTR	not using trust radius to limit displacement [default]

TRUSTRADIUS

• ∗ TRAD	sets the initial value for trust radius - default [0.5]

Warning - When the Trust Radius technique is active, the value of the trust radius could become too small and the geometry optimization process stops with an error message: "TRUST RADIUS TOO SMALL". In this case, we suggest to restart the optimization from the last geometry.

Coordinate system related options

Geometry optimization can be performed in fractional (default) or redundant internal coordinates (see **INTREDUN**). Default fractional coordinates are defined as symmetry allowed directions (atomic positions) and deformations (cell). The latter are related to changes of isotropic volume and of axial ratios.

Some options related to the choice of the coordinate systems are also available:

CRYDEF	crystallographic-like symmetrized cell deformations, corresponding to symmetrized strains of the unit-cell edges (consistent with symmetry). This set of deformations is useful for fixing lattice parameters in constrained optimizations in combination with the keyword FIXDEF
FRACTION	optimization in fractionary coordinates
FRACTIOO	optimization in normalized fractionary coordinates [default when FULLOPTG is requested]
FRACTCOOR	third type of symmetrized fractional coordinates (non-orthogonal; the origin on polar axes must be explicitly fixed by the FIXCOOR option [to be used with constrains])
RENOSAED	renormalize symmetry allowed deformations [default when FULLOPTG is requested]

Optimization procedure control

EXPDE

• ∗ DE	expected energy change used to estimate the initial step [default 10^{-3} Ha, if model 1 initial hessian; 10^{-4} Ha, otherwise]

FINALRUN	action after geometry optimization - integrals classification is based on the last geometry. See page 96
• * ICODE	Action type code:
0	the program stops (default)
1	single-point energy calculation
2	single-point energy and gradient calculation
3	single-point energy and gradient calculation - if convergence criteria on gradients are not satisfied, optimization restarts

FIXDELTE

• * IE	10^{-ie} hartree: threshold on the total energy change for redefining the geometry to which integral classification is referred - see **FIXINDEX**, page 67 - [default -1000, no reclassification]

FIXDELTX

• * DX	RMS (bohr) of the displacement for redefining the geometry to which integral classification is referred - [default: -1, no reclassification]

FIXDEIND	the reference geometry for integrals classification does not change during optimization [default choice]

CELLONLY	only cell parameters are optimized. Default: the cell volume may change (see CVOLOPT to optimize at constant volume)

FITDEGR

• * N	degree of polynomial fitting function for linear search:
	2 parabolic fit [default]
	3 cubic polynomial function
	4 constrained quartic fitting

FULLOPTG	full optimization, atom coordinates and cell parameters. The cell volume may change (see CVOLOPT to optimize at constant volume)

HESEVLIM	limits for the allowed region of hessian eigenvalues (hartree)
• * VMIN	lower limit [default 0.001]
VMAX	upper limit [default 1000.]

ITATOCEL	iterate atom cell optimization: atoms-cell-atoms-cell- . . . .

ITACCONV

• * DE	energy difference threshold for ITATOCEL [default 0.1 * TOLDEE between 2 optimization cycles]

MAXITACE

• * MAXI	max number of iteration cycles in atom/cell iterative optimization [default 100]

MAXCYCLE

• * MAX	maximum number of optimization steps [default 100]

NOGUESS	SCF guess at each geometry point: superposition of atomic densities at each SCF calculation

NRSTEPS

• ∗ DE	number of stored steps to be used in the OLDCG Hessian updating scheme [default: number of degrees of freedom]

RESTART restart geometry optimization from a previous run. Not active with **INTREDUN**. See page 96

SORT sorting of the previous optimization steps information when the OLDCG scheme is active [default:nosort]

Printing options

ONELOG This causes all output to be sent to the standard log file, instead of to SCFOUT.LOG

NOXYZ printing of cartesian coordinates at the end of optimization removed

NOSYMMOPS printing of symmetry operators at the end of optimization removed

PRINTFORCES printing atomic gradients

PRINTHESS printing Hessian information

PRINTOPT prints information on optimization process

PRINT verbose printing

Numerical first derivatives

The nuclear coordinate gradients of the energy can also be computed numerically. A three-point numerical derivative formula is adopted. A finite positive (and then negative) displacement is applied to the desired coordinate and a full SCF calculation is performed. The gradient is then computed as

$$g_i = \frac{E_{\Delta x_i} - E_{-\Delta x_i}}{2\ \Delta x_i}$$

where Δx_i is the finite displacement along the i-coordinate.
Such a computation is very expensive compared to analytical gradients, since the cost is $2 \cdot N \cdot t$, where N is the number of coordinates to be optimized and t the cost of the SCF calculation. Numerical first-derivatives should be avoided whenever possible, but sometimes they are the only way to obtain gradients (i.e. for metals) and therefore to optimize the atoms coordinates.

NUMGRALL geometry opt. for numerical atomic and cell gradient

NUMGRATO geometry optimization - numerical atomic gradients

NUMGRCEL geometry optimization - numerical cell gradients

One choice only, **NUMGRCEL**, **NUMGRATO**, **NUMGRALL**, is allowed.

STEPSIZE step for numerical gradient [default 0.001 au] (developers only)

• ∗ I	integer - step = 10^I (default 7: step=10^7)

Optimization in redundant internal coordinates

INTREDUN	geometry optimization in internal redundant coordinates

An optimization in redundant internal coordinates is performed when the keyword **INTREDUN** is specified.
A symmetrized set of internal coordinates (i.e. bonds, angles and torsions) is defined that contains more coordinates than the requisite internal degrees of freedom.
Redundant internal coordinates are generated according to a hierarchical scheme: bond lengths are firstly identified by using covalent radii. Then, angles are determined on the basis of the irreducible set of distances and, finally, dihedral angles are defined. Note that to define bonds the sum of covalent radii (see page 40) is used. For ionic systems it may be necessary to reduce the default values (see **RAYCOV**, page 40). In case of systems constituted by unconnected pieces (*ie* some molecular crystals or adsorption complexes), pieces are linked to each other by *pseudo* "bond lengths" between the closest pair of atoms belonging to each piece.
There has been substantial controversy in recent years concerning the optimal coordinate system for optimizations.
For molecular systems, it is now well-established that redundant internal coordinates require fewer optimization steps than Cartesian coordinates. However, this is not definitely demonstrated for periodic systems. Nevertheless, the use of internal coordinates can be very useful in several respects: for a chemical intuitive view (e.g. internal coordinates can easily be added), for constrained geometry optimization (see below) and for searching transition states (under development).

By default, optimization of internal redundant coordinates involves both, atomic positions and cell parameters To avoid optimizing cell parameters the keyword **FIXCELL** pag. 91 must be specified.

Before running a geometry optimization in redundant internal coordinates, the set of coordinates generated automatically by CRYSTAL should be checked for consistency. This can be done by specifying the keyword **TESTREDU**.

Optional keywords related to the geometry optimization in redundant internal coordinates are listed below.

Managing with almost linear angles

Linear or almost linear angles (i.e. close to 180^0) can lead to numerical instabilities in the computation of the dihedrals. To avoid this problem a common practice is to split the angle in two ones. The double angles are defined by the angles obtained by projection of the vectors onto two suitable perpendicular planes, in order to avoid the indetermination around 180^0. The threshold value beyond which the almost linear angle is split, is controlled by the keyword **ANGTODOUBLE**.

ANGTODOUBLE	minimum value (degrees) beyond which a double angle is defined
• ∗ AL	value of the angle (degrees) - default [165^0]

The default value is set to 165^0. This means that all angles larger than 165^0 are automatically split into two.
This option can be needed, for instance, when optimizing zeolitic structures where siloxane bridges could change a lot during the geometry minimization. In that case, it is better to reduce the default value to 150^0.

A list of angle to be converted into two can also be explicitly given by specifying

DBANGLIST	list of angles chosen to be converted in double angles - advanced option	
• * MU	number of angles to convert in double	
• * IN(I), I=1,MU)	list of the angles	

The labels used for the angles are those provided by a previous automatic generation of internal coordinates computed in a test run (**TESTREDU** keyword).

Adding internal coordinates - bonds and angles

When some relevant internal coordinates are missing (e.g. intermolecular bonds) they can be added by means of two keywords: **DEFLNGS** and **DEFANGLS**.

DEFLNGS	definition of bond lengths	
• * NL	number of bonds to be added	
	insert NL sets of 5 data to define the bond $\overline{AB}$	II
LA	*label* of the atom A (it must be in the reference cell)	
LB	*label* of the atom B	
I1, I2, I3	indices of the cell where the atom B is located	

DEFANGLS	definition of bond angles	
• * NL	number of angles to be defined	
	insert NL sets of 9 data to define the angle $\widehat{ABC}$	II
LA	*label* of the atom A (it must be in the reference cell)	
LB	*label* of the atom B	
I1, I2, I3	indices of the cell where the atom B is located	
LC	*label* of the atom C	
I1, I2, I3	indices of the cell where the atom C is located	

Other optional keywords

FIXCELL	keep cell parameters fixed in internal coordinates optimization (**INTREDUN**)

STEPBMAT	step used for numerical bmat calculation (developers only)
• * I	integer - step = 10^I (default 7: step=10^7)

TESTREDU	request test run for checking automatic definition of internal coordinates .

TOLREDU	tolerance used to eliminate redundancies (developers only)
• * I	tolerance 10^{-I} (default: 7, 10^{-7})

Optimization with constraints

Along with an unconstrained relaxation of the crystalline structure, options are available to perform different optimizations with constraints.
In particular:

A - Constant volume optimization

B - Fixing lattice parameters

C - Linear constraints between atomic coordinates

D - Partial optimization of atomic positions

E - Fixing internal coordinates

All constraining strategies are compatible with any choice of coordinate system adopted for the optimization process to perform the optimization process. On the other hand, option E is only operative together with the choice of a redundant internal coordinate system (**INTREDUN** pag. 90) to perform the optimization.
The examples at the CRYSTAL Tutorial Project web page illustrate the use of the available keywords for constrained geometry optimizations.

A - Constant volume optimization

CVOLOPT	constant volume optimization.

Only active with **CELLONLY** (cell parameters only optimization) or **FULLOPTG** (atom coordinates and cell parameters optimization).
The volume is kept fixed at the value corresponding to the input unit cell; all cell angles and ratios between cell edges unconstrained by the point-group symmetry are optimized.
Examples: in the tetragonal symmetry, only the c/a ratio, and in the monoclinic symmetry the a/b and b/c ratios and the beta angle, respectively, are optimized.

This option is useful for computing point-by-point E vs V curves by relaxing the crystalline structure at different values of the cell volume. In this case, the keyword **FIXINDEX** must be used to obtain a smooth curve (the reference geometry must correspond either to the smallest volume to be explored, or to the equilibrium structure obtained from a prior optimization run (FULLOPTG).

Warning: if large changes of the individual unit-cell parameters occur in the optimization process, the linear strain approximation may not be strictly obeyed and very small volume variations (of the order of 0.01%) may ensue.

B - Fixing lattice deformations

Linear constraints between unit cell deformations can be set up during optimization by means of the keyword **FIXDEF**:

FIXDEF		optimization with constrained symmetrized cell deformation
• *	NFIXC	number of constraints relating pairs of cell deformations
		insert NFIXC records II
• *	LA,LB	integer sequence number of the two constrained symmetrized cell deformations.
	CA,CB	real coefficients multiplying the two cell deformations in the linear combination constraint. If LA=0, the cell deformation denoted by the second integer (LB) is kept fixed during the optimization (the coefficients in this case can take any value).

FIXDEF can also be combined with the keyword **CRYDEF**, that sets crystallographic-like cell deformations (i.e. $a, b, c, \alpha, \beta, \gamma$) to fix lattice parameters. Integer sequence number given as input refer to the minimal set of lattice parameters:

		1	2	3	4	5	6
cubic		a					
hexagonal		$a,$	c				
rhombohedral	hexagonal cell	$a,$	c				
	rhombohedral cell	$a,$	α				

tetragonal	a,	c				
orthorhombic	a,	b,	c			
monoclinic	a,	b,	c,	β		
	a,	b,	c,	γ		
	a,	b,	c,	α		
triclinic	a,	b,	c,	α,	β,	γ

Note that the labels of the symmetry allowed deformations must correspond to the ones printed in the output file.

As an example, a constrained optimization of the crystalline structure of α-quartz (hexagonal) with the c unit cell edge kept fixed follows

```
QUARTZ ALFA  STO-3G
CRYSTAL
0 0 2
154
0 0 16
4.916  5.4054
2
 14 0.4697      0.          0.
  8 0.4135      0.2669      0.1191
OPTGEOM
FULLOPTG
CRYDEF
FIXDEF
1
0 2 0.0 0.0   : the second lattice parameter, c, is kept fixed
ENDOPT
END
```

C - Linear constraints between atomic coordinates

Linear constraints between atomic coordinates can be set up during optimization by using the keyword **FIXCOOR**.

FIXCOOR	optimization with constrained symmetrized coordinates
• ∗ NFIX	number of constraints relating pairs of coordinates
insert NFIX records	II
• ∗ LA,LB	integer sequence number of the two constrained symmetrized coordinates (sequence numbers are read from the output of PRSYMDIR)
CA,CB	real coefficients multiplying the two coordinates in the linear combination constraint. If LA=0, the coordinate denoted by the second integer (LB) is kept fixed during the optimization (the coefficients in this case can take any value).

Note that the labels of the symmetry allowed directions must correspond to the one printed in the output file (**PRSYMDIR** keyword for coordinates).

In the following example on α-quartz, two constraints are set up on coordinates

```
QUARTZ ALFA    -    Linear constraints between atomic coordinates
CRYSTAL
0 0 2
154
0 0 16
4.916  5.4054
```

```
2
 14 0.4697      0.          0.
  8 0.4135      0.2669      0.1191
OPTGEOM
FULLOPTG
FRACTCOOR
FIXCOOR
2
2 3 1.0 1.0
0 4 0.0 0.0
ENDOPT
END
```

1. The x and y fractional coordinates of Oxygen are forced to change by the same amount, so that their difference remains constant.
2. The z coordinate of Oxygen is kept fixed.

In general, any of the structural parameters can be kept fixed in the optimization process by the combined use of **FIXCOOR** and **FIXDEF** keywords.

D - Partial optimization of atomic positions

FRAGMENT	Partial geometry optimization (default: global optimization)
• ∗ NL	number of atoms "free"
• ∗ LB(L),L=1,NL	*label* of the atoms to move

Optimization is limited to an atomic fragment. Symmetrized cartesian coordinates are generated according to the list of atoms allowed to move. Note that no advantage is taken in the gradient calculation to reduce the number of atoms, i.e. gradients are calculated on the whole system. The symmetrized forces are then computed by using the new set of symmetrized coordinates. See example in section 6.4, page 159.

E - Fixing internal coordinates

Constraints on internal coordinates can be easily imposed during the geometry optimization run.
The following two options allow users to both define and freeze one or more bond lengths or angles:

LNGSFROZEN	explicitly freezes bond lengths	
• ∗ MU	number of bond lengths to freeze	
	insert NL sets of 5 data to define the bond $\overline{AB}$	II
LA	*label* of the atom A (it must be in the reference cell)	
LB	*label* of the atom B	
I1, I2, I3	indices of the cell where the atom B is located	

ANGSFROZEN	definition of bond angles to be frozen	
• ∗ NL	number of angles to be frozen	
	insert NL sets of 9 data to define the angle $\widehat{ABC}$	II
LA	*label* of the atom A (it must be in the reference cell)	
LB	*label* of the atom B	
I1, I2, I3	indices of the cell where the atom B is located	
LC	*label* of the atom C	
I1, I2, I3	indices of the cell where the atom C is located	

According to the list of redundant internal coordinates automatically generated by the code, bond lengths or angles can also be frozen by means of the **FREEZINT** option:

FREEZINT		freeze internal coordinates (active with **INTREDUN** only):
• *	NB	first NB bond length are frozen
	NA	first NA bond angles are frozen
	ND	first ND dihedral angles are frozen (not active)

The list of redundant coordinates can be obtained from a prior run, by inserting the keyword **TESTREDU** (the program stops after the printing).

Note that for a better control over the selected frozen internal coordinates we suggest to use the keywords **LNGSFROZEN** and **ANGSFROZEN**.
Constraint optimization combining internal coordinates and fractional coordinates can also be performed.
For instance, one can keep fixed a bond angle together with the constraint that the x and y fractional coordinates of a given atom change by the same amount. Such a combination of constraining strategies must be used with caution, as it may lead to undesired behaviour in the optimization process.
The constraining of internal coordinates is performed using numerical techniques (particularly in the back-transformation from redundant internal to cartesian coordinate systems) and the fixed values may be affected by some small changes (in general in the order of 10^{-4} au). The following example corresponds to a rigid tetrahedral geometry optimization of α-quartz:

```
QUARTZ ALFA fixing internal coordinates
CRYSTAL
0 0 2
154
0 0 16
4.916  5.4054
2
 14 0.4697      0.          0.
  8 0.4135      0.2669      0.1191
OPTGEOM
INTREDUN
LGNSFROZEN
2
4  1  0 0 0
5  1 -1 0 0
ANGSFROZEN
4
4  1  0 0 0  7  0 0 0
4  1  0 0 0  5  1 0 0
4  1  0 0 0  8  1 0 0
5  1 -1 0 0  8  0 0 0
ENDOPT
END
```

The two independent Si-O bond lengths and then the four O-Si-O angles of the SiO_4 tetrahedron are frozen in order to relax just the Si-O-Si bridges and the dihedral angles.

Notes on geometry optimization

On the integrals classification during a geometry optimization

Truncation of infinite Coulomb and exchange series, based on the overlap between two atomic functions (see chapter 8.11), depends on the geometry of a crystal. With default setting of

thresholds different selection of integrals are evaluated with different geometries. This introduces small discontinuities in the PES, producing artificial noise in the optimization process. To avoid noise in interpolation of PES, the **FIXINDEX** option is always active during optimization. The adopted selection pattern refers to the starting geometry.
If equilibrium geometry is significantly different from the starting point, reference truncation pattern may be inappropriate and the use of proper truncation becomes mandatory.
Since both total energy and gradients are affected by the integrals classification, a single-point energy calculation ought to be run always with the final structure, and integrals classified according to the new final geometry, to calculate correct total energy and gradients.
If during the final run the convergence test on the forces is not satisfied, optimization has to be restarted, keeping the integrals classification based on the new geometry. The **FINALRUN** option has been implemented to this aim.
The three different options of **FINALRUN** allow the following actions, after classification of integrals:

1. single-point energy calculation (correct total energy),
2. single-point energy and gradient calculation (correct total energy and gradients),
3. single-point energy and gradient computation, followed by a new optimization process, starting from the final geometry of the previous one, (used to classify the integrals), if the convergence test is not satisfied.

If the starting and final geometry are close, the energy and gradient calculated from the final geometry, with integral classification based on the initial geometry, are not very different from the values obtained with correct classification of integrals. In some cases (e.g. optimization of the geometry of a surface, with reconstruction) the two geometries are very different, and a second optimization cycle is almost mandatory (ICODE=3 in FINALRUN input). This is strongly recommended.

Optimization of flat surfaces

Often the flat regions of surfaces behave as non quadratic. This may give rise to erratic optimization paths when using the linear minimization to control the step length. In these cases it is recommendable to use the trust radius strategy set by the keyword **ALLOWTRUSTR**. Under this scheme the step is controlled so as to never go out from the quadratically behaved regions of the surface (the trust regions). Additionally one can set the maximum trust radius to a given value **MAXTRADIUS** [def ∞], in order to avoid too large displacements from one point to the next one.
Additional combined test on gradient and energy are adopted for treating special cases:

1. If the gradient criteria are satisfied (but not the displacement criteria) and the energy difference between two steps is below a given threshold (see **TOLDEE**), the optimization stops with a warning message;
2. If both the gradient and displacements criteria are not satisfied, but the energy does not change (**TOLDEE** parameter) for four subsequent steps, the optimization stops with a warning message.

Restart optimization

Restart of geometry optimization is possible for a job which is abruptly terminated (e.g. number of steps exceeded, available cpu time exceeded,...). The optimization restarts from the last complete step of the previous run. Information on optimization is read from file OPTINFO.DAT, and saved in the same file at each step. The density matrix is read from file fort.20, written at the last successful step of the optimization process.
The same input deck as for the initial geometry optimization must be used when the RESTART keyword is added.

Visualizing the optimization process

The geometry of the points scanned during the optimization process is written in file fort.33 E(saved as inpfilename.xyz by the script *runcry06*); it can be read by MOLDEN 2.1.

The program MOLDRAW (http://www.moldraw.unito.it) reads CRYSTAL output and allows many types of visualization, taking into account the crystal structure.

SCF guess

At each geometry point the default guess for SCF is the density matrix calculated at the end of the previous step. If the solution does not correspond to real convergence, but to an energy stabilization due to the techniques applied to help convergence (LEVSHIFT, FMIXING, BROYDEN..), the hamiltonian eigenvalues can be unphysical, and there is no chance to recover the SCF process. In those cases it may be better to use an atomic guess (keyword **NOGUESS**).

Chapter 4

Frequency calculations at Γ point

FREQCALC - Frequency harmonic calculation

This keyword allows frequency calculations at the Γ point. It must be the last keyword in geometry input block. See:

F. Pascale, C.M. Zicovich-Wilson, F. Lopez, B. Civalleri, R. Orlando, R. Dovesi
The calculation of the vibration frequencies of crystalline compounds and its implementation in the CRYSTAL code., J. Comput. Chem. 25 (2004) 888-897

C.M. Zicovich-Wilson, F. Pascale, C. Roetti, V.R. Saunders, R. Orlando, R. Dovesi
The calculation of the vibration frequencies of alpha-quartz: the effect of hamiltonian and basis set., J. Comput. Chem. 25 (2004) 1873-1881

The second derivatives of the energy are computed numerically by using the analytical first derivatives. Frequencies are obtained by diagonalizing the mass-weighted Hessian in cartesian coordinates.
The geometry of the system **must** correspond to a stationary point on the potential energy surface.
Wave function calculations to compute numerical derivatives are carried out exploiting the residual symmetry of the system after displacement.
The default value for SCF convergence criterion on total energy is set to 10^{-9} (TOLDEE in block3 input to modify it - expert users only).
The default choice for DFT grid, when DFT hamiltonian is used, corresponds to **XLGRID** (page 63).
The point group symmetry of the lattice is used to reduce the number of scf+gradient calculation to be performed. Second derivatives calculations are done on the irreducible atoms only. The full hessian matrix is then generated applying the point group symmetry to the irreducible part.
The mass-weighted hessian matrix is diagonalized to obtain eigenvalues, which are converted in frequencies (cm^{-1}), and eigenvectors, i.e. the normal modes.
The first step to compute frequencies is the calculation of the wave function at the equilibrium geometry. SCF guess for wave function calculation for all subsequent geometries defined to compute numerical second derivatives is the density matrix obtained at equilibrium geometry.
MPP doesn't support frequency calculation.

Default choices

Longitudinal optical (LO) frequencies and IR intensities are not evaluated by default. If the **INTENS** keyword is used, intensities are evaluated.
As regards the computation of the IR intensities, they are obtained by means of the Wannier Function (WnF) approach, in which those functions span the occupied manifold and are explicitly constructed in real space. They are at time obtained from the eigenvectors of the one-electron Hamiltonian (Bloch Functions) by numerical integration in reciprocal space through

the definition of a Pack-Monkhorst net.
This procedure leads not to real WnFs, but to an approximation contained into a cyclic space. In the mapping (unfolding) that permits to convert cyclic to real WnFs, CRYSTAL exploits the classification of the lattice vectors made at the very beginning of the SCF calculation that, obviously, does not involve the infinite space, but just a cluster of a finite number of cells, ordered by increasing length (i.e. it covers a close to spherical region of the real space).
In all cases tested, this classification provides sufficient room to represent the matrices needed in the SCF part within the required accuracy. This is also so in what concerns the (post-SCF) computation of the WnFs, apart from very particular cases in which the primitive cell is oblong and the corresponding unfolded cyclic cluster associated to the Monkhorst-Pack net (also very elongated in one direction) does not fit into the real cluster (always close to spherical shape).
IR intensities, Born charges and LO-TO split are evaluated through the Wannier functions, obtained by localizing the Crystalline Orbitals. Localization is very demanding, in terms of memory allocation and CPU time. **NOINTENS**, default choice, avoids intensity calculation, when not necessary.
In order to obtain the LO modes, the high frequency dielectric tensor must be provided (a 3 X 3 matrix in input after **DIELTENS** keyword). The dielectric tensor elements can be obtained from the literature or computed with CRYSTAL using SUPERCEL/FIELD (page 32).
A set of keywords can be used to modify the localisation process (see *properties* input, keyword **LOCALWF**, page 124) They are entered after the **DIPOMOME** keyword. Modification of default choices is not recommended, it should be restricted to developers only.
The frequency input block must be closed by the keyword **END** (or **ENDFREQ**). All the keywords are optional.

Output files

Files written during frequency calculation, to be saved to restart a calculation.

SCFOUT.LOG The output from the wave function and gradient calculation is printed in standard output for the reference geometry only. The output is then written in file SCFOUT.LOG.

FREQINFO.DAT Formatted. Contains information on the hessian. Updated at each point, it is necessary to restart a frequency calculation.

OPTHESS.DAT Formatted. Contains the hessian - to be read by **HESSINP**.

fort.9 Binary. Wave function computed at the equilibrium geometry. Full symmetry exploited by default. When those data are used to restart, they are read from file fort.20, as SCF guess.

fort.13 Binary. Reducible density matrix at central point. To exploit maximum symmetry in numerical second derivatives calculations.

fortran unit 80 Binary. Localized Wannier functions, computed only if IR intensities are requested. Necessary to restart a frequency calculation with IR intensities. Usually the name of the file is fort.80

Optional keywords

- A **ANALYSIS** Analysis of the vibrational modes

- A **DIELISO** Calculation of the LO/TO shifts by using the dielectric tensor. The isotropic dielectric tensor (dielectric constant) should be calculated previously with options (SUPERCELL/FIELD) and (DIEL) applied for each axis of the system.
- ∗ DIEL dielectric constant

- A **DIELTENS** Calculation of the LO/TO shifts by using the dielectric tensor.
- ∗ TENS(1:9) Dielectric tensor matrix TENS (3x3 elements, input by rows: 9 reals (3D)

- A **FRAGMENT** Frequency calculation on a moiety of the system
- ∗ NL number of atoms active for frequencies
- ∗ LB(L),L=1,NL *label* of the active atoms

Frequency calculation can be limited to an atomic fragment, instead of the whole system. Symmetry is removed. If in a fragment there are atoms symmetry related, they must be explicitly defined. A reduced hessian is computed, according to the list of atoms belonging to the fragment. A chemically sound moiety of the system must be considered to avoid random results.

- A **INTENS** calculation of IR intensities through Wannier functions
- A **ISOTOPES** atomic masses modified
- ∗ NL number of atoms whose atomic mass must be modified

II———————— *insert NL records* ————————II

- ∗ LB,AMASS *label* and new atomic mass (amu) of the atom.

II————————————————————————II

When the isotopic mass of one atom symmetry related to others is modified, the symmetry of the electronic wave function is not modified, as the mass of the atoms is not present in the single particle electronic hamiltonian. For instance, if in a methane molecule (point group T_d) we want to substitute H with D, we can redefine the mass of the 1, 2, 3, 4 Hydrogen atoms; if C is the first atom, the corresponding input are:

```
1 H atom        2 H atoms       3 H atoms       4 H atoms

ISOTOPES        ISOTOPES        ISOTOPES        ISOTOPES
1               2               3               4
2     2.000     2     2.000     2     2.000     2     2.000
                3     2.000     3     2.000     3     2.000
                                4     2.000     4     2.000
                                                5     2.000
```

If a single D is inserted, the symmetry is reduced, (point group C_{3v}), the three-fold degeneracy becomes two-fold. When all the four Hydrogens are substituted, the three-fold degeneracy is restored.
If a frequency calculation was performed with standard atomic masses, new frequencies values with different atomic masses for selected atoms can be computed from the hessian already computed, at low computational cost, by inserting the keyword **RESTART**.

- A **MODES** Printing of eigenvectors [default]

- A **NOANALYSIS** No analysis of the vibrational modes [default]

- A **NOINTENS** No calculation of the IR intensities [default choice].

- A **NOMODES** No printing of eigenvectors

- A **NORMBORN** Normalize Born tensor to fulfill sum rule

- A **NUMDERIV** specifies the technique to compute the numerical first-derivatives h(x)=dg(x)/dx of the gradient g(x)=dE(x)/dx
- ∗ N 1 different quotient formula: h(x)=(g(x+t)-g(x))/t t=0.001 Å
  2 Central-difference formula: h(x)=(g(x+t)-g(x-t))/2t t=0.001 Å

- A **PRESSURE** Pressure range for thermodynamic analysis
- ∗ NP,P1,P2 3 reals, NP is the number of intervals in the pressure range P1 to P2 (GPa) [1,0.101325,0.101325]

- A **PRINT** Extended printing active (hessian and hessian eigenvectors)
- A **RESTART** Restart frequency calculation from a previous run. See page 102.
- A **STEPSIZE** Modify the step size of displacement along each cartesian axis
- ∗ STEP step (Å) for numerical derivatives [0.001]

- A **TEMPERAT** Temperature range for thermodynamic analysis
- ∗ NT,T1,T2 3 reals, NT is the number of interval in the range T1 to T2 temperature (K) [1,298.0,298.0]
- A **TESTFREQ** Frequency test run

- A **USESYMM** Maximum space group symmetry used to compute wave function at each point [default]

To be modified by developers only.
- A **DIPOMOME** Calculation of the dipole moment - see Localisation part
  To be modified by developers only.
- A **END** end of the **DIPOMOME** block.

all keywords are optional II

- A **DMACCURA** (Optional) Change the final dipole moment tolerance
- ∗ NTOL Value of the new tolerance as TOLWDM=0.1^{-NTOL}

- A **RELOCAL** (Optional) Relocalize all points in frequency calculations
- A **BOYSCTRL** see **LOCALWF**, page 125
- A **CAPTURE** see **LOCALWF**, page 127
- A **WANDM** see **LOCALWF**, page 129
- A **FULLBOYS** see **LOCALWF**, page 130
- A **MAXCYCLE** see **LOCALWF**, page 125
- A **CYCTOL** see **LOCALWF**, page 124
- A **ORTHNDIR** see **LOCALWF**, page 129
- A **CLUSPLUS** see **LOCALWF**, page 129
- A **PHASETOL** see **LOCALWF**, page 125
- A **RESTART** see **LOCALWF**, page 125
- A **IGSSBNDS** see **LOCALWF**, page 126
- A **IGSSVCTS** see **LOCALWF**, page 126
- A **IGSSCTRL** see **LOCALWF**, page 126

Restart a calculation

A frequency calculation for a job abruptly terminated (e.g. machine crash, exceeded the available cpu time,....). can be restarted exactly from the last point of the previous run.

The same input deck used for the incomplete calculation, with the keyword **RESTART** in the **FREQCALC** input block is submitted. The following files, written by the previous job, must be present:

FREQINFO.DAT information on the part of the hessian already computed.

fort.20 wave function at the equilibrium geometry, with no symmetry, as guess for SCF process (fort.9 saved at the end of single point calculation).

fortran unit 80 (usually file fort.80) localized Wannier functions (if IR intensities are requested).

IR intensities calculation must be present in the first frequency calculation, it can not be inserted in restart only.
The restart option can be used to modify the algorithm used to compute gradients (switch from different quotient formula to Central-difference formula, keyword **NUMDERIV**). In this case the new points only are calculated. The same input deck as for the initial frequency calculation must be used.
Restart can be used to evaluate frequencies for a system with different isotopes of selected atoms (keyword **RESTART** followed by **ISOTOPES** 100).

SCANMODE

R. Dovesi, J. Torres, L. Valenzano

This keyword allows to explore the energy and geometry along selected normal modes.

rec	variable	meaning
• ∗	NMO	—NMO— number of modes to be scanned.
		> 0 SCF calculation at each point along the path - energy is computed
		< 0 only the geometry along the path is computed (no SCF calculation)
	INI	Initial point for the scan
	IFI	Final point for the scan
	STEP	Step given as a fraction of the maximum classical displacement, that corresponds to the 1.0 value
• ∗	N(I),I=1,NMO	sequence number of the modes selected.

Let $|r_0>$ be the equilibrium configuration; then the following configurations are explored: $|r_i>=|r_0>+i\Delta|u>$, where $|u>$ is the eigenvector of the selected mode, i is a positive or negative integer, running from INI to IFI, and Δ is the step. $IFI-INI+1$ is the number of points that will be considered in the $INI*STEP-IFI*STEP$ interval. If the STEP variable is set to 1.0, the maximum classical displacement is computed. This displacement corresponds to the point where the potential energy in the harmonic approximation is equal to the energy of the fundamental vibrational state as follows:

$$V = E_0^{vib}$$

$$\frac{1}{2}kx^2 = \frac{1}{2}\hbar\omega$$

Where x=$|r_{max}\rangle - |r_0\rangle$ and the force constant k is given by:

$$k = \omega^2\mu$$

The final expression of the maximum classical displacement is therefore:

$$x = \sqrt{\frac{\hbar}{\omega\mu}}$$

This option can be useful in two different situations.
Let us consider ν_i as the frequency of the Q_i normal mode:

$\nu_i > 0$ we want to explore the energy curve along Q_i normal mode and check the deviation of the energy from the harmonic behaviour. See example 1;

$\nu_i < 0$ the system is in a transition state. We want to explore the Q_i normal mode in order to find a total energy minimum; usually Q_i is not total-symmetric, the symmetry of the structure needs to be reduced. CRYSTAL determines automatically the subgroup of the original group to which the symmetry of the mode belongs. See example 2.

At each point, the geometry is written in file "SCANmode_number_frequencyvalue_DISP_$i\Delta$" (see below), in a format suitable to be read by the keyword **EXTERNAL** (geometry input, page 11).
The geometry of the system then has to be re-optimized in this new subgroup using as a starting geometry one of those external files (better the one corresponding to the minimum). Frequencies can then be evaluated in the new minimum and the new set of frequencies should contain only positive values (apart from the three referring to translations).

Example 1 - Methane molecule

First run: optimization of the geometry (full input at page 165).
Second run: calculation of the vibrational frequencies of CH_4 in the optimized geometry. The optimized geometry corresponds to a minimum, as all frequencies are positive.

```
    MODES        EIGV           FREQUENCIES     IRREP  IR RAMAN
             (HARTREE**2)   (CM**-1)     (THZ)
    1-   3   -0.1863E-11     -0.2995   -0.0090  (F2 )   A   A
    4-   6    0.7530E-07     60.2270    1.8056  (F1 )   I   I
    7-   9    0.4821E-04   1523.8308   45.6833  (F2 )   A   A
   10-  11    0.6302E-04   1742.3056   52.2330  (E  )   I   A
   12-  12    0.2099E-03   3179.3763   95.3153  (A  )   I   A
   13-  15    0.2223E-03   3272.4193   98.1047  (F2 )   A   A
```

Third run: Scanning of a selected mode.
To explore the 12th normal mode, corresponding to C-H symmetric strectchimg, the following lines must be inserted before the end of geometry input (RESTART to read from external file vibrational modes, computed in 2nd run):

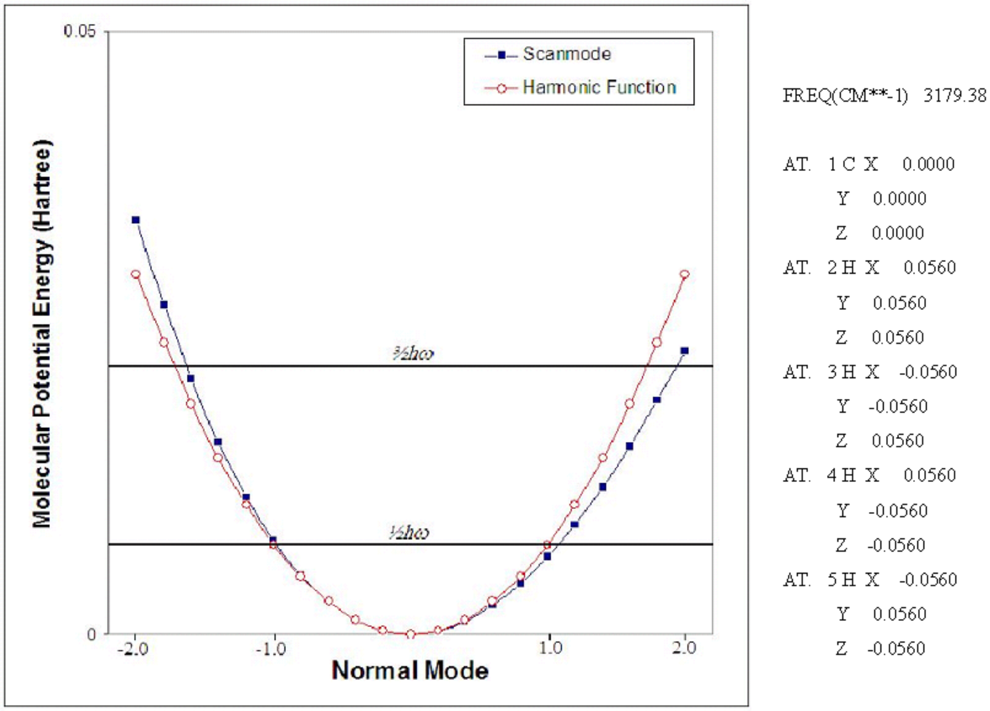
```
FREQCALC
RESTART
SCANMODE
1 -10 10 0.2
12
END
```

The potential energy function as well as its harmonic approximation is computed are represented in the figure. The anharmonicity of C–H stretching is evident.

Example 2 - $PbCO_3$

The spacegroup of this carbonate, as it can be found in the literature [ICSD database], is *Pmcn* (orthorombic lattice).
First run: full optimization of the geometry in *Pmcn* space group (full input at page 165).
Second run: frequency calculation. The output would look as follows:

Figure 4.1: Scanning ot the energy along normal mode 10, ν=3179.3763 cm^{-1}, corresponding to C–H symmetric stretching



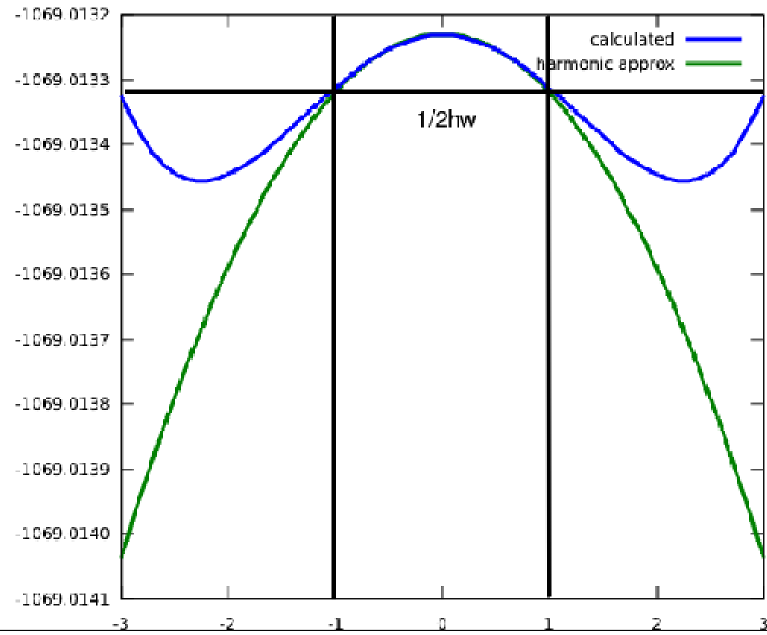

```
    MODES          EIGV          FREQUENCIES      IRREP  IR  RAMAN
                (HARTREE**2)   (CM**-1)     (THZ)
    1-   1    -0.3212E-07     -39.3362    -1.1793   (AU )   I     I
    2-   2    -0.1388E-09      -2.5858    -0.0775   (B3U)   A     I
    3-   3    -0.6924E-10      -1.8262    -0.0547   (B2U)   A     I
    4-   4    -0.2405E-11      -0.3404    -0.0102   (B1U)   A     I
    5-   5     0.4141E-07      44.6637     1.3390   (AG )   I     A
    6-   6     0.4569E-07      46.9137     1.4064   (B3G)   I     A
    7-   7     0.5304E-07      50.5476     1.5154   (B1G)   I     A
 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
   53-  53     0.4245E-04    1429.9950    42.8702   (AU )   I     I
   54-  54     0.4338E-04    1445.5993    43.3380   (B1G)   I     A
   55-  55     0.4340E-04    1445.8649    43.3459   (AG )   I     A
   56-  56     0.4401E-04    1455.9714    43.6489   (B1U)   A     I
   57-  57     0.4408E-04    1457.1539    43.6844   (B3G)   I     A
   58-  58     0.4417E-04    1458.5583    43.7265   (B3U)   A     I
   59-  59     0.4475E-04    1468.2070    44.0157   (B2U)   A     I
   60-  60     0.5007E-04    1553.0286    46.5586   (B2G)   I     A
```

Four negative frequencies are present. Modes 2, 3 and 4 are translations, as results from their small values ($< \ 2\ cm^{-1}$) and from a visual analysis (program MOLDRAW [27]); mode 1, frequency -39.3362 cm^{-1}, corresponds to a maximum along the Q_1 normal coordinate.

Third run: scanning of the first normal mode. The input lines for the frequency calculation block are now the following:

```
FREQCALC
RESTART
SCANMODE
1 -10 10 0.4  scanning of 1 mode, initial point -10, final +10, step 0.4
```

Figure 4.2: Scanning of the energy along normal mode 1, corresponding to a frequency of -39.3362 cm^{-1} (L. Valenzano, unpuplished results)




```
1
END
```

where we are asking to perform the scan of 1 mode (mode 1), computing energy in 21 points in the interval -10/+10 with a step equal to 0.4. Figure 2 shows the energy computed, and the energy in the harmonic approximation.

The optimized geometry of $PbCO_3$ in *Pmcn* space group corresponds to a transition state.

Fourth run: We need to fully re-optimize the geometry of the system with symmetry as a subgroup ($P2_12_12$, space group number 19) of the original space group (*Pmcn*). The geometry, with correct reduced symmetry, is read (EXTERNAL, page 11) from one of the files written during the scan, for instance SCAN1_-39.3361_DISP_-2.400 (scan of mode 1, frequency -39.3361 cm^{-1}, displacement -2.4 the classical amplitude).

Fifth run: After full geometry optimization, we are ready to run a new frequency calculation. The new frequency output looks like (just the first four lines are given):

```
MODES          EIGV            FREQUENCIES      IRREP  IR   INTENS     RAMAN
            (HARTREE**2)    (CM**-1)     (THZ)              (KM/MOL)
1-    1   -0.1504E-09      -2.6917    -0.0807   (B1 )   A (      0.00)    A
2-    2   -0.1414E-09      -2.6097    -0.0782   (B3 )   A (      0.00)    A
3-    3   -0.1690E-11      -0.2853    -0.0086   (B2 )   A (      0.00)    A
4-    4    0.4363E-07      45.8409     1.3743   (A  )   I (      0.00)    A
[. . . . . . . . . . . . . .]
```

Only the three expected negative (translational) modes are present, the fourth negative frequency is not present any more. The $PbCO_3$ structure corresponds now to a minimum in the potential energy surface.

ANHARM - Anharmonic calculation of frequencies of X-H (X-D) bond stretching

rec	variable	meaning
• ∗	LB	*label* of the atom to be displaced (it must have atomic number 1, Hydrogen or Deuterium. The first neighbour (NA) of the LB atom is identified. LB moves along the (NA-LB) direction.
• A	**END**	End of ANHARM input block

This keyword allows the calculation of the anharmonic X-Y stretching. The selected X-Y bond is considered as an independent oscillator. This condition is fulfilled when H or D are involved. It can be used for X-H (or X-D) only.

S. Tosoni, F. Pascale, P. Ugliengo, R. Orlando, V.R. Saunders and R. Dovesi,
"Vibrational spectrum of brucite, Mg(OH)(2): a periodic ab initio quantum mechanical calculation including OH anharmonicity"
Chem. Phys. Lett. **396**, 308-315 (2004)].

Frequencies are calculated as follows:
i) the X-H distance is varied around the equilibrium value, d_0 [default: d_0+ (-0.2, -0.16, -0.06, 0.00, 0.16, 0.24, 0.3 Å)], all other geometrical features being constant (only H moves);
ii) the total potential energy is calculated for each value of the X-H distance [default 7 points];
iii) a polynomial curve of sixth degree is used to best fit the energy points; the root mean square error is well below 10^{-6} hartree;
iv) the corresponding nuclear Schrödinger equation is solved numerically following the method proposed in reference [87]. See P. Ugliengo, "*ANHARM, a program to solve the mono dimensional nuclear Schrödinger equation*", Torino, 1989.

The anharmonicity constant and the harmonic XH stretching frequency are computed from the first vibrational transitions ω_{01} and ω_{02}, as:

$$\omega_e x_e = (2\omega_{01} - \omega_{02})/2$$

$$\omega_e = \omega_{01} + 2\omega_e x_e$$

Stretching of the X-H bond may reduce the symmetry (default). If keyword **KEEPSYMM** is inserted, all equivalent X-H bonds will be stretched, to maintain the symmetry. For example, in CH_4 (point group T_d), KEEPSYMM forces the four CH bonds to stretch in phase; otherwise only the selected C-H bond is stretched, and the symmetry reduced (point group C_{3v}).

Optional keywords of ANHARM input block

ISOTOPES		atomic mass of selected atoms modified
• ∗	NL	number of selected atoms
II		*insert NL records* II
• ∗	LB,AMASS	*label* and new atomic mass (amu) of the atom.
II		II

KEEPSYMM all atoms symmetry equivalent to the selected one are displaced

NOGUESS scf guess at each geometry point: superposition of atomic densities at each scf calculation

POINTS26 26 points: d_{X-H} range: d_0 -0.2 ÷ d_0 +0.3 with a step of 0.02 Å.

PRINT extended printing

PRINTALL printing for programmers

TESTANHA Preliminary test to check if the neighbour(s) of the selected atom is correctly identified and the X-Y direction properly set. No energy calculations is performed.

It has been verified that calculations with 7 points provides very similar results to the ones obtained with 26 points. In the following table, results for POINTS=7 and 26 are reported for three systems. All values are in cm^{-1}.

system		NPOINTS 26	NPOINTS 7
	W_{01}	4358.6	4359.0
HF (molecule)	W_{02}	8607.3	8608.1
	W_e	4468.6	4468.8
	W_eX_e	55.0	54.9
	W_{01}	3325.3	3325.8
$Be(OH)_2$ (bulk)	W_{02}	6406.3	6407.4
	W_e	3569.5	3569.9
	W_eX_e	122.1	122.1
	W_{01}	3637.2	3637.5
$Ca(OH)_2$ (bulk)	W_{02}	7111.4	7111.9
	W_e	3800.3	3800.7
	W_eX_e	81.5	81.6

Chapter 5

Properties

One-electron properties and wave function analysis can be computed from the SCF wave function by running **properties**. At the end of the SCF process, data on the crystalline system and its wave function are stored as unformatted sequential data in file fort.9, and as formatted data in file fort.98 . The wave function data can be transferred formatted from one platform to another (see keyword **RDFMWF**, page 122).
The data in file fort.9 (or fort.98) are read when running **properties**, and cannot be modified. The data include:

1. Crystal structure, geometry and symmetry operators.
2. Basis set.
3. Reciprocal lattice k-points sampling information.
4. Irreducible Fock/KS matrix in direct space (Unrestricted: F_α, F_β).
5. Irreducible density matrix in direct space (Unrestricted: $P_{\alpha+\beta}$ $P_{\alpha-\beta}$).

The **properties** input deck is terminated by the keyword **END**. See Appendix D, page 209, for information on printing.

5.1 Preliminary calculations

In order to compute the one-electron properties it is necessary to access wave function data as binary data set: if binary data are not available in file fort.9, the keyword **RDFMWF**, entered as 1st record, will read formatted data from file fort.98 and write them unformatted in file fort.9.
Full information on the system is generated: :

a.	symmetry analysis information	stored in COMMON areas and modules
b.	reducible Fock/KS matrix	stored on Fortran unit 11
c.	reducible density matrix	
c.1	all electron	stored on Fortran unit 13 (1st record)
c.2	core electron	*stored on Fortran unit 13 (2nd record)*
c.3	valence electron	*stored on Fortran unit 13 (3rd record)*
d.	reducible overlap matrix	stored on Fortran unit 3
e.	Fock/KS eigenvectors	*stored on Fortran unit 10*

1. a, b, c1, d, are automatically computed and stored any time you run the **properties** program.
2. in unrestricted calculations, the total electron density matrix ($\alpha+\beta$) and the spin density matrix ($\alpha - \beta$) are written as a unique record in fortran unit 13.

3. The core and valence electron density matrices (c.2, c.3) are computed *only* by the **NEWK** option when IFE=1. They are stored as sequential data set on Fortran unit 13, after the all electron density matrix. Calculation of Compton profiles and related quantities requires such information.

4. Properties can be calculated using a new density matrix, projected into a selected range of bands (keyword **PBAN**, **PGEOMW**), range of energy (keyword **PDIDE**), or constructed as a superposition of the atomic density matrices relative to the atoms (or ions) of the lattice (keyword **PATO**). In the latter case a new basis set can be used.

   When a specific density matrix is calculated [band projected (**PBAN**), energy projected (**PDIDE**), atomic superposition (**PATO**)], all subsequent properties are calculated using that matrix.

   The option **PSCF** restores the SCF density matrix.

5.2 Properties keywords

RDFMWF wave function data conversion formatted-binary (fort.98 → fort.9)

Preliminary calculations			
NEWK	Eigenvectors calculation	133	I
NOSYMADA	No symmetry Adapted Bloch Functions	73	–
PATO	Density matrix as superposition of atomic (ionic) densities	134	I
PBAN	Band(s) projected density matrix (preliminary **NEWK**)	134	I
PGEOMW	Density matrix from geometrical weights (preliminary **NEWK**)	135	I
PDIDE	Energy range projected density matrix (preliminary **NEWK**)	135	I
PSCF	Restore SCF density matrix	140	–

Properties computed from the density matrix			
ADFT	Atomic density functional correlation energy	111	I
BAND	Band structure	112	I
CLAS	Electrostatic potential maps (point multipoles approximation)	115	I
ECHG	Charge density maps and charge density gradient	119	I
ECH3	Charge density 3D maps	119	I
EDFT	Density functional correlation energy (HF wave function only)	120	I
POLI	Atom and shell multipoles evaluation	135	I
POTM	Electrostatic potential maps	138	I
POTC	Electrostatic properties	137	I
PPAN	Mulliken population analysis	74	
XFAC	X-ray structure factors	141	I

Properties computed from the density matrix (spin-polarized systems)			
ANISOTRO	Hyperfine electron-nuclear spin tensor	111	I
ISOTROPIC	Hyperfine electron-nuclear spin interaction - Fermi contact	123	I
POLSPIN	Atomic spin density multipoles	136	I

Properties computed from eigenvectors (after keyword **NEWK**)			
ANBD	Printing of principal AO component of selected CO	110	I
BWIDTH	Printing of bandwidth	114	I
DOSS	Density of states	118	I
EMDL	Electron momentum distribution - line	121	I
EMDP	Electron momentum distribution - plane maps	121	I
PROF	Compton profiles and related quantities	139	I

New properties

POLARI	Berry phase calculations	143	I
SPOLBP	Spontaneous polarization (Berry phase approach)	145	–
SPOLWF	Spontaneous polarization (localized CO approach)	146	–
PIEZOBP	Piezoelectricity (Berry phase approach) preliminary	142	–
PIEZOWF	Piezoelectricity (localized CO approach) - preliminary	142	–
LOCALWF	Localization of Wannier functions	124	I
DIEL	Optical dielectric constant	116	I
Auxiliary and control keywords			
ANGSTROM	Set input unit of measure to Ångstrom	25	–
BASISSET	Printing of basis set, Fock/KS, overlap and density matrices	113	–
BOHR	Set input unit of measure to bohr	28	–
CHARGED	Non-neutral cell allowed (**PATO**)	47	–
END	Terminate processing of properties input keywords		–
FRACTION	Set input unit of measure to fractional	35	–
MAPNET	Generation of coordinates of grid points on a plane	130	I
NEIGHBOR	Number of neighbours to analyse in **PPAN**	38	I
PRINTOUT	Setting of printing options	40	I
RAYCOV	Modification of atomic covalent radii	40	I
SETINF	Setting of inf array options	42	I
SETPRINT	Setting of printing options	42	I
STOP	Execution stops immediately	43	–
SYMMOPS	Printing of point symmetry operators	46	–
Output of data on external units			
ATOMIR	Coordinates of the irreducible atoms in the cell	112	–
ATOMSYMM	Printing of point symmetry at the atomic positions	28	–
COORPRT	Coordinates of all the atoms in the cell	30	–
CRYAPI_OUT	geometry, BS, direct lattice information	116	–
KNETOUT	Reciprocal lattice information, eigenvalues, eigenvectors	124 obsolete	- –
EXTPRT	Explicit structural/symmetry information	32	–
FMWF	Wave function formatted output. Section 5.2	122	–
INFOGUI	Generation of file with wf information for visualization	123	–
KNETOUT	Reciprocal lattice information + eigenvalues	124	–
MOLDRAW	generation of input file for the program MOLDRAW	36	–

ANBD - Principal AO components of selected eigenvectors

rec	variable	value	meaning
• ∗	NK	n	Number of k points considered.
		0	All the k points are considered.
	NB	n	Number of bands to analyse
		0	All the valence bands + 4 virtual are analysed.
	TOL		Threshold to discriminate the important eigenvector coefficients. The square modulus of each coefficient is compared with TOL.
			if NK > 0 insert II
• ∗	IK(J),J=1,NK		Sequence number of the k points chosen (printed at the top of **NEWK** output)
			if NB > 0 insert II
• ∗	IB(J),J=1,NB		Sequence number of the bands chosen

The largest components of the selected eigenvectors are printed, along with the corresponding AO type and centre.

ADFT/ACOR - *A posteriori* Density Functional atomic correlation energy

The correlation energy of all the atoms not related by symmetry is computed. The charge density is always computed using an Hartree-Fock Hamiltonian (even when the wave function is obtained with a Kohn-Shamm Hamiltonian).
The input block ends with the keyword **END**. Default values are supplied for all the computational parameters.
A new atomic basis set can be entered. It must be defined for *all* the atoms labelled with a different conventional atomic number (not the ones with modified basis set only).

BECKE		Becke weights [default] [65]
	or	
SAVIN		Savin weights [66]

RADIAL		Radial integration information
rec	variable	meaning
• ∗	NR	number of intervals in the radial integration [1]
• ∗	RL(I),I=1,NR	radial integration intervals limits in increasing sequence [4.]
• ∗	IL(I),I=1,NR	number of points in the radial quadrature in the I-th interval [55].

ANGULAR		Angular integration information
rec	variable	meaning
• ∗	NI	number of intervals in the angular integration [default 10]
• ∗	AL(I),I=1,NI	angular intervals limits in increasing sequence. Last limit is set to 9999. [default values 0.4 0.6 0.8 0.9 1.1 2.3 2.4 2.6 2.8]
• ∗	IA(I),I=1,NI	accuracy level in the angular Lebedev integration over the I-th interval [default values 1 2 3 4 6 7 6 4 3 1].
PRINT		printing of intermediate information - no input
PRINTOUT		printing environment (see page 40)
TOLLDENS		
• ∗	ID	DFT density tolerance [default 9]
TOLLGRID		
• ∗	IG	DFT grid weight tolerance [default 18]
NEWBASIS		a new atomic basis set is input
		insert complete basis set input, Section 1.2

ANGSTROM - unit of measure

Unit of measure of coordinates (**ECHG**, **POTM**, **CLAS**) See input block 1, page 25.

ANISOTRO - anisotropic tensor

rec	variable	meaning
• A	keyword	enter one of the following keywords:
• A3	**ALL**	The anisotropic tensor is evaluated for all the atoms in the cell
	or	
• A6	**UNIQUE**	(alias NOTEQUIV) The anisotropic tensor is evaluated for all the non-equivalent atoms in the cell
	or	
• A6	**SELECT**	The anisotropic tensor is evaluated for selected atoms
• ∗	N	number of atoms where to evaluate the tensor
• ∗	IA(I),I=1,N	*label* of the atoms
• A	**PRINT**	extended printing

The anisotropic hyperfine interaction tensor is evaluated. This quantity is given in bohr^{-3} and is transformed into the hyperfine coupling tensor through the relationship [88]

$$\mathbf{T}[\text{mT}] = \frac{1000}{(0.529177 \cdot 10^{-10})^3} \frac{1}{4\pi} \mu_0 \beta_{\text{N}} \text{g}_{\text{N}} \mathbf{T} = 3.4066697 \text{g}_{\text{N}} \mathbf{T}$$

(see **ISOTROPIC** for the units and conversion factors). The elements of the **T** tensor at nucleus A are defined as follows:

$$\text{T}_{ij}^{\text{A}} = \sum_{\mu\nu} \sum_{\text{g}} \text{P}_{\mu\nu\text{g}}^{\text{spin}} \int \varphi_\mu(\boldsymbol{r}) \left(\frac{r_{\text{A}}^2 \delta_{ij} - 3 r_{\text{A}i} r_{\text{A}j}}{r_{\text{A}}^5} \right) \varphi_\nu^{\text{g}}(\boldsymbol{r}) \text{d}\boldsymbol{r}$$

where $\boldsymbol{r}_{\text{A}} = |\boldsymbol{r} - \text{A}|$ and $r_{\text{A}i} = (\boldsymbol{r} - \text{A})_i$ (i^{th} component of the vector).

For extended printing (tensor in original cartesian axes and in principal axis system) insert, before the keyword ANISOTRO:

```
SETPRINT
1
18 1
```

See tests 29, 31, 32, 33.

ATOMIR - coordinates of irreducible atoms

Cartesian and fractional coordinates of the irreducible atoms are printed. No input data required.

ATOMSYMM

See input block 1, page 28

BAND - Band structure

	rec	variable	value	meaning
•	A	TITLE		any string (max 72 characters).
•	*	NLINE	> 0	number of lines in reciprocal space to be explored (max 20)).
		ISS		shrinking factor in terms of which the coordinates of the extremes of the segments are expressed.
		NSUB		total number of k points along the path.
		INZB		first band
		IFNB		last band
		IPLO	0	eigenvalues are not stored on disk.
			$= 1$	formatted output for plotting; see Appendix E, page 212
		LPR66	$\neq 0$	printing of eigenvalues
				add NLINE records
•	*			coordinates of the line extremes in units of $\|\underline{b}_i\|$/ISS
		I1,I2,I3		first point coordinates.
		J1,J2,J3		last point coordinates.

The band structure along a given path n the Brillouin zone is computed.
The data are printed in standard output and (if IPLO = 1) written in file fort.25 (formatted data processed by Crgra2006) and in file BAND.DAT (processed by DLV; see http://www.cse.clrc.ac.uk/cmg/DLV). See Appendix E, page 212).

1. **Warning** : does not run for molecules!! (0D)

2. For a correct interpretation of HF band-structure and DOS's, it must be stressed that the HF eigenvalues are not a good approximation to the optical excitation spectrum of the crystal. However, as discussed in section III.2 of reference [22] and in Chapter 2 of reference [12], the band structures, in conjunction with total and projected DOS's, can be extremely useful in characterizing the system from a chemical point of view.

3. Note on band extremes coordinates: in two-(one-) dimensional cases I3, J3 (I2,I3,J2,J3) are formally input as zero (0). See test 3 and 6.

4. The only purpose of ISS is to express the extremes of the segments in integer units (see tests 8-9). It does not determine the density of k points along the lines, which depends only on NSUB. The number of k points for each line is computed by the program. The step is constant along each line. The step is taken as close as possible to a constant along different lines.

5. If symmetry adapted Bloch functions are used (default option), **BAND** generates symmetry information in k points different from the ones defined by the Monkhosrt net. Eigenvectors computed by NEWK in k points corresponding to the Monkhosrt net are not readable any more. To compute density of states and bands, the sequence must be: BAND - NEWK - DOSS.

See tests 3, 4, 6, 7, 8, 9, 11, 12 and 30.

BASISSET - Printing of basis set and data from SCF

rec	variable	value	meaning
• ∗	NPR		number of printing options.
			if NPR ≠ 0 insert **prtrec** *(see page 42)* II

This option allows printing of the basis set and the computational parameters, and, on request (keyword **PRINTOUT** before **BASISSET**), of the Fock/KS matrix (**FGRED**), the overlap matrix (**OVERLAP**), and the reducible density matrix (**PGRED**), in direct lattice representation.
Warning: the contraction coefficients of the primitive gaussians are different from the ones given in input. See "Normalization coefficients", Appendix F.
Printing options:
59 (Density matrix); 60 (Overlap matrix); 64 (Fock/KS matrix).

BIDIERD - Reciprocal form factors - developers only

This option evaluates the reciprocal form factors (RFF) directly from the direct space density matrix. By using the PROF keyword, followed by the BR keyword, it is possible to obtain the same quantity by Fourier transforming the Compton profiles. As the latter implies numerical integration, the BIDIERD keyword is expected to provide more accurate results.

rec	variable	value	meaning
• ∗	NDIR		number of directions along which the RFF are evaluated
	NPU		number of sampling points along each direction
	STEP		step along each direction
	IMODO	0:	the direction is defined by the Cartesian coordinates (bohr) of a point
		≠ 0:	the direction is defined by the atom label and indices of the cell where the atom is located
	ICASO	1:	the total density matrix is used
		2:	the core density matrix is used
		3:	the valence density matrix is used
	IPLOT	0:	data are not saved for plot
		≠ 0:	data are saved in file fort.25
	LPR	≠ 0:	print the RFF vector for the various directions
			if IMODO=0, insert NDIR records
• ∗	X Y Z		the explored direction is defined by the straight line going from the origin to (X,Y,Z)
			if IMODO≠0, insert NDIR records
• ∗	I XG YG ZG		label of the atom and indices of the cell where the atom is located. The explored direction is defined by the straight line going from the origin to the atom position

Notes:
The explored interval is (NPU−1)×STEP long; X,Y,Z or I,XG,YG,ZG data are just used for defining the direction, **NOT** the length of the explored interval.

BOHR - unit of measure

Unit of measure of coordinates (**ECHG**, **POTM**, **CLAS**) See input block 1, page 28.

BWIDTH - Printing of band width

rec	variable	meaning
• ∗	INZB	first band considered
	0	analysis from first valence band
	IFNB	last band considered
	0	analysis up to first 4 virtual bands

The Fock/KS eigenvalues are ordered in bands following their values. Band crossing is not recognized.

CHARGED - charged reference cell

See input block 2, page 47.
To be used before **PATO**, when new basis set and/or electron configuration of the atoms result in a charged cell.

CLAS - Point charge electrostatic potential maps

rec	variable	value	meaning
• ∗	IDER	0	potential evaluation
		1	calculation of potential and its first derivatives
	IFOR	0	point multipoles have to be evaluated by **POLI** option
		1	point formal charges given as input
			if IFOR ≠ 0 insert II
• ∗	Q(I),I=1,NAF		formal net charge for all the NAF atoms in the unit cell (equivalent and non equivalent, following the sequence printed at the top of the ***properties*** printout)
			insert **MAPNET** input records (page 130)

1. When IDER=0, the electrostatic potential is calculated at the nodes of a 2-dimensional net in a parallelogram-shaped domain defined by the segments AB and BC (see keyword **MAPNET**, page 130). The potential values are written formatted in file fort.25. (see Appendix E, page 212).
2. When IDER ≠ 0, the electrostatic potential gradient is computed at the nodes of the same grid. The x, y and z components are printed on the standard output.
3. The potential is generated by an array of point multipoles up to a maximum order IDIPO defined in the **POLI** option input, or by atomic point charges given in input (IFOR=1; IDIPO = 0 is set in that case).
4. The multipoles *must* be previously computed by running the option **POLI** when IFOR is equal to zero.

COORPRT

See input block 1, page 30.

DENSMAT - First order density matrix $\rho(r, r')$ - developers only

First order density matrix $\rho(r, r')$ along a given path is computed.
The variable r' explores the same interval as r.
For UHF cases two matrices are generated, one corresponding to the total and the other to the spin density matrix.

rec	variable	value	meaning
• ∗	NKN		number of knots in the path (=number of segments+1)
	NPU		number of sampling points along the full path
	IPLOT	0:	data are not saved for plot
		= 1:	data are saved in file fort.25
	IMODO	0:	knot coordinates (x, y, z) in a. u.
		≠ 0:	knots are defined through atom labels
	LPR	≠ 0:	print the $\rho(r, r')$ matrix in integer form (values are multiplied by 10000)
			if IMODO=0, insert NKN records
• ∗	X Y Z		Cartesian coordinates (bohr) of the i-th knot
			if IMODO≠0, insert:
• ∗	DX DY DZ		displacement (bohr) applied to all atoms defining the path
			insert NKN records
• ∗	I XG YG ZG		label of the atom and indices of the cell where the atom is located

- A NPU×NPU square matrix is generated.
- The step between contiguous sampling points belonging to different segments is the same.

- Meaning of the displacement: suppose you want the density matrix corresponding to the π structure of benzene. Define, for example, the path H–C–C–C–H through the atom labels and then displace it along z (if the molecule is in the $x-y$ plane) by an appropriate amount.

CRYAPI_OUT - Geometry, BS, and full wave function information

Geometry, local function Basis Set, overlap, hamiltonian, density matrices n direct lattice are written formatted in file GRED.DAT
Wannier functions (if file fort.80 is present; see keyword **LOCALWF**, page 124) are appended to file GRED.DAT
k points coordinates (Monkhorst sampling net) and eigenvectors (if computed by **NEWK** page 133) in the full Brillouin zone are written formatted in file KRED.DAT.
The scripts ***runcry06/runprop06*** save files GRED.DAT and KRED.DAT (if present) as inpfilename.GRED and inpfilename.KRED
The utility program ***cryapi_inp*** reads and prints the data. The organization of data can be understood from the output of ***cryapi_inp*** and from its source.
See Appendix E, page 219.

DIEL/DIELECT - Optical dielectric constant

rec	variable	meaning
• *A*	**END/ENDDIEL**	end of **DIEL** input block
		optional keywords II
• *A*	**PRINT**	extended output

The electron density must be obtained by applying an electric field (keyword **FIELD**, page 32). The dielectric constant is calculated by using the concept of macroscopic average of the total charge density (see for example Fu *et al.* [25]) and Poisson's equation. The charge density is first averaged with respect to the (infinite) plane orthogonal to the field

$$\overline{\rho}(z) = \frac{1}{A}\int\limits_{A} \rho(z)\, dA \tag{5.1}$$

where $A = |\vec{a} \times \vec{b}|$, and $\vec{a}$ and $\vec{b}$ are the lattice parameters of the supercell orthogonal to the field direction. When a Fourier representation of the charge density is used, the previous equation becomes:

$$\overline{\rho}(z) = \frac{1}{V}\sum_{\ell=-\infty}^{+\infty} F_{00\ell}\, e^{-\imath\frac{2\pi\ell z}{C}} \tag{5.2}$$

$F_{00\ell}$ are structure factors (note that the two first indices are always zero) calculated analytically from the SCF crystalline orbitals depending now on the applied field. The quantity $\overline{\rho}$ is then averaged with respect to the z coordinate

$$\overline{\overline{\rho}}(z) = \frac{1}{\Delta z}\int\limits_{z-\Delta z/2}^{z+\Delta z/2} \overline{\rho}(z')\, dz' \tag{5.3}$$

that is

$$\overline{\overline{\rho}}(z) = \frac{1}{V}\sum_{\ell=-\infty}^{+\infty} F_{00\ell}\, sinc\left(\ell\pi\frac{\Delta z}{C}\right) e^{-\imath\frac{2\pi\ell z}{C}} \tag{5.4}$$

where the *sinc* function is the *cardinal sinus* ($sinc(u) = \frac{\sin(u)}{u}$) and Δz has been chosen equal to c; we can now apply Poisson's equation to $\overline{\overline{\rho}}(z)$:

$$\frac{\partial^2\overline{\overline{V}}(z)}{\partial z^2} = -4\pi\overline{\overline{\rho}}(z) \tag{5.5}$$

or

$$\frac{\partial \overline{\overline{E}}(z)}{\partial z} = 4\pi\overline{\overline{\rho}}(z) \tag{5.6}$$

because

$$\frac{\partial \overline{\overline{V}}(z)}{\partial z} = -\overline{\overline{E}}(z) \tag{5.7}$$

$\overline{\overline{V}}(z)$, $\overline{\overline{F}}(z)$ and $\overline{\overline{\rho}}(z)$ are the mean values of the macroscopic electric potential, electric field and electron density at z position along the electric field direction.
Structure factors can be separated in a real and an imaginary part:

$$F_{00\ell} = F_{00\ell}^{\Re} + \imath F_{00\ell}^{\Im} \tag{5.8}$$

Exploiting the following properties of the structure factors:

$$\begin{aligned} F_{000} &= N_e \quad (\textbf{number of electrons in the supercell}) \\ F_{00\ell}^{\Re} &= F_{00-\ell}^{\Re} \\ F_{00\ell}^{\Im} &= -F_{00-\ell}^{\Im} \end{aligned} \tag{5.9}$$

the real and imaginary parts of $\overline{\overline{\rho}}$ take the following form:

$$\Re\left[\overline{\overline{\rho}}(z)\right] = \frac{N_e}{V} + \frac{2}{V}\sum_{\ell=1}^{+\infty}\left[F_{00\ell}^{\Re}\cos\left(\frac{2\pi\ell z}{C}\right) + F_{00\ell}^{\Im}\sin\left(\frac{2\pi\ell z}{C}\right)\right] sinc\left(\ell\pi\frac{\Delta z}{C}\right) \tag{5.10}$$

$$\Im\left[\overline{\overline{\rho}}(z)\right] = 0 \tag{5.11}$$

As expected, the imaginary part is null. The N_e/V term can be disregarded, as it is exactly canceled by the nuclear charges in the supercell.
According to equation 5.7, the local macroscopic field corresponds to minus the slope of $\overline{\overline{V}}(z)$, it has opposite sign with respect to the imposed outer field, according to the Lenz law, and is obtained from $\overline{\overline{\rho}}(z)$(eq. 5.6):

$$\overline{\overline{E}}(z) = \frac{8\pi}{V}\sum_{\ell=1}^{+\infty}\left[F_{00\ell}^{\Re}\frac{\sin\left(\frac{2\pi\ell z}{C}\right)}{\left(\frac{2\pi\ell}{C}\right)} - F_{00\ell}^{\Im}\frac{\cos\left(\frac{2\pi\ell z}{C}\right)}{\left(\frac{2\pi\ell}{C}\right)}\right] sinc\left(\ell\pi\frac{\Delta z}{C}\right) \tag{5.12}$$

The corresponding macroscopic electric potential can be written as follows:

$$\overline{\overline{V}}(z) = \frac{-8\pi}{V}\sum_{\ell=1}^{+\infty}\left[F_{00\ell}^{\Re}\frac{\cos\left(\frac{2\pi\ell z}{C}\right)}{\left(\frac{2\pi\ell}{C}\right)^2} + F_{00\ell}^{\Im}\frac{\sin\left(\frac{2\pi\ell z}{C}\right)}{\left(\frac{2\pi\ell}{C}\right)^2}\right] sinc\left(\ell\pi\frac{\Delta z}{C}\right) \tag{5.13}$$

Since $-\overline{\overline{E}}$ and E_0 have opposite sign, the ratio $E_0/(E_0+\overline{\overline{E}})$ is larger than one, and characterizes the relative dielectric constant of the material along the direction of the applied field:

$$\epsilon = \frac{E_0}{E_0 + \overline{\overline{E}}} \tag{5.14}$$

The number of structure factors computed for a Fourier representation of the perturbed charge density by default is equal to 300, the structure factors from F_{001} to $F_{00\ 300}$.

The data computed are written in file DIEL.DAT in append mode. See Appendix E, page 213.

DOSS - Density of states

rec	variable	value	meaning
• *	NPRO	0	only total DOS is calculated
		> 0	total DOS and NPRO projected densities are calculated. The maximum number of projections is 15.
	NPT		number of uniformly spaced energy values ($\leq$ LIM019) where DOSs are calculated, from bottom of band INZB to top of band IFNB.
	INZB		first band considered in DOS calculation
	IFNB		last band considered in DOS calculation
	IPLO	0	DOSs are not stored on disk
		1	formatted output on Fortran unit 25 for plotting (Appendix E, page 213).
		2	formatted output on file DOSS.DAT (Fortran unit 24) for plotting (Appendix E, page 214).
	NPOL		number of Legendre polynomials used to expand DOSS (≤ 25)
	NPR		number of printing options to switch on
			if INZB and IFNB < 0 insert II
• *	BMI,BMA		Minimum and maximum energy (hartree) values to span for DOSS. They must be in a band gap
			if NPRO $\neq 0$, insert NPRO records II
• *	N	> 0	DOS projected onto a set of N AOs
		< 0	DOS projected onto the set of all AOs of the N atoms.
	NDM(J),J=1,N		vector NDM identifies the AOs (N>0) or the atoms (N<0) by their sequence number (basis set order)
			if NPR $\neq 0$, insert **prtrec** *(see page 42)* II

Following a Mulliken analysis, the orbital (ρ_μ), atom (ρ_A) and total (ρ_{tot}) density of states can be defined for a closed shell system as follows:

$$\rho_\mu(\epsilon) = 2/V_B \sum_j \sum_\nu \sum_{\underline{g}} \int_{BZ} d\underline{k} S_{\mu\nu}(\underline{k}) a_{\mu j}(\underline{k}) a^*_{\nu j}(\underline{k})\ e^{i\underline{k}\cdot\underline{g}}\ \delta[\epsilon - \epsilon_j(\underline{k})] \tag{5.15}$$

$$\rho_A(\epsilon) = \sum_{\mu \in A} \rho_\mu(\epsilon) \tag{5.16}$$

$$\rho_{tot}(\epsilon) = \sum_A \rho_A(\epsilon) \tag{5.17}$$

where the last sum extends to all the atoms in the unit cell.
Bond population density of states are not computed.

1. **Warning:** do not run for molecules!

2. The **NEWK** option must be executed (to compute Hartree-Fock/KS eigenvectors and eigenvalues) before running **DOSS**. The values of the input parameters IS and ISP of **NEWK** have a consequent effect on the accuracy of the DOSS calculation. Suggested values for IS: from 4 to 12 for 3-D systems, from 6 to 18 for 2-D and 1-D systems (Section 8.7, page 180). ISP must be equal or greater than 2*IS; low values of the ratio ISP/IS can lead to numerical instabilities when high values of NPOL are used.
   If **BAND** is called between **NEWK** and **OSS**, and symmetry adapted Bloch functions are used (default option), the information generated by NEWK is destroyed. To compute density of states and bands, the sequence must be: BAND - NEWK - DOSS.

3. DOSS are calculated according to the Fourier-Legendre technique described in Chapter II.6 of reference 1, and in C. Pisani et al, ([89, 90]). Three computational parameters must be defined: NPOL, IS, ISP. IS and ISP are entered in the **NEWK** option input.

4. NPOL is the number of Legendre polynomials used for the expansion of the DOS. The value of NPOL is related to the values of IS and ISP, first and third input data of **NEWK** option.

Suggested values for NPOL: 10 to 18.

5. **Warning NEWK** with IFE=1 must be run when spin-polarized solutions (**SPINLOCK**, page 78) or level shifter (**LEVSHIFT**, page 71) were requested in SCF, to obtain the correct Fermi energy and eigenvalues spectra.

6. Unit of measure: energy: hartree; DOSS: state/hartree/cell.

Computed data are written in file fort.25 (in Crgra2006 format), and in file DOSS.DAT
Printing options: 105 (density of states and integrated density of states); 107 (symmetrized plane waves).
See tests 3, 4, 5, 6, 7, 8, 9, 11 and 30.

ECH3 - Electronic charge (spin) density on a 3D grid

rec	variable	meaning
• ∗	NP	Number of points along the first direction

if non-3D system
keyword to choose the type of grid on the non-periodic direction(s):

SCALE length scales for non-periodic dimensions	**RANGE** boundary for non-periodic dimensions (au)
if 2D system	
• ∗ ZSCALE	• ∗ ZMIN
	• ∗ ZMAX
if 1D system	
• ∗ YSCALE,ZSCALE	• ∗ YMIN,ZMIN
	• ∗ YMAX,ZMAX
if 0D system	
• ∗ XSCALE,YSCALE,ZSCALE	• ∗ XMIN,YMIN,ZMIN
	• ∗ XMAX,YMAX,ZMAX

The electronic charge or spin density (electron/bohr^3)is computed at a regular 3-dimensional grid of points. The grid is defined by the lattice vectors of the primitive unit cell and user defined extents in non-periodic directions. NP is the number of points along the first lattice vector (or XMAX-XMIN for a molecule). Equally spacing is used along the other vectors. Non-periodic extents may be specified as an explicit range (RANGE) or by scaling the extent defined by the atomic coordinates (SCALE).
Formatted data are written in fortran unit 31 (function value at the grid points) and 32 (general information on the system), in the format required by the visualization program DLV.
See Appendix E, page 218, for description of the format.
PS. The sum of the density values divided by the number of points and multiplied by the cell volume (in bohr, as printed in the output) gives a very rough estimate of the number of electrons.

ECHG - Electronic charge density maps and charge density gradient

rec	variable	value	meaning
• ∗	IDER	n	order of the derivative - < 2
			insert **MAPNET** input records (Section 5.2, page 130)

1. IDER=0
   The electron charge density (and in sequence the spin density, for unrestricted wave functions) is calculated at the nodes of a 2-dimensional net in a parallelogram-shaped domain defined by the segments AB and BC (see keyword **MAPNET**, page 130). The electron density values (electron bohr^{-3}) are written formatted in file fort.25 (see Appendix E, page 212).

2. IDER=1
   electron charge density, x, y, z component of first derivative, and modulus of the derivative, are written. The string in the header is always "MAPN".

3. When the three points define a segment (A≡B or B≡C), function data are written in file RHOLINE.DAT. (see Appendix E, page 212)

4. When IDER $\neq 0$, the charge density gradient is computed at the nodes of the same grid. The x, y and z components are printed on the standard output and written formatted in file fort.25 (see Appendix E, page 212).

5. The electron charge density is computed from the density matrix stored in fortran unit 13. The density matrix computed at the last cycle of **SCF** is the default.

6. Band projected (keyword **PBAN**), energy projected (keyword **PDIDE**) or atomic superposition (keyword **PATO**) density matrices can be used to compute the charge density. The sequence of keywords must be: (**NEWK-PBAN-ECHG**), (**NEWK-PDIDE-ECHG**) or (**PATO-ECHG**).

EDFT/ENECOR - *A posteriori* Density Functional correlation energy

Estimates *a posteriori* the correlation energy via a HF density. It is controlled by keywords. The input block ends with the keyword **END**. All the keywords are optional, as default values for all the integration parameters are supplied by the program, to obtain reasonably accurate integration of the charge density. Please check the integration error printed on the output.

BECKE		Becke weights [default] [65]
	or	
SAVIN		Savin weights [66]

RADIAL		Radial integration information
rec	variable	meaning
• ∗	NR	number of intervals in the radial integration [1]
• ∗	RL(I),I=1,NR	radial integration intervals limits in increasing sequence [4.]
• ∗	IL(I),I=1,NR	number of points in the radial quadrature in the I-th interval [55].

ANGULAR		Angular integration information
rec	variable	meaning
• ∗	NI	number of intervals in the angular integration [default 10]
• ∗	AL(I),I=1,NI	angular intervals limits in increasing sequence. Last limit is set to 9999. [default values 0.4 0.6 0.8 0.9 1.1 2.3 2.4 2.6 2.8]
• ∗	IA(I),I=1,NI	accuracy level in the angular Lebedev integration over the I-th interval [default values 1 2 3 4 6 7 6 4 3 1].
PRINT		printing of intermediate information - no input
PRINTOUT		printing environment (see page 40)
TOLLDENS		
• ∗	ID	DFT density tolerance [default 9]
TOLLGRID		
• ∗	IG	DFT grid weight tolerance [default 18]

EMDL - Electron Momentum Density - line maps

rec	variable	value	meaning
• ∗	N		number of directions (≤ 10)
	PMAX		maximum momentum value (a.u.) for which the EMD is to be calculated
	STEP		interpolation step for the EMD
	IPLO	0	no data stored on disk
		1	formatted output on Fortran unit 25 for plotting (Appendix E, page 214).
		2	formatted output on Fortran unit 24 for plotting (Appendix E, page214).
	LPR113	≠ 0	printing of EMD before interpolation
• ∗	(K(I,J), I=1,3),J=1,N		directions in oblique coordinates
• ∗	NPO		number of orbital projections (≤ 10)
	NPB		number of band projections(≤ 10)
			if NPO ≠ 0 insert NPO sets of records II
• ∗	NO		number of A.O.'s in the I-th projection
• ∗	IQ(I),I=1,NO		sequence number of the A.O.'s in the I-th projection - basis set sequence.
			if NPB ≠ 0 insert NPB sets of records II
• ∗	NB		number of bands in the I-th projection
• ∗	IB(I),I=1,NB		sequence number of the bands in the I-th projection

Warning EMDL does not work for UHF wave functions.
The Electron Momentum Density is calculated along given directions (equation 8.23, page 181). The electron momentum distribution, EMD, is a non-periodic function; it falls rapidly to zero outside the first Brillouin zone. $\rho(\underline{0})$ gives the number of electrons at rest. The oblique coordinates directions given in input refer to the conventional cell, *not* to the primitive cell, for 3D systems.
Example: in a FCC system the input directions refer to the orthogonal unit cell frame (sides of the cube) not to the primitive non-orthogonal unit cell frame.

EMDP - Electron Momentum Density - plane maps

rec	variable	value	meaning
• ∗	NP		number of planes (< 5)
	IS		shrinking factor.
	IPLO	0	no data stored on disk.
		1	formatted output on Fortran unit 25 for plotting
	LPR115		printing of band projections
			insert NP set of records
• ∗	(L1(J),J=1,3), (L2(J),J=1,3)		fractional coordinates of the reciprocal lattice vectors that identify the plane
• ∗	PMX1		maximum p value along the first direction
	PMX2		maximum p value along the second direction
• ∗	NPO		number of orbital projections (≤ 10)
	NPB		number of band projections(≤ 10)
			if NPO ≠ 0 insert NPO set of records II
• ∗	NO		number of A.O.'s in the I-th projection
• ∗	IQ(I),I=1,NO		sequence number of the A.O.'s in the I-th projection - basis set order
			if NPB ≠ 0 insert NPB set of records II
• ∗	NB		number of bands in the I-th projection
• ∗	IB(I),I=1,NB		sequence number of the bands in the I-th projection

Warning EMDP does not work for UHF wave functions.
Calculation of electron momentum density on definite planes (equation 8.23, page 181).
The fractional coordinates of the reciprocal lattice vectors given in input refer to the conventional cell, *not* to the primitive cell, for 3D systems.
Example: in a FCC system the input directions refer to the orthogonal unit cell frame (sides of the cube) not to the primitive non-orthogonal unit cell frame.

END

Terminate processing of ***properties*** input. Normal end of the program ***properties***. Subsequent input records are not processed.

EXTPRT

See input block 1, page 32

FMWF - Wave function formatted output

The keyword **FMWF**, entered in ***properties*** input, writes formatted wave function data (same data are written in file fort.9, unformatted, at the end of SCF) in file fort.98 (LRECL=80). The formatted data can then be transferred to another platform. No input data required.
The resources requested to compute the wave function for a large system (CPU time, disk storage) may require a mainframe or a powerful workstation, while running ***properties*** is not so demanding, at least in terms of disk space. It may be convenient computing the wave function on a given platform, and the properties on a different one.
The keyword **RDFMWF**, entered in the first record of the ***properties*** input deck reads formatted data from file fort.98, and writes unformatted data in file fort.9. The key dimensions of the program computing the wave function and the one computing the properties are checked. If the dimensions of the arrays are not compatible, the program stops, after printing the PARAMETER statement used to define the dimension of the arrays in the program which computed the wave function. The sequence of the operations, when transferring data from one platform to another is the following:

platform	program	input	action
1	***properties***	**FMWF**	wave function formatted to file fort.9898
		ftp file fort.98 from platform 1 to platform 2	
2	***properties***	**RDFMWF**	wf read from file fort.98 (formatted) and written in file fort.9 (unformatted)

FRACTION - unit of measure

Unit of measure of coordinates in the periodic direction (**ECHG**, **POTM**, **CLAS**) See input block 1, page 35.

GRID3D - Selected property computed on a 3D grid

rec	variable	meaning
• ∗	NP	Number of points along the first direction
• A	CHARGE	electronic charge selected - see **ECH3** input
		or
• ∗	POTENTIAL	electronic charge selected - see **POT3** input

The propertty to be computed at the grid points is chosen by a keyword. Input as required by the selected property follows.
Computed data are written, formatted, in fortran unit 31. See Appendix E, page 218, for description of the format.

INFOGUI/INFO - output for visualization

No input data required.
Information on the system and the computational parameters are written formatted in fortran unit 32, in a format suitable for visualization programs.
See Appendix E, page 218, for description of the format.

ISOTROPIC - Fermi contact - Hyperfine electron-nuclear spin interaction isotropic component

rec	variable	meaning
• A	keyword	enter one of the following keywords:
	ALL	Fermi contact is evaluated for all the atoms in the cell
		or
	UNIQUE	Fermi contact is evaluated for all the non-equivalent atoms in the cell
		or
	SELECT	Fermi contact is evaluated for selected atoms
• ∗	N	number of atoms where to evaluate Fermi contact
• ∗	IA(I),I=1,N	*label* of the atoms

The spin density at the nuclei ($\langle\rho^{\rm spin}(\boldsymbol{r}_{\rm N})\rangle$) is evaluated. This quantity is given in $\rm bohr^{-3}$ and is transformed into the hyperfine coupling constant $\rm a_N[mT]$ through the relationship [88]

$$\rm a_N[mT] = \frac{1000}{(0.529177\cdot10^{-10})^3}\frac{2}{3}\;\mu_0\;\beta_N\;g_N\;\langle\rho^{spin}(\boldsymbol{r}_N)\rangle = 28.539649\;g_N\;\langle\rho^{spin}(\boldsymbol{r}_N)\rangle$$

where

$$\mu_0 = 4\pi\cdot10^{-7} = 12.566370614\cdot10^{-7}[\rm T^2J^{-1}m^3] \qquad \text{(permeability of vacuum)}$$

$$\beta_{\rm N} = 5.0507866\cdot10^{-17}[\rm JT^{-1}] \qquad \text{(nuclear magneton)}$$

the nuclear $\rm g_N$ factors for most of the nuclei of interest are available in the code and are taken from [88]. Conversion factors:

$$\rm a_N[MHz] = \frac{a_N[mT]g_e\beta_e}{10^9h[Js]} = 28.02.6\cdot a_N[mT]$$

$$\rm a_N[cm^{-1}] = \frac{a_N\;[MHz]10^8}{c[ms^{-1}]} = 0.33356410\cdot10^{-4}\cdot a_N[MHz]$$

$$\rm a_N[J] = g_e\;\beta_e\;10^{-3}a_N[mT] = 1.856954\cdot10^{-26}a_N[mT]$$

where:

$$\beta_{\rm e} = 9.2740154\cdot10^{-24}\;[\rm JT^{-1}] \qquad \text{(bohr magneton)}$$

$$\rm g_e = 2.002319304386 \qquad \text{(free-electron g factor)}$$

$$\rm c = 2.99792458\cdot10^8\;[ms^{-1}] \qquad \text{(speed of light in vacuum)}$$

$$\rm h = 6.6260755\cdot10^{-34}\;[Js] \qquad \text{(Planck constant)}$$

For extended printing (tensor in original cartesian axes and in principal axis system) insert, before the keyword ISOTROPIC:

```
SETPRINT
1
18 1
```

See tests 29, 31, 32, 33.

KNETOUT - Reciprocal lattice information - Fock/KS eigenvalues

Obsolete. See **CRYAPI_OUT**, page 116.

LOCALWF - Localization of Wannier Functions (WnF)

Wannier functions are computed from Crystalline Orbitals, and then localized, following the method described in [91]. The methods applies to non-conductor systems only.
The localization of Wannier Functions (WnF) is controlled by parameters. Default values are supplied for all parameters.
Optional keywords allow modification of the default choices, restricted to developers only.
The **LOCALWF** block is closed by the **END** keyword. The definition of plotting information (keyword **PRINTPLO**) must be in separated blocks, immediately following the first block. Each block defines the index number of WnF to be computed in a grid of points, followed by data defining the frame inside which the value of localized WnF has to be computed in a grid of points (see **MAPNET**, 130. The package Crgra2006 (http://www.crystal.unito.it/Crgra2006.html) allows plotting the function as contour lines.
For UHF calculations two set of blocks must be inserted for the α and β electrons, each one ending with the keyword **END**.

Definition of the set of bands considered in the localization process

VALENCE

Valence bands are chosen for localization.

OCCUPIED

All the occupied bands are chosen for localization [default].

INIFIBND

rec	variable	value	meaning
• ∗	IBAN		initial band considered for localization
	LBAN		last band considered for localization

BANDLIST

rec	variable	meaning
• ∗	NB	number of bands considered
• ∗	LB(I),I=1,NB	labels of the bands.

Tolerances for short and large cycles

A short cycle is a sequence of wannierization and Boys localization steps. The accuracies in both, the calculation of the Dipole Moments (DM) and the definition of the phases assigned to each periodically irreducible atom, are controlled so that they increase as the localization process evolves. This results in a significant saving of computational cost. Therefore, each time a given criterion is fulfilled, the accuracy in the DM evaluation increases and a new large cycle starts.

CYCTOL

rec	variable	value	meaning
• ∗	ITDP0	> 0	Initial tolerance used to calculate the DM matrix elements: $10^{-\mathrm{ITDP0}}$ [2]
	ITDP	> 0	Final tolerance used to calculate the DM matrix elements: $10^{-\mathrm{ITDP}}$ [5]
	ICONV	> 0	Convergence criterion to finish a large cycle: ABS(ADI(N) - ADI(N-1)) < $10^{-\mathrm{ICONV}}$, where ADI is the atomic delocalization index and N is the short cycle number [5]

PHASETOL

rec	variable	value	meaning
• ∗	ITPH0	> 0	$10^{-\text{ITPH0}}$ is the initial tolerance on the atomic charge population to attribute the phase to atoms in the wannierization step [2]
	ITPH	> 0	$10^{-\text{ITPH}}$ is the final tolerance used to attribute this phase [3]
	ICHTOL	> 0	DM tolerance of the cycle where ITPH0 changes to ITPH. [ITDP0+1]

General Keywords:

RESTART

With this option the WnF of a previous job are read from unit 81 (in the same format as output unit 80). The **RESTART** option set the same choice of the active bands as the previous job (and override any other definition) and the tolerances are, by default, the last attained in the previous calculation. The latter can be changed using **CYCTOL** and **PHASETOL**.

MAXCYCLE

rec	variable	value	meaning
• ∗	NCYC	> 0	maximal number of short cycles for the iterative process [30]

PRINTPLO

rec	variable	value	meaning
• ∗	IPRT	0	Does not print Wannier coefficients [default]
		> 0	Prints Wannier coefficients at each cycle up to the IPRT-th star of direct lattice vectors [0]
	IPRP	0	Prints population analysis only at the end of the localization.
		$\neq 0$	Prints analysis at each W-B cycle [0]
	ITPOP		Only atomic population larger than $10^{-\text{ITPOP}}$ are printed [2]
	IPLOT	0	WnFs are not computed for plot
		$\neq 0$	WnFs are computed in a grid of points, IPLOT being the number of stars of direct lattice vectors taken into account for WnF coefficients. Data are written in file fort.25 [0]

BOYSCTRL

Parameters that control the Boys localization step. Convergence of the process is achieved when the orbital-stability conditions: $B_{st} = 0; A_{st} > 0$, (see Pipek and Mezey 1989 [92]) are fulfilled for all pairs st of WnFs. Additionally, in order to avoid nearly free rotations (for instance in core or lone-pair WnFs) those pairs st with A_{st} close to 0 are not mixed (frozen).

rec	variable	value	meaning
• ∗	IBTOL		$10^{-\text{IBTOL}}$ is the threshold used for the stability condition on B_{st}. [4]
	IBFRZ		If for a pair of WnFs st, $\vert A_{st}\vert \leq 10^{-\text{IBFRZ}}$, then the corresponding WnFs are not mixed. [4]
	MXBCYC		Maximum number of cycles allowed in the Boys localization process [500]

Initial guess options

The iterative localization process of the WnFs needs to start from a reasonable initial guess. By default the starting functions are obtained automatically from the Hamiltonian eigenvectors at the Γ point. When required (pure covalent bonds that link atoms in different unit cells), a pre-localization is performed using a scheme similar to that suggested by Magnasco and Perico (1967) [93].

IGSSCTRL

Parameters used to control the pre-localization of the Γ point eigenvectors.

rec	variable	value	meaning
• ∗	CAPTURE		The capture distance between atoms I and J is given by CAPTURE ∗ (RAYCOV(I) + RAYCOV(J)) (RAYCOV, covalent radius (default value table page 40). An inter-atomic distance lower than the capture indicates that I and J can be covalently bonded. Default value [2.0].
	MPMAXIT		Maximum number of iterations in the pre-localization process 200
	ICNVMP		$10^{-\mathrm{ICNVMP}}$ is the convergence threshold for the Magnasco-Perico pre-localization 8
	IOVPOP		Just those pairs of atoms whose overlap population are greater than $10^{-\mathrm{IOVPOP}}$ are considered covalently bonded 4

The initial guess can be given as input in two mutually exclusive ways, controlled by the keywords **IGSSVCTS** and **IGSSBNDS**:

IGSSVCTS

The eigenvectors and the phases are given explicitly after the LOCALWF block (and before the plot parameters if required), in the following format.

rec	variable	value	meaning
• ∗	NGUES		Number of bands whose phase is pre-assigned such that the involved atoms are to be located in a given cell.
			insert 2 × NGUES records
• ∗	IB		
• ∗	IGAT(I,IB),I=1,NAF		
			Index of the direct lattice vector corresponding to the cell where atom I is expected to have the largest charge population in Wannier IB (NAF is the number of atoms per cell)
			insert:
GUESSV(I),I=NDF*NOCC			
			where NDF is the basis set dimension and NOCC the number of bands considered. GUESSV is a matrix containing the initial guess vectors for the iterative Wannier-Boys procedure (GUESSV is written in free format as a one-dimensional array).

IGSSBNDS

Use this option to explicitly indicate the WnFs that are to be assigned to covalent bonds.

rec	variable	value	meaning
• ∗	NBOND		Number of covalent bonds given as input.
			insert NBOND records
• ∗	NAT1		Label of the first atom of the bond, it is assumed to be located in the reference cell.
	NAT2		Label of the second atom of the covalent bond
	IC1,IC2,IC3		Indices of the cell where atom NAT2 is located
	NBNDORD		Bond Order

CAPTURE

The value of the CAPTURE parameter (see **IGSSCTRL** can be redefined.

rec	variable	value	meaning
• *	CAPTURE		The capture distance between atoms I and J is given by CAPTURE ∗ (RAYCOV(I) + RAYCOV(J)) (RAYCOV, covalent radius (default value table page 40). An inter-atomic distance lower than the capture indicates that I and J can be covalently bonded. Default value [2.0].

Plotting the WnFs

If IPLOT $\neq 0$ insert after the keyword block (defining the localization procedure computational parameters, and terminated by END):

rec	variable	value	meaning
• *	NWF		number of WnF to plot
			insert NWF blocks of data
• *	NUMBWF		sequence number (output order) of the WnF to plot
			MAPNET input data (Section 5.2, page 130)

The WnFs and the WnFs densities (in this order) within the selected regions are given in file fort.25.

1. The **NEWK** option must be executed before running **LOCALWF**, to compute the crystalline orbitals.
2. The number of **k** points required for a good localization depends on the characteristics of the bands chosen. For core electrons or valence bands in non-conducting materials, a IS twice than that used in the SCF part is enough to provide well localized WnFs. For valence bands in semiconductors or conduction bands the **k**-point net is required to be denser, but there are no recipes to determine *a priori* the optimum IS value. However, a necessary condition for the WnFs to be well represented, is that the volume in terms of number of unit cells of the cluster that contains the set of WnFs up to AO coefficients of 10^{-ITDP}, given as output, should be lower than the number of **k**-points in the net (IS**IDIM, being IDIM the dimensionality of the system).
3. The efficiency of the localization can be controlled using the **CYCTOL** parameters. In most cases, increasing ITDP and/or ICONV leads to larger and more accurate localization of the WnFs.
4. The **RESTART** option admits **MAXCYCLE** = 0, then the program just reconstructs all the information about the WnFs given in fortran unit 81 but does not continue the localization. This two options together with a IS=1 in **NEWK** is useful to perform the analysis of the WnFs after localization by means of the **PRINTPLO** option.

Symmetry adaptation of Wannier Functions (WnF)

SYMMWF

The procedure of symmetry adaptation of Wannier Functions [94] can be briefly outlined as follows:

1. WnFs are classified depending on the number of atoms that most contribute to it, on the basis of the atomic population analysis; in particular, WnFs will be defined *bond* or *atomic* if the charge density is mainly concentrated on one or two atoms, respectively. These atoms will be referred in the following as "main" atoms.

2. according to both shell population and symmetry properties of the main atoms, WnFs are grouped into subsets;

3. the symmetry of each subsets is verified;

4. in the case of subsets composed of non-symmetry related WnFs, the WnFs are projected into the sub-space defined by the point group of the subsets (a sub-group of the crystal point group). Each WnF becomes a representative function of one of the rows of the irreducible representation (IRREP) of the sub-group.

5. As a result of this procedure, each WnF is classified by four index (`b,f,p,g`) [bunch, flower, petal and crystal cell] such that a general symmetry operator $\widehat{W}$ of the crystal applied to a WnF gives:
   $\widehat{W}$ (`b,f,p,g`) $= \sum T^{W}_{pp'}$ (`b,f`^{W}`,p',g`^{W})

The WnF symmetrization procedure is mandatory in the case of a subsequent MP2 calculations. A set of optional keywords (to be used by developers only) allows modification of default setting of computational parameters. The **SYMMWF** input block must be closed by the keyword **END**.
To be used by developers only.

rec	variable	value	meaning
• A	**SYMMFLAG**		modification of symmetrization options
• I	IFSYM		Foster-Boys loc. procedure followed by WnFs symmetry adaptation is the default sequence
		1	WnFs symmetry adaptation and then Foster-Boys loc. procedure
		2	read WnFs from a previous run, (*fort.88*)
I	IFSAVE	1	save WnFs on fortran unit *fort.88* [default fort.80]
		2	after symm. adap., do not perform the re-wannierization step
		3	after symm. adap., do not perform the re-wannierization and re-orthogonalization steps
• A	**TOLBON**		redefinition of the tolerance to classify WFs as *bond* or *atomic* - default [0.2]
• F	TOLB		if $abs[pop(1)-pop(2)] < tolb$ then it is a *bond* WnF, where *pop* is the atomic population and 1, 2 are the two main atoms
• A	**TOLSHL**		redefinition of the tolerance to group WnFs according to a shell population analysis (performed on the main atoms) - default [0.01]
• F	DIFFSH		maximal difference between two shells
• A	**TOLPRO**		redefinition of the tolerance for WnF projection onto a given IRREP [0.8]
• F	TOLLIRR		minimal value of the norm of the extracted WnF component on a given IRREP
• A	**SYMVER**		WnF symmetry is verified by means of scalar products performed in a number of crystal cells defined by **g**-max
• I	g-max		value of **g**-max
• A	**FORCE**		in the case of WnF bunches describing double and triple bonds, it allows the choice of a given symmetry
• I	IRR		number of bunches to "force"
			insert IRR records
• 2I	NIRR IJMP		which bunch, numb. of IRREP to skip
• A	**END**		end of SYMMWF block - mandatory

The last two options of the **SYMMFLAG** keyword, (`IFSAVE=2,3`) are intended to maximally preserve the WF symmetry, to a small detriment of the local character.
The use of **FORCE** can be explained as follows: in the case of the acetyl crystal, the second

bunch belongs to a point-group of 2 IRREPs and contains three *bond* WnFs. The previous input cards (`FORCE\1\2 2`) referring to the second subset, yield its symmetrization according to the creation of a 3-dimensional IRREP (the first two IRREPs of the sub-group are skipped) with the three *bond* WnFs acting as basis function.

New keywords - developers only

CLUSPLUS

Upon transformation from Bloch Functions to Wannier Function, the latter are defined within a region with cyclic boundary condition imposed. We call it the "cyclic cluster". The volume of this region depends on the shrinking factor used in the previous NEWK. For instance, if IS=4, then the cyclic cluster in a 3D system will be 4**3 times larger than the primitive cell. For the localization part to work the WnFs are required to be described in the real space, hence the cyclic conditions and the WnFs are mapped onto a cluster in direct space. The size of this cluster where the localization is performed is defined as follows:

1. We define a small cluster, as a spherical region that contains the minimum number of G-vectors that fully map the cyclic cluster. Let' s call RO its radius.
2. As the centroid of some WnFs may be at the border of the reference cell we should consider some additional space in the direct cluster so as to allow the tails to be fully included in the region. This additional distance R1 is calculated as the maximum G-vector modulus of the set of cells at the neighbours of the reference one.
3. The radius of the resulting direct cluster will read: R = R0 + IPLUSCLUS*R1, where IPLUSCLUS is given in input. By default IPLUSCLUS is 5.

rec	variable	value	meaning
• ∗	iplusclus		factor to define the radius of direct cluster

ORTHNDIR

After the WANNIER-BOYS localization the WnFs are not fully orthonormal in direct space (they are just orthonormal within the cyclic cluster). To perform a true localization in direct space (see **FULLBOYS**) a previous re-orthonormalization in direct space is required. This is carried out by constructing the first order approximation of the Lowdin transformation and applying it to the WnFs. This process is performed iteratively up to fulfill a given criterion. ORTHNDIR sets the parameters that control this process.

rec	variable	value	meaning
• ∗	ISTORTH	> 0	number of stars of G-vectors that contains the transformation matrix.
		$= 0$	the number of stars is computed so as to contain the reference cell and all its neighbors [default].
	ITOLORTH	> 0	the overlap matrix elements are computed just between WnF components gt 10**-ITOLORTH in absolute value [default 5].
	NREORTHN	≥ 0	maximum number of iterations [default 10 in *properties*, 0 in *crystal*.
		< 0	the iterative procedure is performed up to the mean normalization error of the WnFs is $<$ 10**NREORTHN in absolute value (Default -7)

WANDM

WANDM controls the computation of the DM matrix elements between WnFs assigned to the reference and the neighboring cells (translational images of the former).

rec	variable	value	meaning
• ∗	INEIGH		controls the extent of the DM matrix by limiting the neighboring cell around the origin considered in the computation of the matrix elements:
		> 0	number of stars of neighboring cells considered for the matrix elements of DM
		< 0	the DM matrix is computed up to star of neighbor ISTAR with the condition that ABS(ALOCLEN(ISTAR)-ALOCLEN(ISTAR-1))¡10**(-ABS(INEIGH)), where ALOCLEN(ISTAR) means ”localization length computed up to star ISTAR”
• ∗	TOLDM		tolerance in the WnF coefficients used to calculate the DM matrix elements (see CYCTOL)

FULLBOYS

rec	variable	value	meaning
• ∗	ITOLWPG	> 0	TOLWPG 10**(-ITOLWPG) tolerance on the DM matrix elements

Request of Foster-Boys localization in direct space. The set of WnFs considered in the calculation of the DM matrix (see WANDM) are orthogonally transformed so as to obtain maximally localized WnFs under the Boys criterion. The resulting functions keep both, orthonormality and translational equivalence.
TOLWPG 10**(-TOLWPG) tolerance on the DM matrix elements to keep and use it in compact form. A small TOLWPG means that only a few DM matrix elements are considered in the localization process, then the calculation is quite fast and not very demanding in memory. A very large value would bring about very accurate LWFs with high computational cost. Recommended values: 4-6.

Bibliography

N. Ashcroft, D. Mermin ”Solid State Physics”, Holt Rinehart and Winston: New York, 1976.

P.-O. Löwdin (Editor) ”Quantum Theory of Atoms, Molecules and the Solid State”, Academic: New York, 1966.

S. F. Boys Rev. Mod. Phys 32 (1960) 296.

J. M. Foster and S. F. Boys Rev. Mod. Phys 32 (1960) 300.

J. Pipek and P. G. Mezey J. Chem. Phys 90 (1989) 4916.

V. Magnasco and A. Perico, J. Chem. Phys. 47 (1967) 971.

MAPNET - coordinates of grid points on a plane

This is a dummy keyword, to explain the way is generated the grid of points in which is evaluated a given function F: charge density and spin density (**ECHG**), electrostatic potential (**CLAS**, **POTM**). The graphic representation of the resulting 2D function is made by external software.

rec	variable	meaning
• ∗	NPY	number of points on the B-A segment.
• *A*	keyword	enter a keyword to choose the type of coordinate:
•	**COORDINA**	
• ∗	XA,YA,ZA	cartesian coordinates of point A
• ∗	XB,YB,ZB	cartesian coordinates of point B
• ∗	XC,YC,ZC	cartesian coordinates of point C
	or	
•	**ATOMS**	
• ∗	IA	*label* of the atom at point A
	AL,AM,AN	indices (direct lattice, input as reals) of the cell where the atom is located
• ∗	IB	*label* of the atom at point B
	BL,BM,BN	indices (direct lattice, input as reals) of the cell where the atom is located
• ∗	IC	*label* of the atom at point C
	CL,CM,CN	indices (direct lattice, input as reals) of the cell where the atom is located
	optional keyword	II
•	**RECTANGU**	definition of a new A'B'C'D' rectangular window, with B'C' on BC, A'D' on AD and diagonals A'C'=B'D'=max(AC,BD) (see Fig 5.1)
	optional keyword	II
•	**MARGINS**	definition of a new A",B",C",D" window including ABCD (or A'B'C'D') (see Fig 5.2)
• ∗	ABM	margins along AB
	CDM	margins along CD
	ADM	margins along AD
	BCM	margins along BC
	optional keyword	II
•	**PRINT**	printing of the values of the function in the net
•	**ANGSTROM**	cartesian coordinates in Ångstrom (default)
•	**BOHR**	cartesian coordinates in bohr
•	**FRACTION**	cartesian coordinates in fractionary units
•	**END**	end of **MAPNET** input block

1. Function F is mapped in a ABCD parallelogram-shaped domain defined by the sides AB and BC of any $\widehat{ABC}$ angle. F is calculated at the n_{AB} * n_{BC} nodes of a commensurate net (n_{AB} and n_{BC} integers).
2. If C ≡ B, F is calculated along the line AB. Data are written in file RHOLINE.DAT E.
3. n_{BC} is set by the program such that all points in the net are as equally spaced as possible ($\delta_{AB} \approx \delta_{BC}$).
4. formatted output is written in file fort.25 (processed by Crgra2006; see Appendix E, page 212.
5. The position of the three points A, B and C can be specified in two alternative ways:

   COORDINA the cartesian coordinates of the three points are given in bohr / Ångstrom / fractionary units (default Ångstrom; see Section 2.1, page 25)

   ATOMS A,B,C correspond to the position of 3 nuclei, identified by their sequence number in the reference cell, and the crystallographic indices of the cell in which they are located (input as real numbers).
6. The symmetry is used to restrict the calculation of the function to the irreducible part of the parallelogram chosen. To maximize the use of symmetry, the points of the net should include the low multiplicity positions in the selected plane. For example, B=(0,0,0), A=(a,0,0), C=(0,b,0) (a,b lattice vectors). Choose NPY=4n+1 for (100) faces of cubic crystals, or NPY = 6n+1 for (0001) faces of hexagonal crystals.

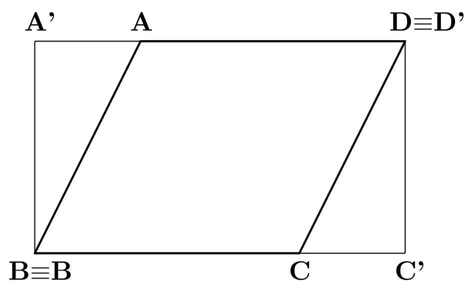


Figure 5.1: Definition of the window where the function F is mapped Effect of optional keyword RECTANGU.

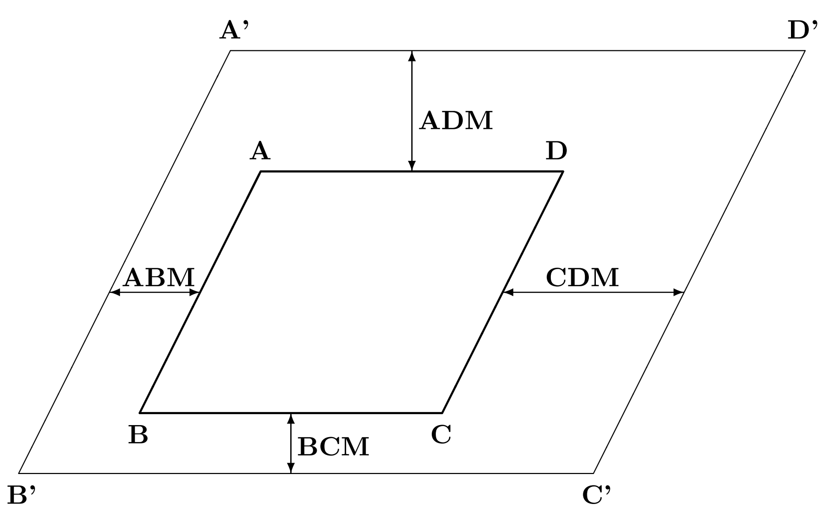


Figure 5.2: Definition of frame around the original window where the function F is mapped. Effect of optional keyword MARGINS.

MOLDRAW

See input block 1, page 36

NEIGHBOR/NEIGHPRT

See input block 1, page 38

NEWK - Fock/KS eigenvectors

rec	variable	value	meaning
			if system is periodic, insert II
• ∗	IS		Shrinking factor for reciprocal space net (Monkhorst net). The number NKF of **k** points, where the Fock/KS matrix is diagonalized, is roughly proportional to IS^{IDIM}/MVF where IDIM denotes the periodic dimensionality of the system, and MVF denotes the number of point symmetry operators (see page 18).
	ISP		Shrinking factor of the secondary reciprocal space net (Gilat net) for the evaluation of the Fermi energy and density matrix.
			if system is periodic and IS=0, insert II
• ∗			Shrinking factors of reciprocal lattice vectors
	IS1		Shrinking factor along B1
	IS2		Shrinking factor along B2
	IS3		Shrinking factor along B3.
• ∗	IFE	0	no Fermi energy calculation is performed;
		1	Fermi energy is computed, by performing integration on the new **k** points net. Total, valence and core density matrices are written on Fortran unit 13
	NPR		number of printing options to switch on
			if NPR ≠ 0 insert **prtrec** *(see page 42)* II

The Fock/KS eigenvectors are computed at a number of **k** points in reciprocal space, defined by the shrinking factor IS, and written unformatted in file fort.10 (in the basis of symmetry adapted Bloch functions) and in file fort.8 (in the basis of AO). Eigenvalues and related information (coordinates of k points in reciprocal lattice, weights etc) are written in file KIBZ.DAT by inserting the keyword **KNETOUT** (page 124). See Appendix C, page 206.

1. The Fock/KS matrix in direct space is always the SCF step final one. If the SCF convergence was poor, and convergence tools were used, eigenvalues and eigenvectors may be different from the ones that could be obtained after one more cycle without any convergence trick.

2. The shrinking factors IS and ISP (Section 8.7, page 180) can be redefined with respect to the ones used in the SCF process. If this value is smaller than the one used in the **scf** step, numerical inaccuracy may occur in the Fourier transform of the Fock/KS matrix, $F_g \rightarrow F_k$ (Chapter 8, equation 8.5).

3. A Fermi energy calculation must be performed (IFE=1) to run **PROF** the Compton profiles option, **PBAN** and **PDIDE** in order to compute the weight associated to each eigenvalue.

4. **Warning** NEWK with IFE=1 must be run to obtain the correct Fermi energy and eigenvalues spectra when a shift of eigenvalues was requested in SCF (**LEVSHIFT**, page 71; **SPINLOCK**, page 78; **BETALOCK**, 57.
   A new density matrix is computed. If the convergence of scf was poor, and convergence tools were used (FMIXING, LEVSHIFT, ..), the density matrix obtained from the eigenvectors computed by NEWK may be different from the matrix that could be calculated with one more scf cycle. Properties depending on the density matrix may be different if computed before or after NEWK.

5. if **BAND** is called after **NEWK**, and symmetry adapted Bloch functions are used (default option), the information generated by NEWK is destroyed. For instance, to compute density of states and bands, the sequence must be: BAND - NEWK - DOSS. The sequence NEWK BAND DOSS will give the error message:

   `NEWK␣MUST␣BE␣CALLED␣BEFORE␣DOSS`

Printing options: 59 (Density matrix - direct lattice); 66 (Hamiltonian eigenvalues); 67 (Hamiltonian eigenvectors).

NOSYMADA

See input block 3, page 73

PARAMPRT - - printing of parametrized dimensions

See input block 1, page 39.

PATO - Density matrix as superposition of atomic densities

rec	variable	value	meaning
• ∗	IBN	0	density matrix computed with the same basis set as in the crystal calculation.
		$\neq 0$	new basis set and/or new electron configuration is given
	NPR	$\neq 0$	printing of the density matrix for the first NPR direct lattice vectors
			if IBN $\neq$ 0 insert basis set input (page 14) II

1. The **PATO** option is used for calculating crystal properties, such as charge density (**ECHG**), structure factors (**XFAC**) with a periodic density matrix obtained as a superposition of atomic solutions (periodic array of non interacting atoms). The density matrix is written in fortran unit 13.

2. The atomic wave function is computed by the atomic program [6], using HF hamiltonian, *s, p, d* orbitals basis set, properly handling the open shell electronic configuration.

3. If the basis set used for the crystalline calculation (given as input of the **integral** part) is not suitable for describing a free- atom or free-ion situation, a new basis set can be supplied (see Section 1.2). When this option is used (IBN.NE.0) the basis set of *all* the atoms with different conventional atomic number has to be provided.

4. The electronic configuration of selected atoms may be modified (**CHEMOD** in basis set input). This allows calculation of the density matrix as superposition of atomic densities or ionic densities, for the same crystal structure.

5. The wave function data stored in file fort.9 at the end of the SCF cycles are not modified. Only the data stored on the temporary data set (reducible density matrix in fortran unit 13 and overlap matrix in fortran unit 3) are modified. The keyword **PSCF** restores the scf density matrix and all the original information (including geometry and basis set).

6. See also **ATOMHF**, input block 3, page 56, and **CHARGED**, input block 2, page 47.

PBAN/PDIBAN - Band(s) projected density matrix

rec	variable	meaning
• ∗	NB	number of bands to consider.
	NPR	printing of the density matrix for the first NPR direct lattice cells.
• ∗	N(I),I=1,NB	sequence number of the bands summed up for the projected density matrix.

A density matrix projected onto a given range of bands is computed and stored in fortran unit 13. The properties will subsequently be computed using such a matrix.
For spin polarized systems, two records are written:
first record, total density matrix (N=n_α + n_β electrons);
second record, spin density matrix (Ns=n_α - n_β electrons).

To be combined *only* with **ECHG** and **PPAN**. Fock/Kohn-Sham eigenvectors and band weights must be precomputed by running **NEWK** and setting IFE=1.

PGEOMW - Density matrix from geometrical weights

A density matrix projected onto the range of bands defined in input (see **PBAN** input instructions) is computed, using the geometrical weights of the **k** points in the reciprocal lattice. The properties will subsequently be computed using such a matrix. All the bands are attributed an occupation number 1., independently of the position of the Fermi energy. The density matrix does not have any physical meaning, but the trick allows analysis of the virtual eigenvectors.
For spin polarized systems, two records are written:
first record, total density matrix (N=n_α + n_β electrons);
second record, spin density matrix (Ns=n_α - n_β electrons).
To be combined *only* with **ECHG** and **PPAN**.
Fock/Kohn-Sham eigenvectors and band weights must be computed by running **NEWK** and setting IFE=1. Symmetry adaptation of Bloch functions is not allowed, the keyword NOSYMADA must be inserted before NEWK.

PDIDE - Density matrix energy projected

rec	variable	meaning
• ∗	EMI,EMAX	lower and upper energy bound (hartree)

A density matrix projected onto a given energy range is computed and stored in fortran unit 13. The properties will subsequently be computed using such a matrix. To be combined *only* with **DOSS**, **ECHG** and **PPAN**. Fock/Kohn-Sham eigenvectors and band weights must be computed by running **NEWK** and setting IFE=1.
The charge density maps obtained from the density matrix projected onto a given energy range give the STM topography [95] within the Tersoff-Haman approximation [96].

POLI - Spherical harmonics multipole moments

rec	variable	value	meaning
• ∗	IDIPO		multipole order (maximum order ℓ=6)
∗	ITENS	1	the quadrupole cartesian tensor is diagonalized
		0	no action
	LPR68		maximum pole order for printing:
		< 0	atom multipoles up to pole IDIPO
		≥ 0	atom and shell multipoles up to pole IDIPO

The multipoles of the shells and atoms in the primitive cell are computed according to a Mulliken partition of the charge density, up to quantum number IDIPO ($0 \leq$ IDIPO ≤ 6). The first nine terms, corresponding to ℓ=0,1,2 (for the definition of higher terms, see Appendix A1, page 170 in reference [22]) are defined as follow:

ℓ	m	
0	0	s
1	0	z
1	1	x
1	-1	y
2	0	$z^2 - x^2/2 - y^2/2$

2	1	$3xz$
2	-1	$3yz$
2	2	$3(x^2 - y^2)$
2	-2	$6xy$
3	0	$(2z^2 - 3x^2 - 3y^2)z$
3	1	$(4z^2 - x^2 - y^2)x$
3	-1	$(4z^2 - x^2 - y^2)y$
3	2	$(x^2 - y^2)z$
3	-2	xyz
3	3	$(x^2 - 3y^2)x$
3	-3	$(3x^2 - y^2)y$

If ITENS=1, the cartesian quadrupole tensor is computed, and its eigenvalues and eigenvectors are printed after diagonalization. The components of the cartesian tensor are: $x^2, y^2, z^2, xy, xz, yz$
Warning: the shell multipoles are *not* printed by default. On request (keyword **POLIPRT**), they are printed in atomic units (electron charge = +1).

POLSPIN - Spin multipole moments

rec	variable	value	meaning
• ∗	IDIPO		multipole order (maximum order ℓ=6)
∗	ITENS	1	the quadrupole cartesian tensor is diagonalized
		0	no action
	LPR68		maximum pole order for printing:
		< 0	atom multipoles up to pole IDIPO
		≥ 0	atom and shell multipoles up to pole IDIPO

The electron spin density is partitioned in atomic contributions according to the Mulliken scheme, and the spherical harmonic atomic multipoles up to the IDIPO angular quantum number are evaluated (see the **POLI** keyword for definition of the multipoles and references). The Cartesian tensor $T_{ij} = \int x_i x_j \, \rho^{spin}(\mathbf{r}) \, d\mathbf{r}$ is computed and diagonalized, and its eigenvalues and eigenvectors are printed. This option may be useful in the analysis of the size, shape and orientation of localized electron holes.

POT3 - Electrostatic potential on a 3D grid

rec	variable	meaning
• ∗	NP	Number of points along the first direction
• ∗	ITOL	penetration tolerance (suggested value: 5) (see **POTM**, page 138)

if non-3D system
keyword to choose the type of grid on the non-periodic direction(s):

SCALE length scales for non-periodic dimensions	**RANGE** boundary for non-periodic dimensions (au)
if 2D system	
• ∗ ZSCALE	• ∗ ZMIN
	• ∗ ZMAX
if 1D system	
• ∗ YSCALE,ZSCALE	• ∗ YMIN,ZMIN
	• ∗ YMAX,ZMAX
if 0D system	
• ∗ XSCALE,YSCALE,ZSCALE	• ∗ XMIN,YMIN,ZMIN
	• ∗ XMAX,YMAX,ZMAX

The electrostatic potential is computed at a regular 3-dimensional grid of points. The grid is defined by the lattice vectors of the primitive unit cell and user defined extents in non-periodic directions. NP is the number of points along the first lattice vector (or XMAX-XMIN for a molecule). Equally spacing is used along the other vectors. Non-periodic extents may be specified as an explicit range (RANGE) or by scaling the extent defined by the atomic coordinates (SCALE).
Formatted data are written in fortran unit 31 (function value at the grid points), in the format required by the visualization program DLV.
See Appendix E, page 218, for description of the format.

POTC - Electrostatic potential and its derivatives

rec	variable		meaning
• ∗	ICA	0	calculation of potential (V), its first derivative (E) and second derivatives (E') in one or more points
		1	not implemented
		2	calculation of $V(z)$, $E(z)$, $E'(z)$ and $\rho(z)$ averaged in the plane at z position (2D only)
		3	calculation of $V(z)$, $E(z)$, $E'(z)$ and $\rho(z)$ averaged in the volume between z–ZD and z+ZD (2D only)
	NPU	n	number of points at which these properties are computed
		0	these properties are computed at the atomic positions defined by IPA value
	IPA	0	calculations are performed at each atomic positions in the cell
		1	calculations are performed just for non equivalent atomic positions in the cell
			if ICA = 0 then
			if NPU > 0 insert NPU records II
• ∗	X,Y,Z		point coordinates (cartesian, bohr)
			if NPU < 0 data are read from file POTC.INP II
			if ICA = 2 insert II
• ∗	ZM,ZP		properties are averaged over NPU planes orthogonal to the z axis from z = ZP to z = ZM by step of (ZP–ZM)/(NPU–1) (bohr)
			if ICA = 3 insert II
• ∗	ZM,ZP		properties are averaged over NPU volumes centered on planes orthogonal to the z axis, same as ICA = 2
	ZD		half thickness of the volume (bohr)

The exact electrostatic potential V, its derivatives E (electric field) and E' (electric field gradient) are evaluated for molecules (0D), slabs (2D) and crystals (3D). Plane and volume averaged properties can be computed for slabs (2D) only. The plane is orthogonal to the z axis.
For ICA = 3, the volume average is performed around a middle plane at z position, from z–ZD to z+ZD, giving a thickness of 2∗ZD.
According to Poisson's law, the charge density $\rho(z)$ is defined as

$$\rho(z) = -\frac{1}{4\pi}\frac{d^2V(z)}{dz^2} = \frac{-E'(z)}{4\pi}$$

If an electric field of intensity E_0 is present (keyword **FIELD**, see page 2.1, only for slabs), the total potential $V_{field}(z)$ is calculated:

$$V_{field}(z) = V(z) - E_0 z$$

where $V(z)$ is the potential of the slab itself and $-E_0 z$ is the perturbation applied.

- ICA = 0 ; NPU > 0; 2D or 3D system
  It is possible to enter the cartesian coordinates (bohr) of the points where the exact value of the properties must be computed. It is useful when applying fitting procedure to obtain formal point charges.

- ICA = 0 ; NPU < 0; 2D or 3D system
  coordinates in bohr are read (free format) from file POTC.INP Data are read in free format.

```
record        type of data      content
  1           1 integer       N, number of points
 2..2+N-1     4 real          x   y   z
```

- ICA ≠ 0; NPU ≠ 0; 2D or 3D system
  The data computed are written in file POTC.DAT. See Appendix E, page 215.

POTM - Electrostatic potential maps and electric field

rec	variable	value	meaning
• *	IDER	0	the electrostatic potential is evaluated
		1	the potential and its first derivatives are evaluated
	ITOL		penetration tolerance (suggested value: 5)
			insert **MAPNET** input records (page 130)

1. When IDER=0, the electrostatic potential is calculated at the nodes of a 2-dimensional net in a parallelogram-shaped domain defined by the segments AB and BC (see keyword **MAPNET**, page 130). The electrostatic potential values are written formatted in file fort.25 (see Appendix E, page 212).
2. When IDER ≠ 0, the electrostatic potential gradient is computed at the nodes of the same grid. The x, y and z components are printed in the standard output, and written formatted in file fort.25 (see Appendix E, page 212).
3. The electrostatic potential at **r** is evaluated [97] by partitioning the periodic charge density $\rho(r)$ in shell contributions ρ_λ^h:
$$\rho(\underline{r}) = \sum_h \sum_\lambda \rho_\lambda(\underline{r} - \underline{h})$$
(h translation vector).
4. The long range contributions are evaluated through a multipolar expansion of $\rho_\lambda(\underline{r} - \underline{h})$ [98]. The short range contributions are calculated exactly.
5. The separation between long and short range is controlled by ITOL: $\rho_\lambda(\mathbf{r}-\mathbf{h})$ is attributed to the short range (exact) region if
$$e^{-\alpha_\lambda(\mathbf{r}-\mathbf{s}_\lambda-\mathbf{h})^2} > 10^{-ITOL}$$
where: α_λ = exponent of the adjoined gaussian of shell λ; $\mathbf{s}_\lambda$ = internal coordinates of shell λ in cell at $\mathbf{h}$.

   The difference between the exact and the approximated potential is smaller than 1% when ITOL=5 (input datum to **POTM**), and IDIPO=4 (input datum to **POLI**), and smaller than 0.01% when ITOL=15 and IDIPO=6 [97, 98].
6. The multipoles of shell charges are computed if **POLI** option was not run before **POTM**.

PPAN/MULPOPAN - Mulliken Population Analysis

See input block 3, page 74.

PRINTOUT - Setting of printing environment

See input block 1, page 40.

PROF - Compton Profiles

rec	variable	value	meaning
• ∗	ICORE	1	core plus valence calculation.
		2	core only calculation.
		3	valence only calculation.
	IVIA	0	valence contribution is computed by numerical integration.
		1	valence contribution is computed analytically.
	NPR		number of printing options to switch on.
	IPLO	0	CP related data are not stored on disk
		1	formatted CP data stored in file fort.25 (Appendix E, page 214)
		2	formatted CP data stored in Fortran unit 24 (Appendix E, page 214)
			if NPR ≠ 0 insert **prtrec** *(see page 42)* II
• A2		**CP**	calculation of Compton profiles (J(q)) along selected directions (eq. 8.27).
• ∗	ND		number of directions (≤ 6).
	REST		maximum value of q for J(q) calculation ($bohr^{-1}$).
	RINT		internal sphere radius ($bohr^{-1}$).
	IRAP		shrinking factor ratio.
• ∗	(KD(J,N), J=1,3), N=1,ND		directions in oblique coordinates; see note 9
• ∗	STPJ		interpolation step (in interpolated Compton profiles calculation).
• A4		**DIFF**	CP difference between all computed directional CPs.
• A2		**BR**	autocorrelation function B(r) calculation (eq. 8.30).
• ∗	RMAX		maximum r value (bohr) at which B(r) is computed
	STBR		step in computation of B(r).
• A4		**CONV**	convolution of the data previously computed (CP, DIFF, BR) (eq. 8.29)
• ∗	FWHM		convolution parameter (a.u.) full width half maximum; $\sigma = \sqrt{(FWHM)^2/(8 \cdot 2log2)}$.
• A4		**ENDP**	End of input records for CP data

The keyword **PROF** starts the calculation of Compton profiles (J(q)) along selected directions (eq. 8.27). The specific keywords **DIFF BR CONV** allow the calculation of the related quantities. The card with the keyword **ENDP** ends the Compton profiles input section.

1. The input of the options must be given in the order in which they appear in the above description. To enter this property, the **CP** option must always be selected after **PROF**, while the others are optional.

2. Core and valence contributions are computed by using different algorithms. Core contribution to CP's is always computed analytically via the Pg matrix (direct lattice summation, equation 8.25); the valence contribution is computed numerically (IVIA=0) by integrating the EMD (equation 8.23). Valence contribution can be evaluated analytically, setting IVIA=1.

3. The numerical integration is extended to a sphere (radius RINT) where EMD is sampled at the points of a commensurate net characterized by a shrinking factor IS (in the IBZ) and at all the points (with modulus less then RINT) obtained from these by applying reciprocal lattice translations.

   It is possible to define a second sphere (with radius REST); in the volume between the two spheres a second net is employed with shrinking factor IS1 greater then IS. IRAP=IS1/IS is given in INPUT card 2; a reasonable value is IRAP=2. The outer contribution is supposed to be the same for different CP's, and is obtained by integrating the average EMD.

4. If ICORE $\neq$ 2 (valence electron CP's are required) the **NEWK** option, with IFE=1, must be run before the **PROF** option, in order to generate the eigenvectors required for the EMD calculation, as well as the weights associated with each k point.

5. If ICORE $\neq$ 2 and IVIA = 0 the CPs are evaluated at points resulting from the IS partition of the reciprocal lattice translators. The interpolation is performed at STPJ intervals (STPJ is given in input).

   If ICORE = 2 or IVIA = 1 the CPs are, in any case, evaluated at points resulting from STPJ intervals.

   IVIA=0 (numerical integration) produces much more accurate results;

   IVIA=1 (analytical integration) is to be used only for molecular calculations or for non conducting polymers.

6. Total CP's are always obtained by summing core and valence contributions.

7. Reasonable values of the integration parameters depend on the system under investigation. The normalization integral of the CP's is a good check of the accuracy of the calculation. For instance, in the case of the valence electron of beryllium (test 9), good values of RINT and IS are 10. a.u. and 4 respectively. In the case of silicon (test 10), good values of the same variables are 8. a.u. and 8 respectively. Much greater RINT values are required in order to include all the core electrons (70. a.u. in the case of silicon, and 25. a.u. in the case of beryllium).

8. BR (autocorrelation function or reciprocal space form factor) should be calculated only for valence electrons. All electron BR are reliable when the normalization integral, after the analytical integration for core electrons contribution, is equal to the number of core electrons.

9. The oblique coordinates directions given in input refer to the conventional cell, *not* to the primitive cell for 3D systems.

   Example: in a FCC system the input directions refer to the orthogonal unit cell frame (sides of the cube) not to the primitive non-orthogonal unit cell frame.

Printing options: 116 (Compton profiles before interpolation); 117 (average EMD before interpolation); 118 (printing of core, valence etc. contribution). The LPRINT(118) option should be used only if ICORE=1, that is, if core plus valence calculation are chosen.

PSCF - Restore SCF density matrix

The wave function data computed at the last SCF cycle are restored in common areas and fortran units 3 (overlap matrix), 11 (Fock/KS matrix), 13 (density matrix). The basis set defined in input block 2 is restored. Any modification in the default settings introduced in ***properties*** is overwritten. No input data required.

RAYCOV - covalent radii modification

See input block 1, page 40

ROTREF - Rotation of eigenvectors and density matrix

This option permits the rotation of the cartesian reference frame before the calculation of the properties.
It is useful, for example, in the population analysis or in the AO projected density of states of systems containing transition metal atoms with partial d occupation.
Consider for example a d^7 occupation as in CoF_2, where the main axis of the (slightly distorted) CoF_6 octahedron in the rutile structure makes a 45^0 angle with the x axis, and lies in the xy plane, so that the three empty β states are a combination of the 5 d orbitals. Re-orienting the octahedron permits to assign integer β occupations to d_{xz} and d_{yz}.

Input for the rotation as for **EIGSHROT** (page 66)

SETINF - Setting of INF values

See input block 1, page 42

SETPRINT - Setting of printing options

See input block 1, page 42.

STOP

Execution stops immediately. Subsequent input records are not processed.

SYMADAPT

See input block 3, page 79

XFAC - X-ray structure factors

rec	variable	value	meaning
• ∗	ISS	> 0	number of reflections whose theoretical structure factors are calculated.
		< 0	a set of non-equivalent reflections with indices h,k,l $< \|ISS\|$ is generated
			if ISS > 0 insert ISS records II
• ∗	H,K,L		Miller indices of the reflection (conventional cell) .

The Fourier transform of the ground state charge density of a crystalline system provides the static structure factors of the crystal, which can be determined experimentally, after taking into account a number of corrective terms, in particular those related to thermal and zero point motion of nuclei:

$$F_{\underline{k}} = \int \rho(\underline{r})\ e^{i\underline{k}\cdot\underline{r}} dr$$

where $\underline{k} \equiv h\ \underline{b}_1 + k\ \underline{b}_2 + l\ \underline{b}_3$. *The Miller indices refer to the conventional cell.* The structure factors are integrated over the primitive cell volume.

5.3 Spontaneous polarization and piezoelectricity

Y. Noel, September 2002 - not fully updated to CRYSTAL06

PIEZOBP - Piezoelectricity (Berry phase approach)

The calculation the piezoelectric constants of a system, can be decompose in few steps. A preliminary run must be performed for the undistorted system ($\lambda = 0$) with the keyword **POLARI**. Then, for a first distorted system ($\lambda = 1$), a second preliminary run (with the keyword **POLARI**) must be performed, followed by third run with the keyword **PIEZOBP** that calculates a approximated value of the piezoelectric constants. The evaluation of the slope $\frac{d\varphi_\alpha}{d\epsilon_{jk}}$ is computed with a single point. For more accuracy, other runs must be done for other distortions (one run with the keyword **POLARI** of the new system, followed by a second run with the keyword **PIEZOBP** with the undistorted and the new distorted systems). Then the mean value of the obtained piezoelectric constants must be performed.

1. First run: preliminary calculation related to system $\lambda = 0$ (undistorted)

Program	Keyword	comments
crystal		see deck 1 for input blocks 1 and 1b
properties	**NEWK**	additional keywords allowed
	POLARI	see above
save Fortran unit 27 as undistord.f27		

2. Second run: preliminary calculation related to system $\lambda = 1$ (distorted)

Program	Keyword	comments
crystal		see deck 1 for input blocks 1 and 1b
properties	**NEWK**	same input as in first run
	POLARI	
save Fortran unit 27 as distord1.f27		

3. Third run: merging of previous data.

copy undistord.f27 to Fortran unit 28		
copy distord1.f27 to fortran unit 29		
Program	Keyword	comments
properties	**PIEZOBP**	

4. Refine the computed value

Repeat 2. and 3. for several distortions.
Then compute the mean value of the piezoelectric constants obtained in each case.

PIEZOWF - Piezoelectricity (localized CO approach)

The calculation the piezoelectric constants of a system, can be decompose in few steps. A preliminary run must be performed for the undistorted system ($\lambda = 0$) with the keyword **LOCALWF**. Then, for a first distorted system ($\lambda = 1$), a second preliminary run (with the keyword **LOCALWF**) must be performed, followed by third run with the keyword **PIEZOWF** that calculates a approximated value of the piezoelectric constants. The evaluation of the slope $\frac{d\varphi_\alpha}{d\epsilon_{jk}}$ is computed with a single point. For more accuracy, other runs must be done for other distortions (one run with the keyword **LOCALWF** of the new system, followed by a second run with the keyword **PIEZOWF** with the undistorted and the new distorted systems). Then the mean value of the obtained piezoelectric constants must be performed.

1. First run: preliminary calculation related to system $\lambda = 0$ (undistorted)

Program	Keyword	comments
crystal		see deck 1 for input blocks 1 and 1b
properties	**NEWK**	additional keywords allowed
	LOCALWF	see above
save Fortran unit 37 as undistord.f37		

2. Second run: preliminary calculation related to system $\lambda = 1$ (distorted)

Program	Keyword	comments
crystal		see deck 1 for input blocks 1 and 1b
properties	**NEWK**	same input as in first run
	LOCALWF	
save Fortran unit 37 as distord1.f37		

3. Third run: merging of previous data.

copy undistord.f37 to Fortran unit 38		
copy distord1.f37 to Fortran unit 39		
Program	Keyword	comments
properties	**PIEZOWF**	

4. Refine the computed value

Repeat 2. and 3. for several distortions.
Then compute the mean value of the piezoelectric constants obtained in each case.

POLARI - Spontaneous polarization (steps 1 and 2)

PHASE - Spontaneous polarization (step 3)

The ferroelectric phases of a ferroelectric material exhibit two possible enantiomorphic non centrosymmetric structures, which can be labelled by the geometric parameters λ=+1 and λ=−1. An external electric field can force the system to change from one structure to the other, passing through a small energy maximum. The centrosymmetric unstable structure which sits in the middle of the λ=+1 and λ=−1 structures can be labelled by the geometric parameters λ=0.
The spontaneous polarization in ferroelectric materials is then evaluated through the Berry phase approach [99, 100] as the polarization difference between one of the two enantiomorphic structures (λ=+1 or λ=−1) and the intermediate geometric structure (λ=0).
The fortran unit 70 is defined as direct access:

```
OPEN(UNIT=IO70,ACCESS='DIRECT',RECL=LREC)
```

Three subsequent runs are required.

1. First run: preliminary calculation related to λ=0 structure

Program	Keyword	comments
crystal		see deck
properties	**NEWK**	see page 133
	POLARI	no input data required
move fortran unit 27 to zero.f27		

2. Second run: preliminary calculation related to λ=+1 (or λ=−1) structure

Program	Keyword	comments
crystal		see deck 2
properties	**NEWK**	same input as in first run
	POLARI	no input data required
move fortran unit 27 to one.f27		

3. Third run: merging of previous data

	fortran unit 9 of previous calculation must be present copy zero.f27 to fortran unit 28 copy one.f27 to fortran unit 29

Program	Keyword	comments
properties	**PHASE**	no input data required
	END	terminate processing of polari keywords

Deck 1

Potassium niobate - $KNbO_3$

CRYSTAL	3D system
0 0 0	IFLAG IFHR IFSO
123	space group, *P4/mmm*
3.997 4.063	lattice parameters
4	4 non equivalent atoms (5 atoms in the primitive cell)
19 0.5 0.5 0.5	Z=19, Potassium; *x, y, z* (multiplicity 1)
8 0.0 0.0 0.5	Z=8, Oxygen I; *x, y, z* (multiplicity 1)
8 0.5 0.0 0.0	Z=8, Oxygen II; *x, y, z* (multiplicity 2)
41 0.0 0.0 0.0	Z=41, Niobium; *x, y, z* (multiplicity 1)
END	end of geometry input records

Deck 2

Potassium niobate - $KNbO_3$

CRYSTAL	3D system
0 0 0	IFLAG IFHR IFSO
123	space group, *P4/mmm*
3.997 4.063	lattice parameters
4	4 non equivalent atoms (5 atoms in the primitive cell)
19 0.5 0.5 0.5	Z=19, Potassium; *x, y, z* (multiplicity 1)
8 0.0 0.0 0.5	Z=8, Oxygen I; *x, y, z* (multiplicity 1)
8 0.5 0.0 0.0	Z=8, Oxygen II; *x, y, z* (multiplicity 2)
41 0.0 0.0 0.0	Z=41, Niobium; *x, y, z* (multiplicity 1)
FRACTION	fractional coordinates
ATOMDISP	displacement of atoms
4	four atoms to be displaced
1 0.0 0.0 -0.023	displacement of atom no. 1 (Potassium)
2 0.0 0.0 -0.042	displacement of atom no. 2 (Oxygen II)
3 0.0 0.0 -0.042	displacement of atom no. 3 (Oxygen II)
4 0.0 0.0 -0.040	displacement of atom no. 4 (Oxygen I)
END	end of geometry input records

1. This subprogram works for 3D systems only.

2. The unit-cell has to contain an even number of electrons.

3. Cell parameters have to be the same for whatever value of the geometric parameter λ. The difference between the λ=+1, λ=0, and λ=−1 structures is only in the atomic positions.

4. Numerical accuracy and computational parameters in input block 3 (such as **TOLINTEG**, **POLEORDR**, etc.) should be the same for the first and the second run.

5. See page 133 for the **NEWK** input, which has to be the same for the first and the second

run. The shrinking factor IS should be at least equal to 4. Fermi energy calculation is not necessary, then set IFE=0.

6. Data evaluated with the keyword **POLARI** in the first two runs do not have any physical meaning if considered independently. Only the output produced choosing the keyword **PHASE** in the third run is significant.

7. When the λ=−1 geometric structure is chosen in the second run, the spontaneous polarization vector obtained at the end will have the same modulus and direction but opposite versus with respect to the vector obtained by choosing the λ=+1 structure.

8. The spontaneous polarization is obtained through the Berry phase approach. Since a phase is defined only in the interval $-\pi$ to $+\pi$, each component of the spontaneous polarization vector is defined to within an integer number (positive or negative) of the correspondent component of the "quantum of polarization" vector, which is automatically shown in the output of the third run.
   Usually there is not need to take into account the quantum of polarization vector, unless the ferroelectric material shows a large value of the spontaneous polarization.
   In case of doubt whether the quantum of polarization vector has to be considered or not, it is possible to evaluate the spontaneous polarization by setting in the second run a geometric structure corresponding to an intermediate geometric parameter, e.g. λ=0.25, and then to extrapolate linearly the result to the λ=1 structure.

`SPOLBP` - Spontaneous polarization (Berry phase approach)

To calculate the spontaneous polarization, a preliminary with the keyword `POLARI` run is needed for each of the two systems $\lambda = 1$ and $\lambda = 0$. Then a third run with the keyword `SPOLBP` gives the difference of polarization between both systems.

1. First run: preliminary calculation related to system $\lambda = 0$

Program	Keyword	comments
crystal		see deck 1 for input blocks 1 and 1b
properties	`NEWK`	additional keywords allowed
	`POLARI`	see above
save Fortran unit 27 as sys0.f27		

2. Second run: preliminary calculation related to system $\lambda = 1$

Program	Keyword	comments
crystal		see deck 1 for input blocks 1 and 1b
properties	`NEWK`	same input as in first run
	`POLARI`	
save Fortran unit 27 as sys1.f27		

3. Third run: merging of previous data.

copy sys0.f27 to Fortran unit 28		
copy sys1.f27 to Fortran unit 29		
Program	Keyword	comments
properties	`SPOLBP`	

SPOLWF - The spontaneous polarization (localized CO approach)

To calculate the spontaneous polarization, two preliminary runs with the keyword **LOCALI** is needed for each of the two systems $\lambda = 1$ and $\lambda = 0$. Then a third run with the keyword **SPOLBP** computes the difference of polarization between both systems.

1. First run: preliminary calculation related to system $\lambda = 0$

Program	Keyword	comments
crystal		see deck 1 for input blocks 1 and 1b
properties	`NEWK`	additional keywords allowed
	`LOCALI`	see above
save Fortran unit 27 as sys0.f27		

2. Second run: preliminary calculation related to system $\lambda = 1$

Program	Keyword	comments
crystal		see deck 1 for input blocks 1 and 1b
properties	`NEWK`	same input as in first run
	`LOCALI`	
save Fortran unit 27 as sys1.f27		

3. Third run: merging of previous data.

copy sys0.f27 to Fortran unit 28		
copy sys1.f27 to Fortran unit 29		
Program	Keyword	comments
properties	`SPOLWF`	

Chapter 6

Input examples

6.1 Standard geometry input

3D - Crystalline compounds - 1st input record keyword: CRYSTAL

Atom coordinates: fractional units of the crystallographic lattice vectors.

Sodium Chloride - NaCl (Rock Salt Structure)

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
225	space group, *Fm3m*, cubic
5.64	a (Å)
2	2 non equivalent atoms
11 .5 .5 .5	Z=11, Sodium, $1/2, 1/2, 1/2$
17 .0 .0 .0	Z=17, Chlorine

Diamond - C (2^{nd} Setting - 48 symmops - 36 with translational component)

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
227	space group, *Fd3m*, cubic
3.57	a (Å)
1	1 non equivalent atom
6 .125 .125 .125	Z=6, Carbon, $1/8, 1/8, 1/8$ (multiplicity 2)

Diamond - C (1^{st} Setting - 48 symmops - 24 with translational component)

CRYSTAL	
0 0 1	IFLAG IFHR IFSO
227	space group 227, *Fd3m*, cubic
3.57	a (Å)
1	1 non equivalent atom
6 .0 .0 .0	Z=6, Carbon (multiplicity 2)

Zinc Blend - ZnS

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
216	space group 216, $F\bar{4}3m$, cubic
5.42	a (Å)
2	2 non equivalent atoms
30 .25 .25 .25	Z=30, Zinc, $(1/4, 1/4, 1/4)$
16 .0 .0 .0	Z=16, Sulphur

Wurtzite - ZnS

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
186	space group 186, $P6_3mc$, hexagonal
3.81 6.23	a,c (Å)
2	2 non equivalent atoms
30 .6666666667 .3333333333 .0	Zinc, $(2/3, 1/3, 0.)$
16 .6666666667 .3333333333 .375	Sulphur, $(2/3, 1/3, 3/8)$

Cuprite - Cu_2O

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
208	space group 208, $P4_232$, cubic
4.27	a (Å)
2	2 non equivalent atoms
8 .0 .0 .0	Z=8, Oxygen
29 .25 .25 .25	Z=29, Copper $(1/4, 1/4, 1/4)$

Aragonite - $CaCO_3$

CRYSTAL	
1 0 0	IFLAG (1, SPGR symbol) IFHR IFSO
P M C N	space group $Pmcn$, orthorhombic
4.9616 7.9705 5.7394	a,b,c (Å)
4	4 non equivalent atoms
20 .25 .4151 .2103	Z=20, Calcium
6 .25 .7627 .085	Z=6, Carbon
8 .25 .9231 .0952	Z=8, Oxygen
8 .4729 .6801 .087	Z=8, Oxygen

Fluorite - CaF_2

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
225	space group 225, $Fm3m$, cubic
5.46	a (Å)
2	2 non equivalent atoms
9 .25 .25 .25	Fluorine
20 .0 .0 .0	Calcium

Cesium chloride - CsCl

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
221	space group 221, $Pm3m$, cubic
4.12	a (Å)
2	2 non equivalent atoms
55 .5 .5 .5	Cesium
17 .0 .0 .0	Chlorine

Rutile - TiO_2

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
136	space group 136, $P4_2/mnm$, tetragonal
4.59 2.96	a, c (Å)
2	2 non equivalent atoms
22 .0 .0 .0	Titanium
8 .305 .305 .0	Oxygen

Graphite - C (Hexagonal)

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
194	space group 194, $P6_3/mmc$, hexagonal
2.46 6.70	a,c (Å)
2	2 non equivalent atoms
6 .0 .0 .25	Carbon, $0, 0, 1/4$
6 .33333333333 .66666666667 .25	Carbon, $1/3, 2/3, 1/4$

Pyrite - FeS_2

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
205	space group 205, $Pa3$, cubic
5.40	a (Å)
2	2 non equivalent atoms
26 .0 .0 .0	Iron
16 .386 .386 .386	Sulphur

Calcite - $CaCO_3$

CRYSTAL	
0 1 0	IFLAG IFHR *(=1, rhombohedral representation)* IFSO
167	space group 167, $R\bar{3}c$, hexagonal
6.36 46.833	a (Å), $\alpha°$
3	3 non equivalent atoms
20 .0 .0 .0	Calcium
6 .25 .25 .25	Carbon
8 .007 .493 .25	Oxygen

Corundum - Al_2O_3 (hexagonal representation)

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
167	space group 167, $R\bar{3}c$, hexagonal
4.7602 12.9933	a,c (Å)
2	2 non equivalent atoms
13 0. 0. 0.35216	Aluminium
8 0.30621 0. 0.25	Oxygen

Corundum - Al_2O_3 (rhombohedral representation)

CRYSTAL	
0 1 0	IFLAG IFHR *(=1, rhombohedral cell)* IFSO
167	space group 167, $R\bar{3}c$, hexagonal
5.12948 55.29155	a (Å), $\alpha°$
2	2 non equivalent atoms
13 0.35216 0.35216 0.35216	Aluminium
8 0.94376 0.25 0.55621	Oxygen

Zirconia - ZrO_2 - monoclinic structure

CRYSTAL	
0 0 1	IFLAG IFHR IFSO *(=1, standard shift of origin)*
14	space group 14, $P2_1/c$, monoclinic
5.03177 5.03177 5.258 90.0	a,b,c (Å), $\beta°$
3	3 non equivalent atoms
240 0.2500 0.0000 0.25000	Zirconium, Pseudopotential (Z' > 200)
208 0.0000 0.2500 0.07600	Oxygen, Pseudopotential
208 -0.500 -0.250 0.07600	Oxygen, Pseudopotential

Zirconia - ZrO_2 - tetragonal structure

CRYSTAL	
0 0 1	IFLAG IFHR IFSO *(=1, standard shift of origin)*
137	space group 137, $P4_2/nmc$, tetragonal
3.558 5.258	a,c (Å)
3	3 non equivalent atoms
240 0.0 0.0 0.0	Zirconium, Pseudopotential (Z' > 200)
208 0.0 -0.5 0.174	Oxygen, Pseudopotential
208 0.5 0.0 0.326	Oxygen, Pseudopotential

Zirconia - ZrO_2 - cubic structure

CRYSTAL	
0 0 1	IFLAG IFHR IFSO *(=1, standard shift of origin)*
225	space group 225, $Fm3m$, cubic
5.10	a (Å)
3	3 non equivalent atoms
240 0.00 0.00 0.00	Z=40 Zirconium, Pseudopotential (Z' > 200)
208 0.25 0.25 0.25	Oxygen, Pseudopotential
208 -0.25 -0.25 -0.25	Oxygen, Pseudopotential

SiO_2, Chabazite

CRYSTAL	
0 1 0	IFLAG IFHR *(=1,rhombohedral representation)* IFSO
166	space group 166 $R\bar{3}m$, hexagonal
9.42 94.47	a (Å), $\beta°$
5	5 non equivalent atoms (36 atoms in the primitive cell)
14 .1045 .334 .8755	Silicon (multiplicity 12)
8 .262 -.262 .0	Oxygen (multiplicity 6)
8 .1580 -.1580 .5000	Oxygen (multiplicity 6)
8 .2520 .2520 .8970	Oxygen (multiplicity 6)
8 .0250 .0250 .3210	Oxygen (multiplicity 6)

SiO_2, Siliceous Faujasite

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
227	space group 227, $Fd3m$, cubic
21.53	a (Å)
5	5 non equivalent atoms (144 atoms in the primitive cell)
14 .1265 -.0536 .0370	Silicon (multiplicity 48)
8 .1059 -.1059 .0	Oxygen (multiplicity 24)
8 -.0023 -.0023 .1410	Oxygen (multiplicity 24)
8 .1746 .1746 -.0378	Oxygen (multiplicity 24)
8 .1785 .1785 .3222	Oxygen (multiplicity 24)

SiO_2, Siliceous Edingtonite

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
115	space group 115, $P\bar{4}m2$, tetragonal
6.955 6.474	a, c (Å)
5	5 non equivalent atoms (15 atoms in the primitive cell)
14 .0 .0 .5000	Silicon (multiplicity 1)
14 .0 .2697 .1200	Silicon (multiplicity 4)
8 .0 .189 .3543	Oxygen (multiplicity 4)
8 .50000 .0 .8779	Oxygen (multiplicity 2)
8 .189 .189 .0	Oxygen (multiplicity 4)

SiO_2, Siliceous Sodalite

CRYSTAL	
0 0 0	IFLAG IFHR IFSO
218	space group 218, $P\bar{4}3n$, cubic
8.950675	a (Å)
3	3 non equivalent atoms (36 atoms in the primitive cell)
14 .25000 .50000 .0	Silicon (multiplicity 6)
14 .25000 .0 .50000	Silicon (multiplicity 6)
8 .14687 .14687 .50000	Oxygen (multiplicity 24)

2D - Slabs (surfaces) - 1st input record keyword: SLAB

A 2D structure can either be created by entering directly the 2D cell parameters and irreducible atoms coordinates to obtain a slab of given thickness (keyword **SLAB** in the first record of the geometry input), or it can be derived from the 3D structure through the keyword **SLABCUT** (page 42), entered in the geometry editing section of 3D structure input. In that case the layer group is automatically identified by the program. The input tests 4-24, 5-25, 6-26 and 7-27 show the two different ways to obtain the same 2D structure.
Atom coordinates: z in Ångstrom, x, y in fractional units of the crystallographic cell translation vectors.

Test05 - graphite 2D (see test 25)

SLAB	
77	layer group (hexagonal)
2.47	lattice vector length (Å)
1	1 non equivalent atom
6 -0.33333333333 0.33333333333 0.	Z=6; Carbon; x,y,z

Beryllium - 3 layers slab

SLAB	
78	layer group (hexagonal)
2.29	lattice vector length (Å)
2	2 non equivalent atoms
4 0.333333333333 0.666666666667 0.	Z=4, Beryllium; $1/3, 2/3, z$
4 0.666666666667 0.333333333333 1.795	Z=4, Beryllium; $2/3, 1/3, z$

Test06 - beryllium - 4 layers slab (see test 26)

SLAB	
72	layer group (hexagonal)
2.29	lattice vector length (Å)
2	2 non equivalent atoms
4 0.333333333333 0.666666666667 0.897499	Z=4, Beryllium;x,y,z
4 0.666666666667 0.333333333333 2.692499	Z=4, Beryllium;x,y,z

---

Test04 - Corundum 001 (0001) 2 layers slab (see test 24)

SLAB	
66	layer group (hexagonal)
4.7602	lattice vector length (Å)
3	3 non equivalent atoms
13 0. 0. 1.9209	Z=13, Aluminum; x,y,z
8 0.333333333 -0.027093 1.0828	Z=8, Oxygen; x,y,z
13 -0.333333333 0.333333333 0.2446	Z=13, Aluminum; x,y,z

Test07 - Corundum 110 (1010) slab (see test 27)

SLAB	
7	layer group (Oblique)
5.129482 6.997933 95.8395	a,b (Å) α (degrees)
6	6 non equivalent atoms
8 -0.25 0.5 2.1124	Z=8, Oxygen; x,y,z
8 0.403120 0.153120 1.9189	Z=8, Oxygen; x,y,z
8 0.096880 0.346880 0.4612	Z=8, Oxygen; x,y,z
8 -0.25 0.00 0.2677	Z=8, Oxygen; x,y,z
13 0.454320 0.397840 1.19	Z=13, Aluminum; x,y,z
13 0.045680 0.102160 1.19	Z=13, Aluminum; x,y,z

MgO (110) 2 layers slab

SLAB	
40	layer group
4.21 2.97692	lattice vectors length (Å)
2	2 non equivalent atoms
12 0.25 0.25 0.74423	Z=12, Magnesium; x,y,z
8 0.75 0.25 0.74423	Z=8, Oxygen; x,y,z

MgO (110) 3 layers slab

SLAB	
37	
4.21 2.97692	lattice vectors length (Å)
4	4 non equivalent atoms
12 0. 0. 1.48846	Z=12, Magnesium; x,y,z
8 0.5 0. 1.48846	Z=8, Oxygen; x,y,z
12 0.5 0.5 0.	Z=12, Magnesium; x,y,z
8 0. 0.5 0.	Z=8, Oxygen; x,y,z

CO on MgO (001) two layers slab - one-side adsorption

SLAB	
55	
2.97692	lattice vector length [4.21/ $\sqrt{2}$] (Å)
6	6 non equivalent atoms
108 0. 0. 4.5625	Z=8, Oxygen; x,y,z
6 0. 0. 3.4125	Z=6, Carbon; x,y,z
12 0. 0. 1.0525	Z=12, Magnesium; x,y,z
8 0.5 0.5 1.0525	Z=8, Oxygen; x,y,z
12 0. 0. -1.0525	Z=12, Magnesium; x,y,z
8 0.5 0.5 -1.0525	Z=8, Oxygen; x,y,z

Two different conventional atomic numbers (8 and 108) are attributed to the Oxygen in CO and to the Oxygen in MgO. Two different basis sets will be associated to the two type of atoms (see Basis Set input, page 14, and test 36).

CO on MgO (001) two layers slab - two-side adsorption

SLAB	
64	
2.97692	lattice vector length (Å)
4	4 non equivalent atoms
108 0.25 0.25 4.5625	Z=8, Oxygen; x,y,z
6 0.25 0.25 3.4125	Z=6, Carbon; x,y,z
12 0.25 0.25 1.0525	Z=12, Magnesium; x,y,z
8 0.75 0.75 1.0525	Z=8, Oxygen; x,y,z

Two different conventional atomic numbers (8 and 108) are attributed to the Oxygen in CO and to the Oxygen in MgO.

Diamond slab parallel to (100) face - nine layers slab

SLAB	
59	
2.52437	lattice vector length (Å)
5	5 non equivalent atoms
6 0. 0. 0.	Z=6, Carbon; x,y,z
6 0.5 0. 0.8925	Z=6, Carbon; x,y,z
6 0.5 0.5 1.785	Z=6, Carbon; x,y,z
6 0. 0.5 2.6775	Z=6, Carbon; x,y,z
6 0. 0. 3.57	Z=6, Carbon; x,y,z

Diamond slab parallel to (100) face - ten layers slab

SLAB	
39	layer group
2.52437 2.52437	lattice vectors length (Å)
5	5 non equivalent atoms
6 0.25 0. 0.44625	Z=6, Carbon; x,y,z
6 0.25 0.5 1.33875	Z=6, Carbon; x,y,z
6 0.75 0.5 2.23125	Z=6, Carbon; x,y,z
6 0.75 0 3.12375	Z=6, Carbon; x,y,z
6 0.25 0. 4.01625	Z=6, Carbon; x,y,z

1D - Polymers - 1st input record keyword: POLYMER

Atom coordinates: y,z in Ångstrom, x in fractional units of the crystallographic cell translation vector.

Test03 - (SN)$_x$ polymer

POLYMER	
4	rod group
4.431	lattice vector length (Å)
2	2 non equivalent atoms
16 0.0 -0.844969 0.0	Z=16, Sulphur; x, y, z
7 0.141600540 0.667077 -0.00093	Z=7, Nitrogen; x, y, z

Water polymer

POLYMER	
1	
4.965635	lattice vector length (Å)
6	6 non equivalent atoms
8 0. 0. 0.	Z=8, Oxygen; x, y, z
1 0.032558 0.836088 -0.400375	Z=1, Hydrogen; x, y, z
1 0.168195 -0.461051 0.	Z=1, Hydrogen; x, y, z
8 0.5 -1.370589 0.	Z=8, Oxygen; x, y, z
1 0.532558 -2.206677 0.400375	Z=1, Hydrogen; x, y, z
1 0.668195 -0.909537 0.	Z=1, Hydrogen; x, y, z

Formamide chain - test40 DFT

POLYMER	
4	rod group
8.774	lattice vector length (Å)
6	6 non equivalent atoms
8 -7.548E-2 5.302E-3 0.7665	Z=8, Oxygen; x, y, z
7 0.1590 -0.8838 0.3073	Z=7, Nitrogen; x, y, z
6 5.627E-2 7.051E-2 0.2558	Z=6, Oxygen; x, y, z
1 0.2677 -0.6952 -9.1548E-2	Z=1, Hydrogen; x, y, z
1 0.1310 -1.8019 0.7544	Z=1, Hydrogen; x, y, z
1 9.244E-2 0.9973 -0.2795	Z=1, Hydrogen; x, y, z

0D - Molecules - 1st input record keyword: MOLECULE

Atom coordinates: x,y,z in Ångstrom.

Test00 - CO molecule

MOLECULE	
1	point group
2	2 non equivalent atoms
6 0. 0. 0.	Z=6, Carbon; x, y, z
8 0.8 0.5 0.4	Z=8, Oxygen; x, y, z

Test01 - CH_4 Methane molecule

MOLECULE	
44	point group
2	2 non equivalent atoms
6 0. 0. 0.	Z=6, Carbon; x, y, z
1 0.629 0.629 0.629	Z=1, Hydrogen; x, y, z

Test02 - $CO(NH_2)_2$ Urea molecule

MOLECULE	
15	point group
5	5 non equivalent atoms
6 0. 0. 0.	Z=6, Carbon; x, y, z
8 0. 0. 1.261401	Z=8, Oxygen; x, y, z
7 0. 1.14824666034 -0.69979	Z=7, Nitrogen; x, y, z
1 0. 2.0265496501 -0.202817	Z=1, Hydrogen; x, y, z
1 0. 1.13408048308 -1.704975	Z=1, Hydrogen; x, y, z

6.2 Basis set input

Optimized basis sets for periodic systems used in published papers are available on WWW:

http://www.crystal.unito.it

All electron Basis sets for Silicon atom

STO-3G

14 3	Z=14, Silicon; 3 shells
1 0 3 2. 0.	Pople BS; s shell; 3G; CHE=2; standard scale factor
1 1 3 8. 0.	Pople BS; sp shell; 3G; CHE=8; standard scale factor
1 1 3 4. 0.	Pople BS; sp shell; 3G; CHE=4; standard scale factor

---

6-21G

14 4	Z=14, Silicon; 4 shells
2 0 6 2. 1.	Pople 6-21 BS; s shell; 6G; CHE=2; scale factor 1 (core AO).
2 1 6 8. 1.	Pople 6-21 BS; sp shell; 6G; CHE=8; scale factor 1 (core AOs).
2 1 2 4. 1.	Pople 6-21 BS; sp shell; 2G; CHE=4; scale factor 1 (inner valence).
2 1 1 0. 1.	Pople 6-21 BS; sp shell; 1G; CHE=0; scale factor 1 (outer valence).

NB. The 4th shell has electron charge 0. The basis functions of that shell are included in the basis set to compute the atomic wave functions, as they correspond to symmetries (angular quantum numbers) occupied in the ground state of the atom. The atomic basis set is: 4s, 3p.

6-21G modified

14 4	Z=14, Silicon; 4 shells
2 0 6 2. 1.	Pople 6-21 BS; s shell; 6G; CHE=2; scale factor 1.
2 1 6 8. 1.	Pople 6-21 BS; sp shell; 6G; CHE=8; scale factor 1.
2 1 2 4. 1.	Pople 6-21 BS; sp shell; 2G; CHE=4; scale factor 1.
0 1 1 0. 1.	free BS; sp shell; 1G; CHE=0; scale factor 1.
0.16 1. 1.	gaussian exponent; s coefficient; p coefficient

3-21G

14 4	Z=14, Silicon; 4 shells
2 0 3 2. 1.	Pople 3-21 BS; s shell; 3G; CHE=2; scale factor 1.
2 1 3 8. 1.	Pople 3-21 BS; sp shell; 3G; CHE=8; scale factor 1.
2 1 2 4. 1.	Pople 3-21 BS; sp shell; 2G; CHE=4; scale factor 1.
2 1 1 0. 1.	Pople 3-21 BS; sp shell; 1G; CHE=0; scale factor 1.

3-21G*

14 5	Z=14, Silicon; 5 shells
2 0 3 2. 1.	Pople 3-21 BS; s shell; 3G; CHE=2; scale factor 1.
2 1 3 8. 1.	Pople 3-21 BS; sp shell; 3G; CHE=8; scale factor 1.
2 1 2 4. 1.	Pople 3-21 BS; sp shell; 2G; CHE=4; scale factor 1.
2 1 1 0. 1.	Pople 3-21 BS; sp shell; 1G; CHE=0; scale factor 1.
2 3 1 0. 1.	Pople 3-21 BS; d shell; 1G; CHE=0; scale factor 1.

NB. The basis functions of the 5th shell, d symmetry, unoccupied in the ground state of Silicon atom, is not included in the atomic wave function calculation.

3-21G modified+polarization

14 5	Z=14, Silicon; 5 shells
2 0 3 2. 1.	Pople 3-21 BS; s shell; 3G; CHE=2; scale factor 1.
2 1 3 8. 1.	Pople 3-21 BS; sp shell; 3G; CHE=8; scale factor 1.
2 1 2 4. 1.	Pople 3-21 BS; sp shell; 2G; CHE=4; scale factor 1.
0 1 1 0. 1.	free BS; sp shell; 1G; CHE=0; scale factor 1.
0.16 1. 1.	gaussian exponent; s contraction coefficient; p contr. coeff.
0 3 1 0. 1.	free BS; d shell; 1G; CHE=0; scale factor 1.
0.5 1.	gaussian exponent; d contraction coefficient.

free basis set

14 4			Z=14, Silicon; 4 shells
0 0 6 2. 1.			free BS; s shell; 6 GTF; CHE=2; scale factor 1.
16115.9	0.00195948		1st gaussian exponent; s contraction coefficient
2425.58	0.0149288		2nd gaussian exponent; s contraction coefficient
553.867	0.0728478		3rd gaussian exponent; s contraction coefficient
156.340	0.24613		4th gaussian exponent; s contraction coefficient
50.0683	0.485914		5th gaussian exponent; s contraction coefficient
17.0178	0.325002		6th gaussian exponent; s contraction coefficient
0 1 6 8. 1.			free BS; sp shell; 6 GTF; CHE=8; scale factor 1.
292.718	-0.00278094	0.00443826	1st gaussian exp.; s contr. coeff.; p contr. coeff.
69.8731	-0.0357146	0.0326679	2nd gaussian exp.; s contr. coeff.; p contr. coeff.
22.3363	-0.114985	0.134721	3rd gaussian exp.; s contr. coeff.; p contr. coeff.
8.15039	0.0935634	0.328678	4th gaussian exp.; s contr. coeff.; p contr. coeff.
3.13458	0.603017	0.449640	5th gaussian exp.; s contr. coeff.; p contr. coeff.
1.22543	0.418959	0.261372	6th gaussian exp.; s contr. coeff.; p contr. coeff.
0 1 2 4. 1.			free BS; sp shell; 2 GTF; CHE=4; scale factor 1
1.07913	-0.376108	0.0671030	1st gaussian exp.; s contr. coeff.; p contr. coeff.
0.302422	1.25165	0.956883	2nd gaussian exp.; s contr. coeff.; p contr. coeff.
0 1 1 0. 1.			free BS; sp shell; 1 GTF; CHE=0; scale factor 1.
0.123	1.	1.	gaussian exp.; s contr. coeff.; p contr. coeff.

Examples of ECP and valence only basis set input

Nickel atom. Electronic configuration: [Ar] 4s(2) 3d(8)

Durand & Barthelat large core

228 4			Z=28,Nickel; 4 shells valence basis set
BARTHE			keyword; Durand-Barthelat ECP
0 1 2 2. 1.			free BS;sp shell;2 GTF;CHE=2;scale factor 1
1.55	.24985	1.	1st GTF exponent;s coefficient;p coefficient
1.24	-.41636	1.	2nd GTF exponent;s coefficient;p coefficient
0 1 1 0. 1.			free BS; sp shell; 1 GTF; CHE=0; scale factor 1
0.0818	1.0	1.	GTF exponent;s coefficient;p coefficient
0 3 4 8. 1.			free BS; d shell; 4 GTF; CHE=8; scale factor 1
4.3842E+01	.03337		1st GTF exponent; d coefficient
1.2069E+01	.17443		2nd GTF exponent; d coefficient
3.9173E+00	.42273		3rd GTF exponent; d coefficient
1.1997E+00	.48809		4th GTF exponent; d coefficient
0 3 1 0. 1.			free BS; d shell; 1 GTF; CHE=0; scale factor 1
0.333	1.		GTF exponent; d coefficient

Hay & Wadt Large Core - [Ar] 4s(2) 3d(8)

228 4			Z=28,Nickel; 4 shells valence basis set
HAYWLC			keyword; Hay-Wadt large core ECP
0 1 2 2. 1.			free BS; sp shell; 2 GTF; CHE=2; scale factor 1
1.257	1.1300E-01	2.6760E-02	exponent,s coefficient,p coefficient
1.052	-1.7420E-01	-1.9610E-02	
0 1 1 0. 1.			second shell,sp type,1 GTF
0.0790	1.0	1.	
0 3 4 8. 1.			third shell,d type,4 primitive GTF
4.3580E+01	.03204		
1.1997E+01	.17577		
3.8938E+00	.41461		
1.271	.46122		
0 3 1 0. 1.			fourth shell,d type,1 GTF
0.385	1.		

Hay & Wadt Small Core - [Ne] 3s(2) 3p(6) 4s(2) 3d(8)

228 6			nickel basis set - 6 shells
HAYWSC			keyword; Hay-Wadt small core ECP
0 1 3 8. 1.			first shell,sp type,3 primitive GTF -
2.5240E+01	-3.7000E-03	-4.0440E-02	exponent,s coefficient,p coefficient
7.2019E+00	-5.3681E-01	-7.6560E-02	
3.7803E+00	4.2965E-01	4.8348E-01	
0 1 2 2. 1.			second shell,sp type,2 primitive GTF
1.40	.84111	.55922	
0.504	.13936	.12528	
0 1 1 0. 1.			third shell,sp type,1 GTF
0.0803	1.0	1.	
0 3 3 8. 1.			fourth shell,d type,4 primitive GTF
4.1703E+01	3.5300E-02		
1.1481E+01	1.8419E-01		
3.7262E+00	4.1696E-01		
0 3 1 0. 1.			fifth shell,d type,1 GTF
1.212	1.		
0 3 1 0. 1.			sixth shell,d type,1 GTF
0.365	1.0		

Free input

228 5			Z=28, nickel basis set - 5 shells (valence only)
INPUT			keyword: free ECP (Large Core)- input follows
10.	5 4 5 2 0		nuclear charge; number of terms in eq. 2.8 and 2.9
344.84100	-18.00000	-1	eq. 2.8, 5 records:
64.82281	-117.95937	0	α, C, n
14.28477	-29.43970	0	
3.82101	-10.38626	0	
1.16976	-0.89249	0	
18.64238	3.00000	-2	eq. 2.9, 4 records $\ell = 0$
4.89161	19.24490	-1	
1.16606	23.93060	0	
0.95239	-9.35414	0	
30.60070	5.00000	-2	eq. 2.9, 5 records $\ell = 1$
14.30081	19.81155	-1	
15.03304	54.33856	0	
4.64601	54.08782	0	
0.98106	7.31027	0	
4.56008	0.26292	0	eq. 2.9, 2 records $\ell = 2$
0.67647	-0.43862	0	basis set input follows - valence only
0 1 1 2. 1.			1st shell: sp type; 1 GTF; CHE=2; scale fact.=1
1.257	1.	1.	exponent, s coefficient, p coefficient
0 1 1 0. 1.			2nd shell: sp type; 1 GTF; CHE=0; scale fact.=1
1.052	1.	1.	
0 1 1 0. 1.			3rd shell: sp type; 1 GTF; CHE=0; scale fact.=1
0.0790	1.0	1.	
0 3 4 8. 1.			4th shell; d type; 4 GTF; CHE=8; scale fact.=1
4.3580E+01	.03204		
1.1997E+01	.17577		
3.8938E+00	.41461		
1.271	.46122		
0 3 1 0. 1.			5th shell; d type; 1 GTF; CHE=0; scale fact.=1
0.385	1.		

6.3 SCF options - SPINEDIT

Example of how to edit the density matrix obtained for a given magnetic solution to define a scf guess with a different magnetic solution.

Deck 1 - ferromagnetic solution

Spinel MnCr2O4

CRYSTAL	
0 0 0	
227	space group number
8.5985	lattice parameter
3	3 non equivalent atoms (14 atoms in the primitive cell)
24 0.500 0.500 0.500	Chromium - x, y, z - multiplicity 4
25 0.125 0.125 0.125	Manganese - x, y, z - multiplicity 2
8 0.2656 0.2656 0.2656	Oxygen - x, y, z - multiplicity 8
END	end of geometry input records - block 1
basis set input terminated by **END**	
UHF	Unrestricted Hartree Fock
TOLINTEG	the default value of the truncation tolerances is modified
7 7 7 7 14	new values for ITOL1-ITOl2-ITOL3-ITOL4-ITOL5
END	end of input block 3
4 0 4	reciprocal lattice sampling (page 18)
SPINLOCK	n_α - n_β is locked to be 22 for 50 cycles.
22 50	All the d electrons are forced to be parallel
LEVSHIFT	a level shifter of 0.3 hartree, maintained after diagonalization,
3 1	causes a lock in a non-conducting solution
MAXCYCLE	the maximum number of SCF cycles is set to 50
50	
PPAN	Mulliken population analysis at the end of SCF cycles
END	

Deck 2 (SCF input only)

4 0 4	
GUESSP	initial guess: density matrix from a previous run
SPINEDIT	elements of the density matrix are modified
2	the diagonal elements corresponding to 2 atoms
5 6	*label* of the 2 atoms (6 is equivalent to 5)
LEVSHIFT	a level shifter of 0.3 hartree, maintained after diagonalization,
3 1	causes a lock in a non-conducting solution
PPAN	Mulliken population analysis at the end of SCF cycles
END	

```
======================================================================
First run - geometry output
======================================================================

 COORDINATES OF THE EQUIVALENT ATOMS (FRACTIONARY UNITS)

   N. ATOM  EQUIVALENT  AT. NUMBER      X          Y          Z

   1   1        1         24 CR     -5.000E-01 -5.000E-01 -5.000E-01
   2   1        2         24 CR     -5.000E-01 -5.000E-01  0.000E+00
   3   1        3         24 CR      0.000E+00 -5.000E-01 -5.000E-01
   4   1        4         24 CR     -5.000E-01  0.000E+00 -5.000E-01

   5   2        1         25 MN      1.250E-01  1.250E-01  1.250E-01
   6   2        2         25 MN     -1.250E-01 -1.250E-01 -1.250E-01

   7   3        1          8 O       2.656E-01  2.656E-01  2.656E-01
   8   3        2          8 O       2.656E-01  2.656E-01 -2.968E-01
   9   3        3          8 O      -2.968E-01  2.656E-01  2.656E-01
  10   3        4          8 O       2.656E-01 -2.968E-01  2.656E-01
  11   3        5          8 O      -2.656E-01 -2.656E-01 -2.656E-01
  12   3        6          8 O      -2.656E-01 -2.656E-01  2.968E-01
  13   3        7          8 O      -2.656E-01  2.968E-01 -2.656E-01
  14   3        8          8 O       2.968E-01 -2.656E-01 -2.656E-01

======================================================================
Ferromagnetic solution: all unpaired electrons with the same spin
======================================================================
 SPIN POLARIZATION - ALPHA-BETA = 22 FOR  50 CYCLES
======================================================================
Convergence on total energy reached in 33 cycles (level shifter active)
======================================================================
 CYCLE  33 ETOT(AU) -7.072805900367E+03 DETOT -8.168E-07 DE(K)  9.487E+00

======================================================================
Population analysis  -  ferromagnetic solution
======================================================================
 MULLIKEN POPULATION ANALYSIS
 ALPHA+BETA ELECTRONS - NO. OF ELECTRONS  210.000000

  ATOM    Z CHARGE  SHELL POPULATION
                      s     sp     sp     sp     sp      d      d
   1 CR  24 21.884  2.000  8.047  2.251  4.487  1.331  3.078   .690
   5 MN  25 23.147  2.000  8.081  2.170  4.299  1.489  4.478   .629
   7 O    8  9.521  1.996  2.644  2.467  2.414

 MULLIKEN POPULATION ANALYSIS
 ALPHA-BETA ELECTRONS - NO. OF ELECTRONS   22.000000

  ATOM    Z CHARGE  SHELL POPULATION
                      s     sp     sp     sp     sp      d      d
   1 CR  24  3.057   .000  -.002   .011   .027  -.011  2.790   .242
   5 MN  25  4.925   .000  -.003   .019   .055  -.052  4.408   .498
   7 O    8  -.010   .000   .003  -.014   .002
===============================================================================
===============================================================================
Second  run  - Anti ferromagnetic solution:
Integrals calculation not affected by the spin state

Cr (atoms 1-2-3-4) unpaired electrons spin alpha;
Mn (atoms 5 and 6) unpaired electrons spin beta
===============================================================================

RESTART FROM A PREVIOUS RUN DENSITY MATRIX
SPIN INVERSION IN SPIN DENSITY MATRIX FOR ATOMS:    5    6

======================================================================
Convergence on total energy reached in 15 cycles
```

```
======================================================================
 CYCLE  15 ETOT(AU) -7.072808080821E+03 DETOT -4.930E-07 DE(K)  6.694E-06

=====================================uuuu============================
Population analysis  -  anti ferromagnetic solution
======================================================================
 MULLIKEN POPULATION ANALYSIS
ALPHA+BETA ELECTRONS - NO. OF ELECTRONS  210.000000

 ATOM    Z CHARGE  SHELL POPULATION
                     s     sp     sp     sp     sp      d      d
   1 CR  24 21.884  2.000  8.047  2.251  4.487  1.331  3.078   .690
   5 MN  25 23.149  2.000  8.081  2.170  4.299  1.489  4.479   .631
   7 O    8  9.521  1.997  2.644  2.467  2.414

 MULLIKEN POPULATION ANALYSIS
 ALPHA-BETA ELECTRONS - NO. OF ELECTRONS    2.000000
  ATOM    Z CHARGE  SHELL POPULATION
                     s     sp     sp     sp     sp      d      d
   1 CR  24  3.049   .000  -.002   .011   .027  -.012  2.785   .240
   5 MN  25 -4.917   .000   .003  -.018  -.055   .054 -4.406  -.495
   7 O    8  -.045   .000  -.024  -.013  -.008
=================================================================================
```

6.4 Geometry optimization - OPTGEOM

Crystal geometry input section (block1) for the geometry optimization of the urea molecule:

▷ Example

`Urea Molecule`	Title
`MOLECULE`	Dimension of the system
`15`	Point group (C_{2v})
`5`	Number of non equivalent atoms
`6 0.  0.  0.`	Atomic number and cartesian coordinates
`8 0.  0.  1.261401`	
`7 0.  1.148247 -0.699790`	
`1 0.  2.026550 -0.202817`	
`1 0.  1.134080 -1.704975`	
`OPTGEOM`	Keyword to perform a geometry optimization
`ENDOPT`	End of geometry optimization input block
`END`	end og geometry input
`Basis set input`	As in test 12
`END`	End of basis set input section
`END`	block3 input - Molecule - no information on sampling in **K** space

Crystal output contains additional information on the optimization run after the initial part of the geometry output:

```
 . . . . . . . . .. . . . . . . . . . . . . . . . .

BERNY OPTIMIZATION CONTROL

 MAXIMUM GRADIENT COMPONENT   0.00045 MAXIMUM DISPLACEMENT COMPONENT     0.00030
 R.M.S. OF GRADIENT COMPONENT 0.00180 R.M.S. OF DISPLACEMENT COMPONENTS  0.00120
 THRESHOLD ON ENERGY CHANGE 0.100E-06 EXTRAPOLATING POLYNOMIAL ORDER         2
 MAXIMUM ALLOWED NUMBER OF STEPS  100 SORTING OF ENERGY POINTS:             NO
 ANALYTICAL GRADIENTS
 . . . . . . . . .. . . . . . . . . . . . . . . . .

 SYMMETRY ALLOWED INTERNAL DEGREE(S) OF FREEDOM:   7

 . . . . . . . . .. . . . . . . . . . . . . . . . .
```

At the first step of the optimization, the **Crystal** standard output contains both energy (complete SCF cycle) and gradient parts. At the end of the first step, a convergence check is performed on the initial forces and the optimization stops if the criteria are already satisfied. For the subsequent steps, only few lines on the optimization process are reported in standard output, namely: current geometry, total energy and gradients, and convergence tests (SCF output is routed to file SCFOUT.LOG).
At each optimization step, xxx, the geometry is written in file optcxxx (in a format suitable to be read with EXTERNAL keyword). Optimization step can be restarted from any step geometry, by renaming optcxxx as fort.34.
The standard output for the urea molecule geometry optimization looks as follows:

```
OPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPT
*******************************************************************************
GEOMETRY OPTIMIZATION - POINT    2
*******************************************************************************
ATOMS IN THE ASYMMETRIC UNIT    5 - ATOMS IN THE UNIT CELL:    8
    ATOM          X(ANGSTROM)         Y(ANGSTROM)         Z(ANGSTROM)
*******************************************************************************
  1 T   6 C     0.000000000000E+00  0.000000000000E+00  2.645266012706E-02
  2 T   8 O     0.000000000000E+00  0.000000000000E+00  1.241474126876E+00
  3 T   7 N     0.000000000000E+00  1.150483100972E+00 -7.044307566681E-01
  4 F   7 N     0.000000000000E+00 -1.150483100972E+00 -7.044307566681E-01
  5 T   1 H     0.000000000000E+00  2.022583078191E+00 -2.043778206895E-01
  6 F   1 H     0.000000000000E+00 -2.022583078191E+00 -2.043778206895E-01
  7 T   1 H     0.000000000000E+00  1.135517317174E+00 -1.702036316144E+00
  8 F   1 H     0.000000000000E+00 -1.135517317174E+00 -1.702036316144E+00

T = ATOM BELONGING TO THE ASYMMETRIC UNIT

INTRACELL NUCLEAR REPULSION (A.U.)  1.2463005288098E+02

TOTAL ENERGY(HF)(AU)(  11) -2.2379435865343E+02 DE-4.8E-08 DP 7.2E-06

SYMMETRY ALLOWED FORCES (ANALYTICAL) (DIRECTION, FORCE)

   1  4.0854048E-02    2 -2.8460660E-02    3  1.4184257E-03    4 -3.0361419E-03
   5 -1.7599295E-02    6 -1.3809310E-02    7  6.7962224E-03

GRADIENT NORM     0.055108  GRADIENT THRESHOLD     0.500000

MAX GRADIENT      0.040854  THRESHOLD              0.000450 CONVERGED NO
RMS GRADIENT      0.020829  THRESHOLD              0.000300 CONVERGED NO
MAX DISPLAC.      0.024990  THRESHOLD              0.001800 CONVERGED NO
RMS DISPLAC.      0.015649  THRESHOLD              0.001200 CONVERGED NO
```

When all four convergence tests are satisfied, optimization is completed. The final energy and the optimized structure are printed after the final convergence tests.

```
******************************************************************
* OPT END - CONVERGED * E(AU):  -2.237958289701E+02  POINTS   14 *
******************************************************************
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .. . .

FINAL OPTIMIZED GEOMETRY - DIMENSIONALITY OF THE SYSTEM      0
(NON PERIODIC DIRECTION: LATTICE PARAMETER FORMALLY SET TO 500)
*******************************************************************************
ATOMS IN THE ASYMMETRIC UNIT    5 - ATOMS IN THE UNIT CELL:    8
    ATOM          X(ANGSTROM)         Y(ANGSTROM)         Z(ANGSTROM)
*******************************************************************************
  1 T   6 C     0.000000000000E+00  0.000000000000E+00  3.468988750953E-02
  2 T   8 O     0.000000000000E+00  0.000000000000E+00  1.230143233209E+00
  3 T   7 N     0.000000000000E+00  1.143750090534E+00 -7.056136525307E-01
  4 F   7 N     0.000000000000E+00 -1.143750090534E+00 -7.056136525307E-01
  5 T   1 H     0.000000000000E+00  2.001317638364E+00 -2.076003454226E-01
  6 F   1 H     0.000000000000E+00 -2.001317638364E+00 -2.076003454226E-01
  7 T   1 H     0.000000000000E+00  1.157946292824E+00 -1.696084062406E+00
```

```
  8 F   1 H     0.000000000000E+00 -1.157946292824E+00 -1.696084062406E+00

 T = ATOM BELONGING TO THE ASYMMETRIC UNIT

 INTRACELL NUCLEAR REPULSION (A.U.)  1.2541002823701E+02

 ****   4 SYMMOPS - TRANSLATORS IN FRACTIONARY UNITS
  V INV                  ROTATION MATRICES                       TRANSLATOR
  1  1  1.00  0.00  0.00  0.00  1.00  0.00  0.00  0.00  1.00    0.00  0.00  0.00
  2  2 -1.00  0.00  0.00  0.00 -1.00  0.00  0.00  0.00  1.00    0.00  0.00  0.00
  3  3 -1.00  0.00  0.00  0.00  1.00  0.00  0.00  0.00  1.00    0.00  0.00  0.00
  4  4  1.00  0.00  0.00  0.00 -1.00  0.00  0.00  0.00  1.00    0.00  0.00  0.00
```

The final geometry is both printed in the **Crystal** output and written in file fort.34. The following input defines the geometry, reading data from file fort.34 (keyword **EXTERNAL**, input block1, page 11).

▷ Example

`Urea Molecule`	Title
`EXTERNAL`	Geometry read from file fort.34
optional keywords	
`END`	End of the geometry input section

Optimization can restart, by adding the keyword **RESTART** in the same input deck used for the first optimization run. Information to restart are read from file OPTINFO.DAT, updated after each optimization cycle.

▷ Example

Initial input	**Restart input**
`Urea Molecule`	`Urea Molecule`
`MOLECULE`	`MOLECULE`
`15`	`15`
`5`	`5`
`6 0.  0.  0.`	`6 0.  0.  0.`
`8 0.  0.  1.261401`	`8 0.  0.  1.261401`
`7 0.  1.148247 -0.699790`	`7 0.  1.148247 -0.699790`
`1 0.  2.026550 -0.202817`	`1 0.  2.026550 -0.202817`
`1 0.  1.134080 -1.704975`	`1 0.  1.134080 -1.704975`
`OPTGEOM`	`OPTGEOM`
`ENDOPT`	`RESTART`
`END`	`ENDOPT`
	`END`

Partial optimization

In order to optimize the coordinates of the hydrogens in urea molecule, the following input must be entered:

▷ Example

`Urea Molecule`	Title
`MOLECULE`	Dimension of the system
`15`	Point group (C_{2v})
`5`	Number of non equivalent atoms
`6 0.  0.  0.`	Atomic number and cartesian coordinates
`8 0.  0.  1.261401`	
`7 0.  1.148247 -0.699790`	
`1 0.  2.026550 -0.202817`	
`1 0.  1.134080 -1.704975`	
`KEEPSYMM`	maintain symmetry in subsequent operations
`OPTGEOM`	Keyword to perform a geometry optimization
`ATOMFREE`	Keyword for a partial optimization
`2`	Number of atoms to be optimized
`5 7`	Label of the atoms to be optimized
`ENDOPT`	End of the geometry optimization input block
`END`	End of the geometry input section

The atoms allowed to move are the two hydrogens irreducible, 5 and 7. The symmetry is maintained, atoms 6 and 7 are allowed to move.

```
*******************************************************************************
* PARTIAL OPTIMIZATION - ATOMS FREE TO MOVE    2  INPUT LIST :

  5(  1)   7(  1)

SYMMETRY IS KEPT

ATOM   5 AND   6 ARE LINKED BY SYMMOP  2
ATOM   7 AND   8 ARE LINKED BY SYMMOP  2
*******************************************************************************
```

If the symmetry is not maintained (default, no KEEPSYMM before OPTGEOM) the symmetry operators linking atoms 5 and 6, and atoms 7 and 8, are removed.

```
*******************************************************************************
* PARTIAL OPTIMIZATION - ATOMS FREE TO MOVE    2  INPUT LIST :

  5(  1)   7(  1)

SYMMETRY MAY BE BROKEN

THE NUMBER OF SYMMETRY OPERATORS HAS BEEN REDUCED FROM  4 TO  2
******************************************************************************
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .. . .
SYMMETRY ALLOWED INTERNAL DEGREE(S) OF FREEDOM:   4
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .. . .
```

Final run

During optimization process, the classification of the integrals is done with reference to the input geometry, and used for all subsequent wave function calculations.
In some cases, when the optimized geometry is far from the original one, the series truncation defined with reference to the starting geometry may be inhomogeneous if applied to the final geometry (see keyword **FIXINDEX** for explanation). In those cases, the total energy computed for one geometry, with integrals selected according to a different geometry, may be not correct.
A single point calculation, with the final optimized geometry, allows to check if that is the case.

A safe procedure to check if that geometry corresponds to a real energy minimum, is to perform a second optimization process, with same truncation criteria, starting from the geometry obtained in the first optimization (read from file fort.34, keyword **EXTERNAL**, page 11). The keyword **FINALRUN** starts the process automatically. (it does not work with Pcrystal)

A typical example is the geometry optimization of a surface, described with a slab model. The optimization process may lead to a structure significantly different from the one cut from the bulk, when there is surface relaxation. As an example, the geometry optimization of the surface (001) of the α -Al_2O_3 is reported.

▷ Example - Optimization of surface

α -Al_2O_3 - (001) surface	title
CRYSTAL	dimension of the system
0 0 0	
167	space group
4.7602 12.9933	lattice parameters
2	number of irreducible atoms
13 0. 0. 0.35216	fractional coordinates of first atom
8 0.30624 0. 0.25	fractional coordinates of second atom
SLABCUT	3D→ 2D
0 0 1	(h, k, l) Miller indices of the surface
1 6	number of layers, starting from the first classified
OPTGEOM	Keyword to perform a geometry optimization
FINALRUN	keyword to check gradients vs true series truncation
3	new optimization if convergence criteria are not satisfied
ENDOPT	end of the geometry optimization input block
END	end of the geometry input section

Neighbors analysis on the initial geometry obtained with SLABCUT

```
N = NUMBER OF NEIGHBORS AT DISTANCE R   11 cycles
   ATOM  N     R/ANG      R/AU   NEIGHBORS (ATOM LABELS AND CELL INDICES)
  1 AL   3     1.8551     3.5057    2 O    0 0 0   3 O    0 0 0   4 O    0 1 0
  1 AL   3     3.2192     6.0834    5 AL   0 0 0   5 AL   1 1 0   5 AL   0 1 0
  1 AL   3     3.2219     6.0885    2 O   -1 0 0   3 O    1 1 0   4 O    0 0 0
  1 AL   3     3.4295     6.4808    7 O    0 1 0   8 O    0 1 0   9 O    0 0 0
  1 AL   3     3.4990     6.6121    6 AL   0 0 0   6 AL  -1 0 0   6 AL   0 1 0
  1 AL   1     3.8419     7.2601   10 AL   0 0 0

Total energy   E = -1399.7999027 hartree
```

Series truncation is defined with reference to that geometry. Optimization begins. After 11 cycles convergence on gradients and displacements is satisfied.
Neighbors analysis on the optimized geometry:

```
N = NUMBER OF NEIGHBORS AT DISTANCE Ra     6 cycles
   ATOM  N     R/ANG      R/AU   NEIGHBORS (ATOM LABELS AND CELL INDICES)
  1 AL   3     1.6886     3.1911    2 O    0 0 0   3 O    0 0 0   4 O    0 1 0
  1 AL   1     2.6116     4.9351   10 AL   0 0 0
  1 AL   3     2.8198     5.3286    7 O    0 1 0   8 O    0 1 0   9 O    0 0 0
  1 AL   3     3.0425     5.7494    5 AL   0 0 0   5 AL   1 1 0   5 AL   0 1 0
```

```
  1 AL   3     3.0430     5.7504   6 AL   0 0 0   6 AL  -1 0 0   6 AL   0 1 0
  1 AL   3     3.1214     5.8987   2 O   -1 0 0   3 O    1 1 0   4 O    0 0 0

 Total energy   E = -1400.1148194 hartree
```

A large geometrical relaxation occurred during the optimization: the aluminium atoms move toward the core of the slab. In this case both the total energy and gradients should be recalculate using truncation series which refer to the final relaxed geometry.
This crucial step is automatically performed if the keyword FINALRUN is present in the input file. If this is the case, CRYSTAL03 checks for the numerical consistency and it will find that the gradients do not match the requested convergence criteria. At the end of the new optimization the total energy is:

```
Total energy E = -1400.1193593 hartree
```

In this case, FINALRUN was followed by the keyword ICODE=3. This means that the geometry optimization restarts from the very last step of the previous geometry optimization with truncation series criteria defined relative to that geometry. After 6 new optimization cycles, convergence criteria are satisfied.
Neighbors analysis on the final run optimized geometry

```
 N = NUMBER OF NEIGHBORS AT DISTANCE R
    ATOM  N     R/ANG      R/AU   NEIGHBORS (ATOM LABELS AND CELL INDICES)
  1 AL   3     1.6863     3.1867   2 O    0 0 0   3 O    0 0 0   4 O    0 1 0
  1 AL   1     2.5917     4.8976  10 AL   0 0 0
  1 AL   3     2.8095     5.3092   7 O    0 1 0   8 O    0 1 0   9 O    0 0 0
  1 AL   3     3.0382     5.7414   5 AL   0 0 0   5 AL   1 1 0   5 AL   0 1 0
  1 AL   3     3.0387     5.7424   6 AL   0 0 0   6 AL  -1 0 0   6 AL   0 1 0
  1 AL   3     3.1215     5.8987   2 O   -1 0 0   3 O    1 1 0   4 O    0 0 0

 Total energy   E = -1400.1194545 hartree
```

The final geometry is printed, and written in file fort.34.
A final check on total energy can be done with the following input:

```
alpha-Al2O3 (corundum) 001 2 LAYERS (3D-->2D)
EXTERNAL
OPTGEOM
ENDOPT
END
```

The keyword EXTERNAL routes the basic geometry input stream to file fort.34, written at the end of the optimization run.
No optimization starts, convergence criteria are already satisfied.

```
 Total energy   E = -1400.1194544 hartree
```

/sectionScanning of vibrational modes - SCANMODE

Methane molecule

```
MOLECULE
44
2
6      0.000000000000E+00  0.000000000000E+00  0.000000000000E+00
1      6.252140000000E-01  6.252140000000E-01  6.252140000000E-01
FREQCALC
SCANMODE
-1 -1 0 0.1
12
ENDfreq
END
6 3
0 0 3 2. 1.
172.256000         0.617669000E-01
25.9109000         0.358794000
5.53335000         0.700713000
0 1 2 4. 1.
3.66498000        -0.395897000         0.236460000
0.770545000          1.21584000         0.860619000
0 1 1 0. 1.
0.195857000          1.00000000          1.00000000
1 2
0 0 2 1. 1.
5.44717800         0.156285000
0.824547000         0.904691000
0 0 1 0. 1.
0.183192000          1.00000000
99 0
END
TOLINTEG
20 20 20 20 20
END
TOLDEE
11
END
```

$PbCO_3$

```
PBCO3 - frequency calculation
CRYSTAL
1 0 0
P m c n
5.20471446     8.45344758     6.16074145
4
282    2.500000000000E-01  4.175726169487E-01 -2.463557995068E-01
6      2.500000000000E-01 -2.363341497085E-01 -8.558132726424E-02
8      2.500000000000E-01 -8.360585350428E-02 -9.431628799197E-02
8      4.648370653436E-01 -3.129222129903E-01 -8.842925698155E-02
FREQCALC
RESTART
SCANMODE
1 -40 40 0.1
1
END
```

```
END
282 4
HAYWLC
0 1 2 4. 1.
  1.335104   -0.1448789   -0.1070612
  0.7516086  1.0          1.0
0 1 1 0. 1.
  0.5536686  1.0          1.0
0 1 1 0. 1.
  0.1420315  1.0          1.0
0 3 1 0. 1.
  0.1933887  1.0
6 4
0 0 6 2.0 1.0
 3048.0    0.001826
  456.4    0.01406
  103.7    0.06876
   29.23   0.2304
    9.349  0.4685
    3.189  0.3628
0 1 2 4.0 1.0
    3.665 -0.3959 0.2365
    0.7705 1.216  0.8606
0 1 1 0.0 1.0
    0.26   1.0    1.0
0 3 1 0.0 1.0
    0.8    1.0
8 4
0 0 6 2.0 1.0
     .5484671660D+04   .1831074430D-02
     .8252349460D+03   .1395017220D-01
     .1880469580D+03   .6844507810D-01
     .5296450000D+02   .2327143360D+00
     .1689757040D+02   .4701928980D+00
     .5799635340D+01   .3585208530D+00
0 1 3 6.0 1.0
     .1553961625D+02  -.1107775490D+00   .7087426820D-01
     .3599933586D+01  -.1480262620D+00   .3397528390D+00
     .1013761750D+01   .1130767010D+01   .7271585770D+00
0 1 1 0.0 1.0
     .2700058226D+00   .1000000000D+01   .1000000000D+01
0 3 1 0.0 1.0
     .8000000000D+00   .1000000000D+01
99 0
ENDBS
SCFDIR
DFT
B3LYP
RADIAL
1
4.0
99
ANGULAR
5
0.1667 0.5 0.9 3.5 9999.0
6 10 14 18 14
END
```

```
SHRINK
6 6
LEVSHIFT
5 0
TOLDEE
10
MAXCYCLE
200
ENDSCF
```

Chapter 7

Basis set

The most common source of problems with CRYSTAL is probably connected with the basis set. It should never be forgotten that ultimately the basis functions are Bloch functions, modulated over the infinite lattice: any attempt to use large uncontracted molecular or atomic basis sets, with very diffuse functions can result in the wasting of computational resources. The densely packed nature of many crystalline structures gives rise to a large overlap between the basis functions, and a quasi-linear dependence can occur, due to numerical limitations.
The choice of the basis set (BS) is one of the critical points, due to the large variety of chemical bonding that can be found in a periodic system. For example, carbon can be involved in covalent bonds (polyacetylene, diamond) as well as in strongly ionic situations (Be_2C, where the Mulliken charge of carbon is close to -4).
Many basis sets for lighter elements and the first row transition metal ions have been developed for use in periodic systems. A selection of these which have been used in published work are available on WWW:

http://www.crystal.unito.it

We summarize here some general considerations which can be useful in the construction of a BS for periodic systems.
It is always useful to refer to some standard basis set; Pople's STO-nG, 3-21G and 6-21G have proved to be good starting points. A molecular minimal basis set can in some cases be used as it is; larger basis sets must be re-optimized specifically for the chemical periodic structure under study.
Let us explore the adequacy of the molecular BS for crystalline compounds and add some considerations which can be useful when a molecular BS must be modified or when an *ex novo* crystalline BS is defined.

7.1 Molecular BSs performance in periodic systems

Two sets of all electron basis sets are included in CRYSTAL (see Chapter 1.2):

1. Minimal STO-nG basis set of Pople and co-workers
   obtained by fitting Slater type orbitals with n contracted GTFs (n from 2 to 6, atomic number from 1 to 54) [101, 102, 103, 104].

   The above BSs are still widely used in spite of the poor quality of the resulting wave function, because they are well documented and as a rule provide quite reasonable optimized geometries (due to fortuitous cancellation of errors) at low cost.

2. "Split valence" 3-21 and 6-21 BSs.
   The core shells are described as a linear combination of 3 (up to atomic number 54) or 6 (up to atomic number 18) gaussians; the two valence shells contain two and one gaussians, respectively [105, 106]. Exponents (s and p functions of the same shell share the same exponent) and contraction coefficients have been optimized variationally for the isolated atoms.

A single set of polarization functions (p,d) can be added without causing numerical problems. Standard molecular polarization functions are usually also adequate for periodic compounds.

When free basis sets are chosen, two points should be taken into account:

1. From the point of view of CPU time, basis sets with *sp* shells (*s* and *p* functions sharing the same set of exponents) can give a saving factor as large as 4, in comparison with basis sets where s and p have different exponents.

2. As a rule, extended atomic BSs, or 'triple zeta' type BSs should be avoided. Many of the high quality molecular BSs (Roos, Dunning, Huzinaga) cannot be used in CRYSTAL without modification, because the outer functions are too diffuse. One should not forget that the real basis functions are Bloch functions.

Let us consider in more detail the possibility of using molecular BS for periodic systems. We can refer to five different situations:

Core	functions	
Valence	functions:	molecular crystals
		covalent crystals
		ionic crystals
		metals.

7.2 Core functions

In this case standard (contracted) molecular BSs can be adopted without modification, because even when very strong crystal field effects are present, the deformation of inner states is small, and can be correctly described through the linear variational parameters in SCF calculation. An adequate description of the core states is important in order to avoid large basis set superposition errors.

7.3 Valence functions

Molecular crystals

Molecular BSs, minimal and split-valence, are perfectly adequate. Tests have been performed on bulk urea [107] and oxalic acid, where the molecules are at relatively small distances, with STO-3G, 6-21, 6-21* and 6-21** BSs presenting no problem.

Covalent crystals.

Standard minimal and split valence BSs are usually adequate. In the split valence case the best exponent of the most diffuse shell is always slightly higher than the one proposed for molecules; in general it is advisable to re-optimize the exponent of this shell. This produces a slightly improved basis, while reducing the cost of the calculation. Let us consider for example the 6-21 basis set for carbon (in diamond) and silicon (bulk).
At an atomic level, the best exponent of the outer shell is 0.196 and 0.093 for C and Si, respectively. Optimization of the valence shell has been repeated in the two crystalline compounds. The innermost valence shell is essentially unaltered with respect to the atomic solution; for the outer single-gaussian shell the best exponent is around 0.22 and 0.11 bohr^{-2} for carbon and silicon, as shown in Table 7.1. The last entry of Table 7.1 refers to "catastrophic" behaviour: the low value of the exponent generates unphysical states.
A set of 5 polarization single-gaussian d functions can be added to the 6-21G basis (6-21G* BS); the best exponents for the solid are very close to those resulting from the optimization in molecular contexts: 0.8 for diamond [108] and 0.45 for silicon.

Basis sets for III-V and IV-IV semiconductors (all electron and valence electron (to be associated with effective core pseudopotentials) are given in references [109, 110].

Table 7.1: Total energy per cell and number of computed bielectronic integrals in 10^6 units (N), as a function of the exponent α (bohr^{-2}) of the most diffuse shell for carbon and silicon.

Diamond			Silicon		
a	N	Et	a	N	Et
0.296	58	-75.6633	0.168	46	-577.8099
0.276	74	-75.6728	0.153	53	-577.8181
0.256	83	-75.6779	0.138	72	-577.8231
0.236	109	-75.6800	0.123	104	-577.8268
0.216	148	-75.6802	0.108	151	-577.8276
0.196	241	-75.6783	0.093	250	-577.8266
0.176	349	catastrophe	0.078	462	catastrophe

Ionic crystals.

Cations

The classification of covalent or ionic crystals is highly conventional, many systems being midway. Let us first consider totally ionic compounds, such as LiH, MgO, or similar. For these systems the cation valence shell is completely empty. Therefore, for cations it is convenient to use a basis set containing the core functions plus an additional sp shell with a relatively high exponent. For example, we used for Mg in MgO and for Li in LiH (Li_2 O and Li_3 N) a 'valence' sp shell with exponent 0.4-0.3 and 0.5-0.6, respectively [24, 18].
The crystalline total energies obtained by using only core functions for Li or Mg and by adding a valence shell to the cation differ by 0.1 eV/atom, or less. This figure is essentially the same for a relatively large range of exponents of the valence shell (say 0.5-0.2 for Mg) [18].
It can be difficult (or impossible) to optimize the exponents of nearly empty shells: the energy decreases almost linearly with the exponent. Very low exponent values can give rise to numerical instabilities, or require the calculation of an enormous number of integrals (selected on the basis of overlap criteria). In the latter cases, when the energy gain is small ($\Delta E \leq 1$ m hartree for $\Delta\alpha = 0.2$ bohr^{-2}), it is convenient to use a relatively large exponent.

Anions

Reference to isolated ion solutions is only possible for halides, because in such cases the ions are stable even at the HF level. For other anions, which are stabilized by the crystalline field (H^-, O^{2-}, N^{3-} and also C^{4-}), the basis set must be re-designed with reference to the crystalline environment. For example, let us consider the optimization of the O^{2-} BS in Li_2O [24]. Preliminary tests indicated the fully ionic nature of the compound; the point was then to allow the valence distribution to relax in the presence of the two extra electrons. We started from a standard STO-6G BS. Two more gaussians were introduced in the 1s contraction, in order to improve the virial coefficient and total energy, as a check of wave function quality. The 6 valence gaussians were contracted according to a 411 scheme; the exponents of the two outer independent gaussians and the coefficients of the four contracted ones were optimized. Whereas the two most diffuse gaussians are more diffuse than in the neutral isolated atom (α=0.45 and 0.15 to be compared with α=0.54 and 0.24 respectively), the rest of the O^{2-} valence shell is unchanged with respect to the atomic situation. The introduction of d functions in the oxygen basis-set causes only a minor improvement in energy (1 10^{-4} hartree/cell, with a population of 0.02 electrons/atom in the cell). Ionic BSs for H and N can be found in reference 1.
For anions, re-optimization of the most diffuse valence shell is mandatory; when starting from a standard basis set, the most diffuse (or the two most diffuse) gaussians must be allowed to relax.

From covalent to ionics

Intermediate situations must be considered individually, and a certain number of tests must be performed in order to verify the adequacy of the selected BSs.
Let us consider for example α-quartz (SiO_2) and corundum (Al_2O_3). The exponent of the outer shell for the 2 cations in the 6-21G BS is 0.093 (Si) and 0.064 (Al), respectively; in both cases this function is too diffuse (in particular in the Al case it causes numerical catastrophes). For quartz, re-optimization in the bulk gives α=0.15 bohr^{-2} for Si (the dependence of total energy per Si atom on α is much smaller than the one resulting from Table 7.1; note too that the cost at α=0.15 is only 50% of the one at α=0.09). On the contrary, the best molecular and crystalline exponent (α=0.37) for oxygen coincide. Corundum is more ionic than quartz, and about 2 valence electrons are transferred to oxygen. In this case it is better to eliminate the most diffuse valence shell of Al, and to use as independent functions the two gaussians of the inner valence shells (α=0.94 and 0.20 bohr^{-2}, respectively [111]).

Metals

Very diffuse gaussians are required to reproduce the nearly uniform density characterizing simple metallic systems, such as lithium and beryllium. This is the worse situation, where a full optimization of the atomic basis set is probably impossible. Functions which are too diffuse can create numerical problems, as will be discussed below.
The optimization procedure can start from 6-21 BS; the most diffuse valence shell (exponent 0.028 for Li and 0.077 for Be) can be dropped and the innermost valence shell (exponents 0.54 and 0.10 for Li, and 1.29 and 0.268 for Be) can be split.

Table 7.2: Example of BS for metallic lithium and beryllium derived from the standard 6-21G BS

.

Lithium shell	Exp.	Coeff.		Beryllium shell	Exp.	Coeff.	
s	642.418	0.00215096		s	1264.50	0.00194336	
	96.5164	0.0162677			189.930	0.0148251	
	22.0174	0.0776383			43.1275	0.0720662	
	6.1764	0.246495			12.0889	0.237022	
	1.93511	0.467506			3.80790	0.468789	
sp	0.640	1.	1.		1.282	1.	1.
sp	0.10	1.	1.		0.27	1.	1.

At this point the outer gaussian of the 6G core contraction, with very similar exponents (0.64 and 1.28) to those of the innermost valence shell (0.54 and 1.29), can be used as an independent (sp) function, and the innermost valence shell can be eliminated.
The resulting (reasonable) BS, derived from the split valence standard one, is reported in Table 7.2. Finally, the most diffuse gaussian can be optimized; in the two cases the minimum has not been found owing to numerical instabilities.
See [112] for a more extensive discussion of the metallic lithium case.

7.4 Hints on crystalline basis set optimization

In the definition of a valence shell BS, each exponent can be varied in a relatively narrow range: in the direction of higher exponents, large overlaps with the innermost functions may occur (the rule of thumb is: exponents must be in a ratio not too far from 3; ratios smaller than

2 can give linear dependence problems); proceeding towards lower exponents, one must avoid large overlaps with a high number of neighbours (remember: the basis functions are Bloch functions).
Diffuse gaussian orbitals play a critical role in HF-LCAO calculations of crystals, especially the three-dimensional ones; they are *expensive, not always useful, in some cases dangerous.*

- Cost.

The number of integrals to be calculated increases dramatically with decreasing exponents; this effect is almost absent in molecular calculations. Table 7.1 shows that the cost of the calculation (number of bielectronic integrals) for silicon (diamond) can increase by a factor 10 (6) simply by changing the exponent of the most diffuse single-gaussian from 0.168 to 0.078 (0.296 to 0.176). The cost is largely dominated by this shell, despite the fact that large contractions are used for the 1s, 2sp and the innermost valence shell.
A high number of contracted primitives tremendously increases the integrals computation time.

- Usefulness.

In atoms and molecules a large part of the additional variational freedom provided by diffuse functions is used to describe the tails of the wave function, which are poorly represented by the $e^{-\alpha r^2}$ decay of the gaussian function. On the contrary, in crystalline compounds (in particular 3D non-metallic systems), low exponent functions do not contribute appreciably to the wave function, due to the large overlap between neighbours in all directions. A small split valence BS such as the 6-21G one, is nearer to the variational limit in crystals than in molecules.

- Numerical accuracy and catastrophic behaviour.

In some conditions, during the SCF (periodic) calculation, the system 'falls' into non-physical states, characterized by very low single particle and total energies (see for example the last entry in Table 7.1 and the above discussion on metals).
This behaviour, generically interpreted in our early papers as due to 'linear dependence', is actually due to poor accuracy in the treatment of the Coulomb and exchange series. The exchange series is much more delicate, for two reasons: first, long range contributions are not taken into account (whereas the long range Coulomb contributions are included, although in an approximate way); second, the "pseudoverlap" criteria associated with the two computational parameters ITOL4 and ITOL5 mimic only in an approximate way the real behaviour of the density matrix.
The risks of "numerical catastrophes" increase rapidly with a decreasing exponent; higher precision is required in order to obtain physical solutions.
For non-metallic systems, and split-valence type BSs, the default computational conditions given in section 1.3 are adequate for the optimization of the exponents of the valence shell and for systematic studies of the energy versus volume curves.
For metallic systems, the optimization of the energy versus exponent curve could require extremely severe conditions for the exchange series and, as a consequence, for the reciprocal space net. Reasonable values of the valence shell exponent (say 0.23 for beryllium and 0.10 for lithium, see Table 7.2), though not corresponding to a variational minimum, are reasonably adequate for the study of the structural and electronic properties of metallic systems (see reference 1).

7.5 Check on basis-set quasi-linear-dependence

In order to check the risk of linear dependence of Bloch functions, it is possible to calculate the eigenvalues of the overlap matrix in reciprocal space by running **integrals** and entering the keyword **EIGS** (input block 3, page 65). Full input (general information, geometry, basis set, SCF) is to be entered.

The overlap matrix in direct space is Fourier transformed at all the k points generated in the irreducible part of the Brillouin zone, and diagonalized. The eigenvalues are printed.
The higher the numerical accuracy obtained by severe computational conditions, the closer to 0 can be the eigenvalues without risk of numerical instabilities. Negative values indicate numerical linear dependence. The program stops after the check (even if negative eigenvalues are not detected).
The Cholesky reduction scheme [69] requires basis functions linearly independent. A symptom of numerical dependence may produce an error message in RHOLSK or CHOLSK while running **scf**.

Chapter 8

Theoretical framework

8.1 Basic equations

CRYSTAL is an *ab initio* Hartree-Fock LCAO program for the treatment of periodic systems. *LCAO*, in the present case, means that each Crystalline Orbital , $\psi_i(\mathbf{r};\mathbf{k})$, is a linear combination of Bloch functions (BF), $\phi_\mu(\mathbf{r};\mathbf{k})$, defined in terms of local functions, $\varphi_\mu(\mathbf{r})$ (here referred to as Atomic Orbitals, AOs).

$$\psi_i(\mathbf{r};\mathbf{k}) = \sum_\mu a_{\mu,i}(\mathbf{k})\phi_\mu(\mathbf{r};\mathbf{k}) \tag{8.1}$$

$$\phi_\mu(\mathbf{r};\mathbf{k}) = \sum_{\mathbf{g}} \varphi_\mu(\mathbf{r}-\mathbf{A}_\mu-\mathbf{g})\ e^{i\mathbf{k}\cdot\mathbf{g}} \tag{8.2}$$

$\mathbf{A}_\mu$ denotes the coordinate of the nucleus in the zero reference cell on which φ_μ is centred, and the $\sum_{\mathbf{g}}$ is extended to the set of all lattice vectors $\mathbf{g}$.
The local functions are expressed as linear combinations of a certain number, n_G, of individually normalized (basis set) Gaussian type functions (GTF) characterized by the same centre, with fixed coefficients, d_j and exponents, α_j, defined in the input:

$$\varphi_\mu(\mathbf{r}-\mathbf{A}_\mu-\mathbf{g}) = \sum_j^{n_G} d_j\ G(\alpha_j;\mathbf{r}-\mathbf{A}_\mu-\mathbf{g}) \tag{8.3}$$

The AOs belonging to a given atom are grouped into *shells*, λ. The shell can contain all AOs with the same quantum numbers, n and ℓ, (for instance 3s, 2p, 3d shells), or all the AOs with the same principal quantum number, n, if the number of GTFs and the corresponding exponents are the same for all of them (mainly sp shells; this is known as the *sp shells constraint*). These groupings permit a reduction in the number of auxiliary functions that need to be calculated in the evaluation of electron integrals and therefore increase the speed of calculation.
A single, normalized, s-type GTF, G_λ, is associated with each shell (the *adjoined Gaussian* of shell λ). The α exponent is the smallest of the α_j exponents of the Gaussians in the contraction. The adjoined Gaussian is used to estimate the AO overlap and select the level of approximation to be adopted for the evaluation of the integrals.
The expansion coefficients of the Bloch functions, $a_{\mu,i}(\mathbf{k})$, are calculated by solving the matrix equation for each reciprocal lattice vector, $\mathbf{k}$:

$$\mathbf{F}(\mathbf{k})\mathbf{A}(\mathbf{k}) = \mathbf{S}(\mathbf{k})\mathbf{A}(\mathbf{k})\mathbf{E}(\mathbf{k}) \tag{8.4}$$

in which $\mathbf{S}(\mathbf{k})$ is the overlap matrix over the Bloch functions, $\mathbf{E}(\mathbf{k})$ is the diagonal energy matrix and $\mathbf{F}(\mathbf{k})$ is the Fock matrix in reciprocal space:

$$\mathbf{F}(\mathbf{k}) = \sum_{\mathbf{g}} \mathbf{F}^{\mathbf{g}}\ e^{i\mathbf{k}\cdot\mathbf{g}} \tag{8.5}$$

The matrix elements of $\mathbf{F^g}$, the Fock matrix in direct space, can be written as a sum of one-electron and two-electron contributions in the basis set of the AO:

$$F_{12}^{\mathbf{g}} = H_{12}^{\mathbf{g}} + B_{12}^{\mathbf{g}} \tag{8.6}$$

The one electron contribution is the sum of the kinetic and nuclear attraction terms:

$$H_{12}^{\mathbf{g}} = T_{12}^{\mathbf{g}} + Z_{12}^{\mathbf{g}} = \langle \varphi_1^{\mathbf{0}} \mid \widehat{T} \mid \varphi_2^{\mathbf{g}} \rangle + \langle \varphi_1^{\mathbf{0}} \mid \widehat{Z} \mid \varphi_2^{\mathbf{g}} \rangle \tag{8.7}$$

In core pseudopotential calculations, $\hat{Z}$ includes the sum of the atomic pseudopotentials.
The two electron term is the sum of the Coulomb and exchange contributions:

$$B_{12}^{\mathbf{g}} = C_{12}^{\mathbf{g}} + X_{12}^{\mathbf{g}} =$$

$$\sum_{3,4} \sum_{\mathbf{n}} P_{3,4}^{\mathbf{n}} \sum_{\mathbf{h}} [(\varphi_1^{\mathbf{0}} \varphi_2^{\mathbf{g}} \mid \varphi_3^{\mathbf{h}} \varphi_4^{\mathbf{h+n}}) - \frac{1}{2} (\varphi_1^{\mathbf{0}} \varphi_3^{\mathbf{h}} \mid \varphi_2^{\mathbf{g}} \varphi_4^{\mathbf{h+n}})] \tag{8.8}$$

The Coulomb interactions, that is, those of electron-nucleus, electron-electron and nucleus-nucleus, are individually divergent, due to the infinite size of the system. The grouping of corresponding terms is necessary in order to eliminate this divergence.
The $\mathbf{P^n}$ density matrix elements in the AOs basis set are computed by integration over the volume of the Brillouin zone (BZ),

$$P_{3,4}^{\mathbf{n}} = 2 \int_{BZ} d\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{n}} \sum_j a_{3j}^*(\mathbf{k}) a_{4j}(\mathbf{k}) \theta(\epsilon_F - \epsilon_j(\mathbf{k})) \tag{8.9}$$

where a_{in} denotes the i-th component of the n-th eigenvector, θ is the step function, ϵ_F, the Fermi energy and ϵ_n, the n-th eigenvalue. The total electronic energy per unit cell is given by:

$$E^{elec} = \frac{1}{2} \sum_{1,2} \sum_{\mathbf{g}} P_{12}^{\mathbf{g}} (H_{12}^{\mathbf{g}} + B_{12}^{\mathbf{g}}) \tag{8.10}$$

A discussion of the different contributions to the total energy is presented in [113, 97] and in Chapter 11 of reference [12].

$$E^{coul} = \frac{1}{2} \sum_{1,2} \sum_{\mathbf{g}} P_{12}^{\mathbf{g}} \sum_{3,4} \sum_{\mathbf{n}} P_{3,4}^{\mathbf{n}} \sum_{\mathbf{h}} [(\varphi_1^{\mathbf{0}} \varphi_2^{\mathbf{g}} \mid \varphi_3^{\mathbf{h}} \varphi_4^{\mathbf{h+n}})] \tag{8.11}$$

$$E^{exch} = -\frac{1}{4} \sum_{1,2} \sum_{\mathbf{g}} P_{12}^{\mathbf{g}} \sum_{34} \sum_{\mathbf{n}} P_{34}^{\mathbf{n}} \sum_{\mathbf{h}} [(\varphi_1^{\mathbf{0}} \varphi_3^{\mathbf{h}} \mid \varphi_2^{\mathbf{g}} \varphi_4^{\mathbf{h+n}})] \tag{8.12}$$

8.2 Remarks on the evaluation of the integrals

The approach adopted for the treatment of the Coulomb and exchange series is based on a few simple ideas and on a few general tools, which can be summarized as follows:

1. Where possible, terms of the Coulomb series are aggregated so as to reduce the number of integrals to be evaluated;

2. Exchange integrals which will combine with small density matrix elements are disregarded;

3. Integrals between non-overlapping distributions are approximated;

4. Approximations for large integrals must be very accurate; for small integrals large percentage errors can be accepted;

5. Selection must be very efficient, because a large number of possible terms must be checked (adjoined Gaussians are very useful from this point of view).

8.3 Treatment of the Coulomb series

For the evaluation of the Coulomb contributions to the total energy and Fock matrix, correct coupling of electron-nucleus and electron-electron interactions is essential. The computational technique for doing so was presented by Dovesi et al [114] and by Saunders et al. [97]. It may be summarized as follows.
Consider the Coulomb bielectronic contribution to the Fock matrix ($C^{\mathbf{g}}_{12}$) and to the total energy :

$$E^{coul}_{ee} = \frac{1}{2}\sum_{1,2}\sum_{\mathbf{g}} P^{\mathbf{g}}_{12}\sum_{3,4}\sum_{\mathbf{n}} P^{\mathbf{n}}_{3,4}\sum_{\mathbf{h}}[(\varphi^{\mathbf{0}}_1\varphi^{\mathbf{g}}_2 \mid \varphi^{\mathbf{h}}_3\varphi^{\mathbf{h}+\mathbf{n}}_4) \tag{8.13}$$

Seven indices are involved in equation 8.13; four of them (1, 2, 3 and 4) refer to the AOs of the unit cell; in principle, the other three ($\mathbf{g}$, $\mathbf{n}$ and $\mathbf{h}$) span the infinite set of translation vectors: for example, $\varphi^{\mathbf{g}}_2(\mathbf{r})$ is AO number 2 in cell $\mathbf{g}$. P is the density matrix; the usual notation is used for the bielectronic integrals. Due to the localized nature of the basis set, the total charges, q_1 and q_2, associated with the two pseudo-overlap distributions: $\{G_{1\mathbf{0}}G_{2\mathbf{g}}\}$ and $\{G_{3\mathbf{h}}G_{4\mathbf{h}+\mathbf{n}}\}$, decay exponentially to zero with increasing $|\mathbf{g}|$ and $|\mathbf{n}|$ (for example, G_1 is the adjoined gaussian of the shell to which φ_1 belongs).
A *Coulomb overlap* parameter, S_c, can be defined in such a way that when either q_1 or q_2 are smaller than S_c, the bielectronic integral is disregarded, and the sum over $\mathbf{g}$ or $\mathbf{n}$ truncated. The ITOL1 input parameter is defined as **ITOL1**$=-log_{10}S_c$. The same parameter value is used for selecting overlap, kinetic, and multipole integrals.
The problem of the $\mathbf{h}$ summation in equation 8.13 is more delicate, $\mathbf{h}$ being related to the distance between the two interacting distributions. The multipolar expansion scheme illustrated below is particularly effective when large unit cell or low dimensionality systems are considered. The electron-electron and electron-nuclei series ($C^{\mathbf{g}}_{12}$ and $Z^{\mathbf{g}}_{12}$) can be rearranged as follows:

1. Mulliken shell *net* charge distributions are defined as :

$$\rho_\lambda(\mathbf{r}-\mathbf{h}) \equiv \{\lambda\}' \equiv \{\lambda\} - Z_\lambda = \sum_{3\in\lambda}\sum_{4\mathbf{n}} P^{\mathbf{n}}_{34}\ \varphi_3(\mathbf{r}-\mathbf{h})\ \varphi_4(\mathbf{r}-\mathbf{h}-\mathbf{n}) - Z_\lambda \tag{8.14}$$

   where Z_λ is the fraction of nuclear charge formally attributed to shell λ, and $\{\lambda\}$ is the electron charge distribution of shell λ.

2. Z and C contributions are reordered:

$$C^{\mathbf{g}}_{12} + Z^{\mathbf{g}}_{12} = \sum_{\lambda}\sum_{\mathbf{h}}\int d\mathbf{r}\, d\mathbf{r}'\varphi^{\mathbf{0}}_1(\mathbf{r})\ \varphi^{\mathbf{g}}_2(\mathbf{r})\ |\mathbf{r}-\mathbf{r}'-\mathbf{h}|^{-1}\ \rho_\lambda(\mathbf{r}'-\mathbf{h}) \tag{8.15}$$

3. For a given shell λ, there is a finite set B_λ of $\mathbf{h}$ vectors for which the two interacting distributions overlap; in this B_λ zone (*bielectronic zone*), all the bielectronic integrals are evaluated explicitly. In the outer, infinite region which we define as M_λ, complementary to B_λ (the *mono-electronic zone*), ρ_λ can be expanded in multipoles and the series can be evaluated to infinity analytically, using Ewald's method combined with recursion formulae [97].

The resulting expression for the Coulomb contribution to the Fock matrix is:

$$\begin{aligned} C^{\mathbf{g}}_{12} + Z^{\mathbf{g}}_{12} = \quad & \textstyle\sum_\lambda\{\sum^{B_\lambda}_{\mathbf{h}}[\sum_{3\in\lambda}\sum_4\sum_{\mathbf{n}} P^{\mathbf{n}}_{34}(\varphi^{\mathbf{0}}_1\varphi^{\mathbf{g}}_2 \mid \varphi^{\mathbf{h}}_3\varphi^{\mathbf{h}+\mathbf{n}}_4) + \\ - \quad & \textstyle\sum_{\ell,m}\gamma^m_\ell(\mathbf{A}_\lambda;\{\lambda\})\Phi^m_\ell(12\mathbf{g};A_\lambda+\mathbf{h})] + \\ + \quad & \textstyle\sum_{\mathbf{h}}\sum_{\ell,m}\gamma^m_\ell(A_\lambda;\{\lambda\}')\Phi^m_\ell(12\mathbf{g};\mathbf{A}_\lambda+\mathbf{h})\} \end{aligned} \tag{8.16}$$

where:

$$\gamma^m_\ell(\mathbf{A}_\lambda;\{\lambda\}) = \int d\mathbf{r}\ \rho_\lambda(\mathbf{r}-\mathbf{A}_\lambda)N^m_\ell X^m_\ell(\mathbf{r}-\mathbf{A}_\lambda) \tag{8.17}$$

$$\Phi^m_\ell(12\mathbf{g};\mathbf{A}_\lambda+\mathbf{h}) = \int d\mathbf{r}\varphi^{\mathbf{0}}_1(\mathbf{r})\varphi^{\mathbf{g}}_2(\mathbf{r})X^m_\ell(\mathbf{r}-\mathbf{A}_\lambda-\mathbf{h})\ |\mathbf{r}-\mathbf{A}_\lambda-\mathbf{h}|^{-2\ell-1} \tag{8.18}$$

The Ewald term in eq. 8.16 includes zones $B_\lambda + M_\lambda$. The contribution from B_λ is subtracted. The X_ℓ^m functions entering in the definition of the multipoles and field terms are real, solid harmonics, and N_ℓ^m, the corresponding normalization coefficients.
The advantage of using equation 8.16 is that many four-centre (long-range) integrals can be replaced by fewer three-centre integrals.
The attribution of the interaction between $\rho_1 = \{1\mathbf{0},\, 2\mathbf{g}\}$ and ρ_λ to the *exact*, short-range or to the *approximate*, long-range zone is performed by comparing the penetration between ρ_1 and ρ_λ with the ITOL2 input parameter (if **ITOL2**$> -\log S_{1\lambda}$, then ρ_λ is attributed to the *exact* B_λ zone).
The multipolar expansion in the approximate zone is truncated at $L = \ell^{max}$. The default value of L is 4; the maximum possible value is 6, the minimum suggested value, 2 (defined via the input keyword **POLEORDR**, input block 3, page 73).

8.4 The exchange series

The exchange series does not require particular manipulations of the kind discussed in the previous section for the Coulomb series, but needs a careful selection of the terms contributing appreciably to the Fock operator and to the total energy [115]. The exchange contribution to the total energy can be written as follows:

$$E^{ex} = \frac{1}{2}\sum_{12}\sum_{\mathbf{g}} P_{12}^{\mathbf{g}}\, [-\frac{1}{2}\sum_{34}\sum_{\mathbf{n}} P_{34}^{\mathbf{n}} \sum_{\mathbf{h}} (\varphi_1^{\mathbf{0}}\varphi_3^{\mathbf{h}} \mid \varphi_2^{\mathbf{g}}\varphi_4^{\mathbf{h}+\mathbf{n}})] \tag{8.19}$$

where the term in square brackets is the exchange contribution to the $12\mathbf{g}$ element of the direct space Fock matrix. E^{ex} has no counterpart of opposite sign as the Coulomb term has; hence, it must converge by itself.
The $\mathbf{h}$ summation can be truncated after a few terms, since the $\{\varphi_1^{\mathbf{0}}\varphi_3^{\mathbf{h}}\}$ overlap distribution decays exponentially as $\mathbf{h}$ increases. Similar considerations apply to the second charge distribution. In CRYSTAL, the $\mathbf{h}$ summation is, therefore, truncated when the charge associated with either $\{G_1\mathbf{0}\ G_3\mathbf{h}\}$ or $\{G_2\mathbf{g}\ G_4\mathbf{h}+\mathbf{n}\}$ is smaller than $10^{-\mathbf{ITOL3}}$.
The situation is more complicated when $\mathbf{g}$ and $\mathbf{n}$ summations are analysed. Let us consider the leading terms at large distance, corresponding to $\varphi_1{=}\varphi_3, \varphi_2{=}\varphi_4$, $\mathbf{h} = \mathbf{0}$ and $\mathbf{n} = \mathbf{g}$:

$$e_{12}^{\mathbf{g}} = -1/4(P_{12}^{\mathbf{g}})^2(1\mathbf{0}\,1\mathbf{0}|2\mathbf{g}\,2\mathbf{g}) = -(p^{\mathbf{g}})^2/(4|\mathbf{g}|) \tag{8.20}$$

(Here $p^{\mathbf{g}}$ indicates the dominant P matrix element at long range). Since the number of terms per unit distance of this kind increases as $|\mathbf{g}|^{d-1}$, where d is the dimensionality of the system, it is clear that the convergence of the series depends critically on the long range behaviour of the bond order matrix.
Cancellation effects, associated in particular with the oscillatory behaviour of the density matrix in metallic systems, are not predominant at long range. Even if the actual behaviour of the P matrix elements cannot be predicted because it depends in a complicated way on the physical nature of the compound [90], on orthogonality constraints and on basis set quality, the different range of valence and core elements can be exploited by adopting a *pseudoverlap* criterion. This consists in truncating $\mathbf{g}$ summations when the $\int d\mathbf{r}\varphi_1^{\mathbf{0}}\varphi_2^{\mathbf{g}}$ overlap is smaller than a given threshold, defined as $P_{ex}^{\mathbf{g}}$ (where **ITOL4** = -log_{10} ($P_{ex}^{\mathbf{g}}$)) and also truncating the $\mathbf{n}$ summation when $\int d\mathbf{r}\varphi_3^{\mathbf{0}}\varphi_4^{\mathbf{n}}$ overlap is smaller than the threshold, $P_{ex}^{\mathbf{n}}$ (**ITOL5** = -log_{10} ($P_{ex}^{\mathbf{n}}$)).
Despite its partially arbitrary nature, this criterion presents some advantages with respect to other more elaborate schemes: it is similar to the other truncation schemes (ITOL1, ITOL2, ITOL3), and so the same classification tables can be used; it is, in addition, reasonably efficient in terms of space occupation and computer time.
This truncation scheme is symmetric with respect to the $\mathbf{g}$ and $\mathbf{n}$ summations. However, if account is not taken of the different role of the two summations in the SC (Self Consistent) stage, distortions may be generated in the exchange field as felt by charge distributions $\varphi_1\varphi_2^T$, where T labels the largest (in modulus) $\mathbf{g}$ vector taken into account according to ITOL4. This distortion may be variationally *exploited*, and unphysically large density matrix elements build

up progressively along the SC stage, eventually leading to catastrophic behaviour (see Chapter II.5 of reference [22] for a discussion of this point). In order to overcome this problem, the threshold, $P_{ex}^{\mathbf{n}}$ (**ITOL5**) for $\mathbf{n}$ summation must be more severe than that for $\mathbf{g}$ summation (**ITOL4**). In this way, all the integrals whose second pseudo charge $\int d\mathbf{r}\varphi_3^{\mathbf{0}}\varphi_4^{\mathbf{n}}$ is larger than $P_{ex}^{\mathbf{n}}$ are taken into account. A difference in the two thresholds ranging from three to eight orders of magnitude is sufficient to stabilize the SC behaviour in most cases.

8.5 Bipolar expansion approximation of Coulomb and exchange integrals

We may now return to the partition of the $\mathbf{h}$ summation in the Coulomb series shown in equation 8.13. Consider one contribution to the charge distribution of electron 1, centred in the reference cell: $\rho^{\mathbf{0}} = \varphi_1^{\mathbf{0}}\varphi_2^{\mathbf{g}}$; now consider the charge distribution $\rho_\lambda(\mathbf{h})$ of shell λ centred in cell $\mathbf{h}$ (equation 8.14). For small $|\mathbf{h}|$ values, ρ_λ and $\rho^{\mathbf{0}}$ overlap, so that all the related bielectronic integrals must be evaluated exactly, one by one; for larger values of $|\mathbf{h}|$, ρ_λ is external to ρ^0, so that all the related bielectronic integrals are grouped and evaluated in an approximate way through the multipolar expansion of ρ_λ .
However, in many instances, although ρ_λ is not external to ρ^0, the two-centre $\varphi_3^{\mathbf{h}}\varphi_4^{\mathbf{h}+\mathbf{n}}$ contributions to ρ_λ are external to $\rho^0 = \varphi_1^{\mathbf{0}}\varphi_2^{\mathbf{g}}$; in this case, instead of exactly evaluating the bielectronic integral, a two-centre truncated bipolar expansion can be used (see Chapter II.4.c in reference [22] and references therein).
In order to decide to which zone a shell may be ascribed, we proceed as follows: when, for a given pair of shells $\lambda_1^{\mathbf{0}}\lambda_2^{\mathbf{g}}$, shell $\lambda_3^{\mathbf{h}}$ is attributed to the B (*bielectronic*) zone, the penetration between the products of adjoined Gaussians $G_1^{\mathbf{0}}G_2^{\mathbf{g}}$ and $G_3^{\mathbf{h}}G_4^{\mathbf{h}+\mathbf{n}}$ is estimated: the default value of the penetration parameter is 14, and the block of bielectronic integrals is attributed accordingly to the b_e (*exact*) or to the b_b (*bipolar*) zone. The set of $\mathbf{h}$ vectors defining the B zone of $\rho^{\mathbf{0}}= \{12\mathbf{g}\}$ and $\rho_\lambda \equiv \{\lambda_3\}$ is then split into two subsets, which are specific for each partner λ_4^l of λ_3.
A similar scheme is adopted for the selected exchange integrals (see previous section) whose pseudo charges do not overlap appreciably. The default value of the penetration parameter is 10.
The total energy change due to the bipolar expansion approximation should not be greater than 10^{-4} hartree/atom; exact evaluation of all the bielectronic integrals (obtained by setting the penetration parameter value > 20000) increases the computational cost by a factor of between 1.3 and 3. Multipolar expansion is very efficient, because the following two conditions are fulfilled:

1. A general algorithm is available for reaching high ℓ values easily and economically [114, 97]. The maximum allowed value is ℓ=6.

2. The multipolar series converges rapidly, either because the interacting distributions are nearly spherical (shell expansion), or because their functional expression is such that their multipoles are zero above a certain (low) ℓ value.

8.6 Exploitation of symmetry

Translational symmetry allows the factorization of the eigenvalue problem in periodic calculations, because the Bloch functions are a basis for irreducible representations of the translational group.
In periodic calculations, point symmetry is exploited to reduce the number of points for which the matrix equations are to be solved. Point symmetry is also explicitly used in the reconstruction of the Hamiltonian, which is totally symmetric with respect to the point group operators of the system.
In the HF-CO-LCAO scheme, the very extensive use of point symmetry allows us to evaluate bielectronic and mono-electronic integrals with saving factors as large as h in the number of bielectronic integrals to be computed or h^2 in the number of those to be stored for the SCF part

of the calculation, where h is the order of the point group. The main steps of the procedure [116] can be summarized as follows:

- The set of Coulomb and exchange integrals whose 3,4 indices identify translationally equivalent pairs of AOs, so that the associated element of the density matrix P_{34} is the same, are summed together to give D_{1234} elements:

$$D_{1,2T;3,4Q} = \sum_{\mathbf{n}} [(\varphi_1^{\mathbf{0}} \varphi_2^{\mathbf{g}} \mid \varphi_3^{\mathbf{h}} \varphi_4^{\mathbf{h}+\mathbf{n}}) - 1/2(\varphi_1^{\mathbf{0}} \varphi_3^{\mathbf{h}} \mid \varphi_2^{\mathbf{g}} \varphi_4^{\mathbf{h}+\mathbf{n}})] \tag{8.21}$$

- The products of AOs $\varphi_1\varphi_2$ (and $\varphi_3\varphi_4$) are classified in symmetry-related sets; using the fact that the Fock matrix is totally symmetric, only those quantities are evaluated whose indices $1, 2$ refer to the first member of a symmetry set. The corresponding saving factor is as large as h.
- Using the symmetry properties of the density matrix, D quantities referring to $3, 4$, couples belonging to the same symmetry set (and with the same $1, 2\mathbf{g}$ index) can be combined after multiplication by appropriate symmetry matrices, so that a single quantity for each $3, 4$ symmetry set is to be stored, with a saving factor in storage of the order of h.
- The symmetry $\mathrm{P}_{34}^{\mathbf{n}} = \mathrm{P}_{43}^{-\mathbf{n}}$ is exploited.
- The symmetry $\mathrm{F}_{12}^{\mathbf{g}} = \mathrm{F}_{21}^{-\mathbf{g}}$ is exploited.

Symmetry-adapted Crystalline Orbitals

A computational procedure for generating space-symmetry-adapted Bloch functions, when BF are built from a basis of local functions (AO), is implemented in the CRYSTAL98 code. The method, that applies to any space group and AOs of any quantum number, is based on the diagonalization of Dirac characters. For its implementation, it does not require character tables or related data as an input, since the information is automatically generated starting from the space group symbol and the AO basis set. Formal aspects of the method, not available in textbooks, are discussed in:

C. Zicovich-Wilson and R. Dovesi
On the use of Symmetry Adapted Crystalline Orbitals in SCF-LCAO periodic calculations. I. The construction of the Symmetrized Orbitals
Int. J. Quantum Chem. **67**, 299–310 (1998)

C. Zicovich-Wilson and R. Dovesi
On the use of Symmetry Adapted Crystalline Orbitals in SCF-LCAO periodic calculations. II. Implementation of the Self-Consistent-Field scheme and examples
Int. J. Quantum Chem. **67**, 311–320 (1998).

The following table presents the performance obtained with the new method. In all cases convergence is reached in ten cycles.

System	Chabazite			Pyrope	Faujasite
Space Group		$R\bar{3}m$		$Ia3d$	$Fd3m$
N. of atoms		36		80	144
N. of AOs		432		1200	1728
N. symmetry operators	12	6	3	48	48
CPU time (sec) on IBM RISC-6000/365					
integrals	447	900	1945	4286	815
Atomic BF(ABF) **scf** (total)	1380	2162	4613	24143	50975
Atomic BF **scf** (diagonalization)	898	898	898	19833	44970
Symmetry Adapted BF (SABF) **scf** (total)	526	1391	4335	3394	2729
Symmetry Adapted BF **scf** (diagonalization)	42	97	570	312	523
ABF/SABF **scf** time	2.62	1.55	1.06	7.11	18.7

8.7 Reciprocal space integration

The integration in reciprocal space is an important aspect of *ab initio* calculations for periodic structures. The problem arises at each stage of the self-consistent procedure, when determining the Fermi energy, ϵ_F, when reconstructing the one-electron density matrix, and, after self-consistency is reached, when calculating the density of states (DOS) and a number of observable quantities. The P matrix in direct space is computed following equation 8.9. The technique adopted to compute ϵ_F and the P matrix in the SCF step is described in reference [117]. The Fourier-Legendre technique presented in Chapter II.6 of reference [22] is adopted in the calculation of total and projected DOS. The Fermi energy and the integral in equation 8.9 are evaluated starting from the knowledge of the eigenvalues, ϵ_n(k) and the eigenvectors, $a_{\mu n}(\mathbf{k})$, at a certain set of sampling points, $\{\kappa\}$. In 3D crystals, the sampling points belong to a lattice (called the *Monkhorst net*, [19]) with basis vectors $\mathbf{b}_1/s_1,\ \mathbf{b}_2/s_2,\ \mathbf{b}_3/s_3$, where $\mathbf{b}_1, \mathbf{b}_2$ and $\mathbf{b}_3$ are the ordinary reciprocal lattice vectors; s_1, s_2 and s_3 (input as IS1, IS2 and IS3) are integer *shrinking factors*. Unless otherwise specified, IS1=IS2=IS3=IS. In 2D crystals, IS3 is set equal to 1; in 1D crystals both IS2 and IS3 are set equal to 1. Only points of the Monkhorst net belonging to the irreducible part of the Brillouin Zone (BZ) are considered, with associated geometrical weights, w_i.
In the selection of the κ points for non-centrosymmetric crystal, time-reversal symmetry is exploited ($\epsilon_n(\kappa) = \epsilon_n(-\kappa)$).
The number of inequivalent sampling points, κ_i, is asymptotically given by the product of the shrinking factors divided by the order of the point group. In high symmetry systems and with small s_i values, it may be considerably larger because many points lie on symmetry planes or axes.
Two completely different situations (which are automatically identified by the code) must now be considered, depending on whether the system is an insulator (or zero gap semiconductor), or a conductor. In the former case, all bands are either fully occupied or vacant. The identification of ϵ_F is elementary, and the Fourier transform expressed by equation 8.9 is reduced to a weighted sum of the integrand function over the set $\{\kappa_\mathbf{i}\}$ with weights w_i, the sum over n being limited to occupied bands.
The case of conductors is more complicated; an additional parameter, ISP, enter into play. ISP (or ISP1, ISP2, ISP3) are *Gilat shrinking factors* which define a net *Gilat net* [20, 21] completely analogous to the Monkhorst net. The value of ISP is larger than IS (by up to a factor of 2), giving a denser net.
In high symmetry systems, it is convenient to assign IS *magic* values such that all low multiplicity (high symmetry) points belong to the Monkhorst lattice. Although this choice does not correspond to maximum efficiency, it gives a safer estimate of the integral.
The value assigned to ISP is irrelevant for non-conductors. However, a non-conductor may give rise to a conducting structure at the initial stages of the SCF cycle, owing, for instance, to a very unbalanced initial guess of the density matrix. The ISP parameter must therefore be defined in all cases.

8.8 Electron momentum density and related quantities

Three functions may be computed which have the same information content but different use in the discussion of theoretical and experimental results; the momentum density itself, $\rho(\underline{p})$ or EMD; the Compton profile function, J($\underline{p}$) or CP; the autocorrelation function, or reciprocal space form factor, B($\underline{r}$) or BR.
With reference to a Crystalline-Orbital (CO)-LCAO wave function, the EMD can be expressed as the sum of the squared moduli of the occupied COs in a momentum representation, or equivalently, as the diagonal element of the six-dimensional Fourier transform of the one electron density matrix from configuration to momentum space:

$$\rho(\underline{p}) = 1/V_{BZ} \sum_j \int_{BZ} d\underline{k} \; |\psi_j(\underline{k},\underline{p})|^2 \; \theta[\epsilon_F - \epsilon_j(\underline{k})] = \tag{8.22}$$

$$= \sum_j \sum_{\mu\nu} e^{-i\underline{p}\cdot(\underline{s}_\mu - \underline{s}_\nu)} a_{\mu j}(\underline{p}^0) a^*_{\nu j}(\underline{p}^0) \chi_\mu(\underline{p}) \chi^*_\nu(\underline{p}) \; \theta[\epsilon_F - \epsilon_j(\underline{p}^0)] \tag{8.23}$$

$$\rho(\underline{p}) = N^{-1} \int d\underline{r} d\underline{r}' e^{-i\underline{p}\cdot(\underline{r}' - \underline{r})} \rho(\underline{r} - \underline{r}') \tag{8.24}$$

$$= \sum_{\mu\nu} \sum_{\underline{g}} P^{\underline{g}}_{\mu\nu} e^{-i\underline{p}\cdot(\underline{g} + \underline{s}_\mu - \underline{s}_\nu)} \chi_\mu(\underline{p}) \chi^*_\nu(\underline{p}) \tag{8.25}$$

In the above equations $\underline{p}^0$ is the value of momentum in the Brillouin zone (BZ), which is related to $\underline{p}$ by a reciprocal lattice vector $\underline{K}$, $\underline{s}_\mu$ is the fractional coordinate of the χ_μ centre, and $\chi_\mu(\underline{p})$ is the Fourier transform of $\chi_\mu(\underline{r})$, calculated analytically:

$$\chi_\mu(\underline{p}) = \int dr \chi_\mu(\underline{r}) \; e^{i\underline{p}\cdot\underline{r}} \tag{8.26}$$

Equation 8.25 is used by default to compute the core band contribution, and equation 8.23 the valence band contribution.
The CPs are obtained by 2D integration of the EMD over a plane through $\underline{p}$ and perpendicular to the $\underline{p}$ direction. After indicating with $\underline{p}_\perp$ the general vector perpendicular to $\underline{p}$, we have:

$$J(\underline{p}) = \int d\underline{p}'_\perp \rho(\underline{p} + \underline{p}'_\perp) \tag{8.27}$$

It is customary to make reference to CPs as functions of a single variable p, with reference to a particular direction $< hkl >$ identified by a vector

$$\mathbf{e} = (h\mathbf{a}_1 + k\mathbf{a}_2 + l\mathbf{a}_3)/|(h\mathbf{a}_1 + k\mathbf{a}_2 + l\mathbf{a}_3)|$$

We have:

$$J_{<hkl>}(p) = J(p \; \underline{e}) \tag{8.28}$$

The function $J_{<hkl>}(p)$ will be referred to as directional CPs.
The weighted average of the directional CPs over all directions is the average CP.

In the so called impulse approximation, $J_{<hkl>}(p)$ may be related to the experimental CPs, after correction for the effect of limited resolution as a convolution of the "infinite resolution" results, $J^0_{<hkl>}(p)$, with a normalized function characterized by a given standard deviation σ:

$$J^\sigma_{<hkl>}(p) = \int_{-\infty}^{+\infty} dp' J^0_{<hkl>}(p') g_\sigma(p - p') \tag{8.29}$$

In CRYSTAL g_σ is a gaussian function with standard deviation σ. Once the directional CPs are available, the numerical evaluation of the corresponding autocorrelation function, or reciprocal space form factor, B($\underline{r}$) is given by the 1D Fourier Transform:

$$B_{<hkl>}(r) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dp J_{<hkl>}(p) e^{i \; pr} \tag{8.30}$$

The average Compton profile can be evaluated from the average EMD:

$$\overline{J(q)} = \int_0^q \rho(\underline{p}) p dp \tag{8.31}$$

and can be used for the evaluation of the kinetic energy:

$$KE = \int_0^\infty p^2 \overline{J(p)} dp \tag{8.32}$$

8.9 Elastic Moduli of Periodic Systems

The elastic constants are *second* derivatives of the energy density with respect to strain components:

$$C_{ij} = 1/V \cdot \frac{\partial^2 E}{\partial \epsilon_i \partial \epsilon_j} \tag{8.33}$$

where V is the volume of the cell. The energy derivatives must be evaluated numerically. Particular care is required in the selection of the computational parameters and of the points where the energy is evaluated, in order to avoid large numerical errors in the fitting procedure (**FIXINDEX**, page 67; **OPTGEOM**, page 82).

See **http://www.crystal.unito.it** ⇒ **tutorials** ⇒ **Elastic and piezoelectric tensors**

When the unit cell is deformed, the point group is reduced to a subgroup of the original point group (see examples below). The new point group is automatically selected by the code. Off-diagonal (partial derivatives) elastic constants can be computed as linear combinations of single-variable energy curves. For example, for a cubic system, C_{12} can be obtained from B=$(C_{11} + 2C_{12})/3$ and $(C_{11} - C_{12})$ (see examples below). Following the deformation of the unit cell, internal relaxation of the atoms may be necessary (depending on the space group symmetry) See test 20, referring to Li_2 O.
The analysis of the point group at the atomic positions (printed by the **ATOMSYMM** option, page 28) is useful in finding the atomic coordinates to be relaxed. Examples of deformation strategies are discussed in references [24, 118].
In a crystalline system a point $\mathbf{r}$ is usually defined in terms of its fractionary components:

$$\mathbf{r} = \mathbf{h}\ \mathbf{L}_p$$

where :

$$\mathbf{L}_p = \begin{bmatrix} \mathbf{l}_1 \\ \mathbf{l}_2 \\ \mathbf{l}_3 \end{bmatrix} = \begin{bmatrix} l_{1x} & l_{1y} & l_{1z} \\ l_{2x} & l_{2y} & l_{2z} \\ l_{3x} & l_{3y} & l_{3z} \end{bmatrix} \tag{8.34}$$

$$V = det(\mathbf{L}_p)$$

$\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3$ are the fundamental vectors of the primitive cell, $\mathbf{h}$ is the fractional vector and V the cell volume.
$\mathbf{L}_p$ can be computed from the six cell parameters $a, b, c, \alpha, \beta, \gamma$. For instance, the matrix $\mathbf{L}_p$ for a face centred cubic lattice with lattice parameter a has the form:

$$\mathbf{L}_p = \begin{bmatrix} 0 & a/2 & a/2 \\ a/2 & 0 & a/2 \\ a/2 & a/2 & 0 \end{bmatrix}$$

Under an elastic strain, any particle at $\mathbf{r}$ migrates microscopically to $\mathbf{r}'$ according to the relation:

$$\mathbf{r}' = \mathbf{r}\ (\mathbf{I} + \epsilon)$$

where ϵ is the symmetric Lagrangian elastic tensor.
In the deformed crystalline system:

$$\mathbf{r}' = \mathbf{h}\ \mathbf{L}'_p$$

$$\mathbf{L}'_p = (\mathbf{I} + \epsilon)\mathbf{L}_p \tag{8.35}$$

or:

$$\mathbf{L}'_p = \mathbf{L}_p + \mathbf{Z} \tag{8.36}$$

where

$$\mathbf{Z} = \epsilon\ \mathbf{L}_p$$

$$V' = det(\mathbf{L}'_p)$$

The deformation may be constrained to be volume-conserving, in which case the lattice vectors of the distorted cell must be scaled as follows:

$$\mathbf{L}_p" = \mathbf{L}'_p\dot{(}V/V')^{1/3} \tag{8.37}$$

If a non-symmetric Lagrangian elastic tensor, η, is used, instead of ϵ, the deformation is the sum of a strain (ϵ) and a rotation (ω) of the crystal:

$$\epsilon = (\eta + \eta^{+})/2$$

$$\omega = (\eta - \eta^{+})/2$$

The total energy of the crystal is invariant to a pure rotation, which allows non-symmetric η matrices to be employed. However, a non-symmetric deformation will lower the symmetry of the system, and therefore increase the complexity of the calculation, since the cost required is roughly inversely proportional to the order of the point group.
The elastic constants of a crystal are defined as the second derivatives of the energy with respect to the elements of the infinitesimal Lagrangian strain tensor ϵ.
Let us define, according to the Voigt convention:

$$\begin{array}{ll} \epsilon_1 = \epsilon_{11} & \epsilon_4 = \epsilon_{32} + \epsilon_{23} \\ \epsilon_2 = \epsilon_{22} & \epsilon_5 = \epsilon_{13} + \epsilon_{31} \\ \epsilon_3 = \epsilon_{33} & \epsilon_6 = \epsilon_{12} + \epsilon_{21} \end{array}$$

A Taylor expansion of the energy of the unit cell to second order in the strain components yields:

$$E(\epsilon) = E(\mathbf{0}) + \sum_i^6 \frac{\partial E}{\partial \epsilon_i}\epsilon_i + 1/2\sum_{i,j}^6 \frac{\partial^2 E}{\partial \epsilon_i \partial \epsilon_j}\epsilon_i\epsilon_j \tag{8.38}$$

If $E(\mathbf{0})$ refers to the equilibrium configuration the first derivative is zero, since there is no force on any atom in equilibrium. The elastic constants of the system can be obtained by evaluating the energy as a function of deformations of the unit cell parameters. The indices of the non-zero element(s) (in the Voigt convention) of the ϵ matrix give the corresponding elastic constants.

Examples of ϵ matrices for cubic systems

Consider a face-centred cubic system, for example Li_2O, with the Fm3m space group.
For cubic systems there are only three independent elastic constants (C_{11}, C_{12} and C_{44}), as the symmetry analysis shows that:

$$\begin{array}{llll} C_{11} & = C_{22} & = C_{33}; & \\ C_{44} & = C_{55} & = C_{66}; & \\ C_{12} & = C_{13} & = C_{23}; & \\ C_{ij} & = 0 & \text{for } i = 1,6, \quad j = 4,6 \quad \text{and } i \neq j. \end{array}$$

Calculation of C_{11}

The ϵ matrix for the calculation of C_{11} is

$$\epsilon = \begin{bmatrix} \delta & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The energy expression is:

$$E(\delta) = E(0) + 1/2\frac{\partial^2 E}{\partial \epsilon_1^2}\delta^2 + \cdots = a + b\delta^2 + c\delta^3 \cdots$$

where a, b, c are the coefficients of a polynomial fit of E versus δ, usually truncated to fourth order (see examples below). Then

$$C_{11} = 1/V \frac{\partial^2 E}{\partial \epsilon_1^2} = \frac{2b}{V}$$

The above distortion reduces the number of point symmetry operators to 12 (tetragonal distortion).

Calculation of $C_{11} - C_{12}$

The ϵ matrix for the calculation of the $C_{11} - C_{12}$ combination has the form:

$$\epsilon = \begin{bmatrix} \delta & 0 & 0 \\ 0 & -\delta & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The energy expression is:

$$\begin{aligned} E(\epsilon_1, \epsilon_2) &= E(0,0) + 1/2 \frac{\partial^2 E}{\partial \epsilon_1^2} \delta^2 + 1/2 \frac{\partial^2 E}{\partial \epsilon_2^2} \delta^2 - \frac{\partial^2 E}{\partial \epsilon_1 \partial \epsilon_2} \delta^2 + \cdots = \\ &= E(0,0) + V(C_{11} - C_{12})\delta^2 + \cdots = a + b\delta^2 + \cdots \end{aligned}$$

Then $C_{11} - C_{12} = b/V$
With the previous form of the ϵ matrix the number of point symmetry operators is reduced to 8, whereas the following ϵ matrix reduces the number of point symmetry operators to 16:

$$\epsilon = \begin{bmatrix} \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & -2\delta \end{bmatrix}$$

$$E(\epsilon_1, \epsilon_2, \epsilon_3) = E(0,0,0) + 3V(C_{11} - C_{12})\delta^2 + \cdots = a + b\delta^2 + \cdots$$

and $(C_{11} - C_{12}) = b/3V$

Calculation of C_{44}

Monoclinic deformation, 4 point symmetry operators.

The ϵ matrix has the form:

$$\epsilon = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & x \\ 0 & x & 0 \end{bmatrix}$$

The energy expression is ($\delta = 2x$) (see Voigt convention and equation 8.38)

$$E(\epsilon_4) = E(0) + 1/2 \frac{\partial^2 E}{\partial \epsilon_4^2} \delta^2 + \cdots = E(0) + 2 \frac{\partial^2 E}{\partial \epsilon_4^2} x^2 + \cdots = a + bx^2 + \cdots$$

so that $C_{44} = b/2V$.

Calculation of C_{44}

Rhombohedral deformation, 12 point symmetry operators.

The ϵ matrix has the form:

$$\epsilon = \begin{bmatrix} 0 & x & x \\ x & 0 & x \\ x & x & 0 \end{bmatrix}$$

The energy expression is ($\delta = 2x, C_{45} = C_{46} = C_{56} = 0$)

$$E(\epsilon_4, \epsilon_5, \epsilon_6) = E(0) + 3/2\frac{\partial^2 E}{\partial \epsilon_4^2}\delta^2 + \cdots = E(0) + 6\frac{\partial^2 E}{\partial \epsilon_4^2}x^2 + \cdots = a + bx^2 + \cdots$$

so that $C_{44} = b/6V$.

Bulk modulus

The bulk modulus can be evaluated simply by varying the lattice constant, (1 in cubic systems) without the use of the ϵ matrix, and fitting the curve $E(V)$.
If the ϵ matrix is used, the relation between B and C_{ij} (cubic systems) must be taken into account:

$$B = (C_{11} + 2C_{12})/3$$

The ϵ matrix has the form:

$$\epsilon = \begin{bmatrix} \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$$

and the energy:

$$E(\epsilon) = E(\mathbf{0}) + 3/2\frac{\partial^2 E}{\partial \epsilon_1^2}\delta^2 + 3\frac{\partial^2 E}{\partial \epsilon_1 \partial \epsilon_2}\delta^2 = \quad (8.39)$$

$$= E(\mathbf{0}) + \frac{3V}{2}[C_{11} + 2C_{12}]\delta^2 \quad (8.40)$$

so that $B = \frac{2}{9V}b$

N.B. Conversion factors:
1 hartree $Å^{-3}$ = 4359.74812 GPa
1 GPa = 1 GN m^{-2} = 1 GJ m^{-3} = 10^{10} dyne cm^{-2} = 10^{-2} Mbar.

8.10 Spontaneous polarization through the Berry phase approach

The electronic phase of a system λ in the direction 1, $\varphi_{el}^{(\lambda,1)}$, can be written as:

$$\varphi_{el}^{(\lambda,1)} = \frac{1}{s2s3} \sum_{j2,j3} \sum_{j_1} \Delta\varphi_{j_1,j_2,j_3}^{(\lambda,1)}(\mathbf{k}) \tag{8.41}$$

The electronic contribution to the polarization of a system λ can be written as :

$$\mathbf{P}_{el}^{(\lambda)} = \frac{1}{\Omega^{(\lambda)}} \left(B^{(\lambda)}\right)^{-1} \varphi_{el}^{(\lambda)} \tag{8.42}$$

Where $(B^{(\lambda)})^{-1}$ is the reciprocal lattice vectors components inverse matrix and $\varphi_{el}^{(\lambda)}$ the electronic phase difference vector of a system λ (which components are $\varphi_{el}^{(\lambda,i)}$). The nuclear contribution to the polarization of a system λ, $\mathrm{P}_{nuc}^{(\lambda)}$ can also be written as:

$$\mathbf{P}_{nuc}^{(\lambda)} = \frac{1}{\Omega^{(\lambda)}} \sum_{A} \mathbf{R}_{A}^{(\lambda)} \cdot Z_A \tag{8.43}$$

where $\mathbf{R}_{A}^{(\lambda)}$ and Z_A are the position vector and the nuclear charge of the atom A respectively of the system λ. The total polarization is the sum of these two contributions and can be written as

$$\mathbf{P}_{tot}^{(\lambda)} = \mathbf{P}_{nuc}^{(\lambda)} + \mathbf{P}_{el}^{(\lambda)} \tag{8.44}$$

The spontaneous polarization is the difference between the systems $\lambda = 1$ and $\lambda = 0$

$$\mathbf{P} = \mathbf{P}_{tot}^{(\lambda)} - \mathbf{P}_{tot}^{(\lambda)} \tag{8.45}$$

Spontaneous polarization through the localized crystalline orbitals approach

The electronic contribution to the polarization of a system λ, $\mathrm{P}_{el}^{(\lambda)}$, can be written as

$$\mathbf{P}_{el}^{(\lambda)} = \frac{e}{\Omega^{(\lambda)}} \sum_{\mu} \langle \mathbf{r}_\mu \rangle \tag{8.46}$$

Where $\langle \mathbf{r}_\mu \rangle$ is the centroid of the Wannier function μ.
The nuclear contribution to the polarization of a system λ, $\mathrm{P}_{nuc}^{(\lambda)}$ can also be written as

$$\mathbf{P}_{nuc}^{(\lambda)} = \frac{1}{\Omega^{(\lambda)}} \sum_{A} \mathbf{R}_{A} \cdot Z_A \tag{8.47}$$

where $\mathbf{R}_A$ and Z_A are the position vector and the nuclear charge of the atom A respectively. The total polarization is the sum of these two contributions and can be written as

$$\mathbf{P}_{tot}^{(\lambda)} = \mathbf{P}_{nuc}^{(\lambda)} + \mathbf{P}_{el}^{(\lambda)} \tag{8.48}$$

The spontaneous polarization is the difference between the both systems $\lambda = 1$ and $\lambda = 0$:

$$\mathbf{P} = \mathbf{P}_{tot}^{(1)} - \mathbf{P}_{tot}^{(2)} \tag{8.49}$$

To calculate the spontaneous polarization, a preliminary run is needed for each of the two systems $\lambda = 1$ and $\lambda = 0$. Then a third run with the keyword `SPOLWF` gives the difference of polarization between systems $\lambda = 1$ and $\lambda = 0$.

8.11 Piezoelectricity through the Berry phase approach

The electronic phase vector of a system λ, is given by (2.1). The nuclear phase vector of a system λ, $\varphi^{(\lambda)}_{nuc}$, can be written as

$$\varphi^{(\lambda)}_{nuc} = \Omega^{(\lambda)}\ B^{(\lambda)}\mathbf{P}^{(\lambda)}_{nuc} \tag{8.50}$$

Where $B^{(\lambda)}$ reciprocal lattice vectors components matrix.The last equation can be simplified thanks to (8.43):

$$\varphi^{(\lambda)}_{nuc} = B^{(\lambda)} \sum_A \mathbf{R}^{(\lambda)}_A \cdot Z_A \tag{8.51}$$

So the phase vector of a system λ, $\varphi^{(\lambda)}$ is:

$$\varphi^{(\lambda)} = \varphi^{(\lambda)}_{nuc} + \varphi^{(\lambda)}_{el} \tag{8.52}$$

The proper piezoelectric constants can be obtained by:

$$\tilde{e}_{ijk} = -\frac{1}{2\pi}\frac{1}{\Omega}\sum_\alpha \frac{d\varphi_\alpha}{d\epsilon_{jk}}\ a_{\alpha,i} \tag{8.53}$$

Where φ_α is projection of the phase φ along the α direction and $a_{\alpha,i}$ is the component of a lattice vector a_α along the cartesian axis i . To obtain the improper piezoelectric constants, the following correction must done:

$$e_{ijk} = \tilde{e}_{ijk} + \delta_{ij}P_k - \delta_{jk}P_i \tag{8.54}$$

In the piezoelectric constants calculations the $\frac{d\varphi_\alpha}{d\epsilon_{jk}}$ term is evaluated numerically. The calculated term is:

$$\frac{d\varphi_\alpha}{d\epsilon_{jk}} \simeq \frac{\Delta\varphi_\alpha}{\Delta\epsilon_{jk}} = \frac{\varphi^{(1)}_\alpha - \varphi^{(0)}_\alpha}{\epsilon^{(1)}_{jk} - \epsilon^{(0)}_{jk}} \tag{8.55}$$

Piezoelectricity through the localized crystalline orbitals approach

The electronic phase vector of a system λ, is given by:

$$\varphi^{(\lambda)}_{el} = \Omega^{(\lambda)}\ B^{(\lambda)}\mathbf{P}^{(\lambda)}_{el} \tag{8.56}$$

Where $B^{(\lambda)}$ reciprocal lattice vectors components matrix. The nuclear phase vector of a system λ, $\varphi^{(\lambda)}_{nuc}$, can be written as

$$\varphi^{(\lambda)}_{nuc} = \Omega^{(\lambda)}\ B^{(\lambda)}\mathbf{P}^{(\lambda)}_{nuc} \tag{8.57}$$

The last equation can be simplified thanks to 8.43:

$$\varphi^{(\lambda)}_{nuc} = B^{(\lambda)} \sum_A \mathbf{R}^{(\lambda)}_A \cdot Z_A \tag{8.58}$$

So the phase vector of a system λ, $\varphi^{(\lambda)}$ is:

$$\varphi^{(\lambda)} = \varphi^{(\lambda)}_{nuc} + \varphi^{(\lambda)}_{el} \tag{8.59}$$

The proper piezoelectric constants can be obtained by:

$$\tilde{e}_{ijk} = -\frac{1}{2\pi}\frac{1}{\Omega}\sum_\alpha \frac{d\varphi_\alpha}{d\epsilon_{jk}}\ a_{\alpha,i} \tag{8.60}$$

Where φ_α is projection of the phase φ along the α direction and $a_{\alpha,i}$ is the component of a lattice vector a_α along the cartesian axis i . To obtain the improper piezoelectric constants, the following correction must done:

$$e_{ijk} = \tilde{e}_{ijk} + \delta_{ij}P_k - \delta_{jk}P_i \tag{8.61}$$

In the piezoelectric constants calculations the $\frac{d\varphi_\alpha}{d\epsilon_{jk}}$ term is evaluated numerically. The calculated term is:

$$\frac{d\varphi_\alpha}{d\epsilon_{jk}} \simeq \frac{\Delta\varphi_\alpha}{\Delta\epsilon_{jk}} = \frac{\varphi_\alpha^{(1)} - \varphi_\alpha^{(0)}}{\epsilon_{jk}^{(1)} - \epsilon_{jk}^{(0)}} \tag{8.62}$$

Appendix A

Symmetry groups

A.1 Labels and symbols of the space groups

The labels are according to the International Tables for Crystallography [13]. The symbols are derived by the standard SHORT symbols, as shown in the following examples:

Symbol		Input to CRYSTAL
P $\bar{6}$ 2 m	→	`P␣-6␣2␣M` ;
P 6_3 m	→	`P␣63␣M`.

For the groups 221-230 the symbols are according to the 1952 edition of the International Tables, *not* to the 1982 edition. The difference involves the 3 axis: 3 (1952 edition); $\bar{3}$ (1982 edition) (Example group 221: 1952 ed. → P m 3 m ; 1982 ed. → P m $\bar{3}$ m)

IGR	symbol
Triclinic lattices	
1	$P1$
2	$P\bar{1}$
Monoclinic lattices	
3	$P2$
4	$P2_1$
5	$C2$
6	Pm
7	Pc
8	Cm
9	Cc
10	$P2/m$
11	$P2_1/m$
12	$C2/m$
13	$P2/c$
14	$P2_1/c$
15	$C2/c$
Orthorhombic lattices	
16	$P222$
17	$P222_1$
18	$P2_12_12$
19	$P2_12_12_1$
20	$C222_1$
21	$C222$
22	$F222$
23	$I222$
24	$I2_12_12_1$
25	$Pmm2$
26	$Pmc2_1$
27	$Pcc2$
28	$Pma2$
29	$Pca2_1$
30	$Pnc2$
31	$Pmn2_1$
32	$Pba2$
33	$Pna2_1$
34	$Pnn2$
35	$Cmm2$
36	$Cmc2_1$

IGR	symbol
37	$Ccc2$
38	$Amm2$
39	$Abm2$
40	$Ama2$
41	$Aba2$
42	$Fmm2$
43	$Fdd2$
44	$Imm2$
45	$Iba2$
46	$Ima2$
47	$Pmmm$
48	$Pnnn$
49	$Pccm$
50	$Pban$
51	$Pmma$
52	$Pnna$
53	$Pmna$
54	$Pcca$
55	$Pbam$
56	$Pccn$
57	$Pbcm$
58	$Pnnm$
59	$Pmmn$
60	$Pbcn$
61	$Pbca$
62	$Pnma$
63	$Cmcm$
64	$Cmca$
65	$Cmmm$
66	$Cccm$
67	$Cmma$
68	$Ccca$
69	$Fmmm$
70	$Fddd$
71	$Immm$
72	$Ibam$
73	$Ibca$
74	$Imma$

IGR	symbol
Tetragonal lattices	
75	$P4$
76	$P4_1$
77	$P4_2$
78	$P4_3$
79	$I4$
80	$I4_1$
81	$P\bar{4}$
82	$I\bar{4}$
83	$P4/m$
84	$P4_2/m$
85	$P4/n$
86	$P4_2/n$
87	$I4/m$
88	$I4_1/a$
89	$P422$
90	$P42_12$
91	$P4_122$
92	$P4_12_12$
93	$P4_222$
94	$P4_22_12$
95	$P4_322$
96	$P4_32_12$
97	$I422$
98	$I4_122$
99	$P4mm$
100	$P4bm$
101	$P4_2cm$
102	$P4_2nm$
103	$P4cc$
104	$P4nc$
105	$P4_2mc$
106	$P4_2bc$
107	$I4mm$
108	$I4cm$
109	$I4_1md$
110	$I4_1cd$
111	$P\bar{4}2m$
112	$P\bar{4}2c$

IGR	symbol
113	$P\bar{4}2_1m$
114	$P\bar{4}2_1c$
115	$P\bar{4}m2$
116	$P\bar{4}c2$
117	$P\bar{4}b2$
118	$P\bar{4}n2$
119	$I\bar{4}m2$
120	$I\bar{4}c2$
121	$I\bar{4}2m$
122	$I\bar{4}2d$
123	$P4/mmm$
124	$P4/mcc$
125	$P4/nbm$
126	$P4/nnc$
127	$P4/mbm$
128	$P4/mnc$
129	$P4/nmm$
130	$P4/ncc$
131	$P4_2/mmc$
132	$P4_2/mcm$
133	$P4_2/nbc$
134	$P4_2/nnm$
135	$P4_2/mbc$
136	$P4_2/mnm$
137	$P4_2/nmc$
138	$P4_2/ncm$
139	$I4/mmm$
140	$I4/mcm$
141	$I4_1/amd$
142	$I4_1/acd$
Trigonal lattices	
143	$P3$
144	$P3_1$
145	$P3_2$
146	$R3$
147	$P\bar{3}$
148	$R\bar{3}$
149	$P312$
150	$P321$
151	$P3_112$
152	$P3_121$
153	$P3_212$
154	$P3_221$
155	$R32$
156	$P3m1$
157	$P31m$
158	$P3c1$
159	$P31c$
160	$R3m$
161	$R3c$
162	$P\bar{3}1m$
163	$P\bar{3}1c$
164	$P\bar{3}m1$
165	$P\bar{3}c1$
166	$R\bar{3}m$
167	$R\bar{3}c$
Hexagonal lattices	
168	$P6$
169	$P6_1$
170	$P6_5$
171	$P6_2$
172	$P6_4$
173	$P6_3$
174	$P\bar{6}$
175	$P6/m$
176	$P6_3/m$
177	$P622$
178	$P6_122$
179	$P6_522$
180	$P6_222$
181	$P6_422$
182	$P6_322$
183	$P6mm$
184	$P6cc$
185	$P6_3cm$
186	$P6_3mc$
187	$P\bar{6}m2$
188	$P\bar{6}c2$
189	$P\bar{6}2m$
190	$P\bar{6}2c$
191	$P6/mmm$
192	$P6/mcc$
193	$P6_3/mcm$
194	$P6_3/mmc$
195	$P23$
Cubic lattices	
196	$F23$
197	$I23$
198	$P2_13$
199	$I2_13$
200	$Pm3$
201	$Pn3$
202	$Fm3$
203	$Fd3$
204	$Im3$
205	$Pa3$
206	$Ia3$
207	$P432$
208	$P4_232$
209	$F432$
210	$F4_132$
211	$I432$
212	$P4_332$
213	$P4_132$
214	$I4_132$
215	$P\bar{4}3m$
216	$F\bar{4}3m$
217	$I\bar{4}3m$
218	$P\bar{4}3n$
219	$F\bar{4}3c$
220	$I\bar{4}3d$
221	$Pm\bar{3}m$
222	$Pn\bar{3}n$
223	$Pm\bar{3}n$
224	$Pn\bar{3}m$
225	$Fm\bar{3}m$
226	$Fm\bar{3}c$
227	$Fd\bar{3}m$
228	$Fd\bar{3}c$
229	$Im\bar{3}m$
230	$Ia\bar{3}d$

A.2 Labels of the layer groups (slabs)

The available layer groups belong to a subset of the 230 space groups. Therefore they can be identified by the corresponding space group.
The first column gives the label to be used in the input card (IGR variable).
The second column gives the Hermann-Mauguin symbol of the corresponding space group (generally the short one; the full symbol is adopted when the same short symbol could refer to different settings). The third column gives the Schoenflies symbol. The fourth column the number of the corresponding space group, according to the International Tables for Crystallography. The number of the space group is written in parentheses when the orientation of the symmetry operators does not correspond to the first setting in the I. T.

IGR	Hermann Mauguin	Schoenflies	N	IGR	Hermann Mauguin	Schoenflies	N
	Oblique lattices (P)			41	$Pbam$	D_{2h}^9	55
1	$P1$	C_1^1	1	42	$Pmaa$	D_{2h}^3	(49)
2	$P\bar{1}$	C_i^1	2	43	$Pman$	D_{2h}^7	(53)
3	$P112$	C_2^1	(3)	44	$Pbma$	D_{2h}^1	(57)
4	$P11m$	C_s^1	(6)	45	$Pbaa$	D_{2h}^8	(54)
5	$P11a$	C_s^2	(7)	46	$Pban$	D_{2h}^4	50
6	$P112/m$	C_{2h}^1	(10)	47	$Cmmm$	D_{2h}^9	65
7	$P112/a$	C_{2h}^4	(13)	48	$Cmma$	D_{2h}^2	67
	Rectangular lattices (P or C)				Square lattices (P)		
8	$P211$	C_2^1	(3)	49	$P4$	C_4^1	75
9	$P2_111$	C_2^2	(4)	50	$P\bar{4}$	S_4^1	81
10	$C211$	C_2^3	(5)	51	$P4/m$	C_{4h}^1	83
11	$Pm11$	C_s^1	(6)	52	$P4/n$	C_{4h}^3	85
12	$Pb11$	C_s^2	(7)	53	$P422$	D_4^1	89
13	$Cm11$	C_s^3	(8)	54	$P42_12$	D_4^2	90
14	$P2/m11$	C_{2h}^1	(10)	55	$P4mm$	C_{4v}^1	99
15	$P2_1/m11$	C_{2h}^2	(11)	56	$P4bm$	C_{4v}^2	100
16	$C2/m11$	C_{2h}^3	(12)	57	$P\bar{4}2m$	D_{2d}^1	111
17	$P2/b11$	C_{2h}^4	(13)	58	$P\bar{4}2_1m$	D_{2d}^3	113
18	$P2/b11$	C_{2h}^5	(14)	59	$P\bar{4}m2$	D_{2d}^5	115
19	$P222$	D_2^1	16	60	$P\bar{4}b2$	D_{2d}^7	117
20	$P222$	D_2^2	(17)	61	$P4/mmm$	D_{4h}^1	123
21	$P2_12_12$	D_2^3	18	62	$P4/nbm$	D_{4h}^3	125
22	$C222$	D_2^6	21	63	$P4/mbm$	D_{4h}^5	127
23	$Pmm2$	C_{2v}^1	25	64	$P4/nmm$	D_{4h}^7	129
24	$Pma2$	C_{2v}^4	28		Hexagonal lattices (P)		
25	$Pba2$	C_{2v}^8	32	65	$P3$	C_3^1	143
26	$Cmm2$	C_{2v}^1	35	66	$P\bar{3}$	C_{3i}^1	147
27	$P2mm$	C_{2v}^1	(25)	67	$P312$	D_3^1	149
28	$P2_1am$	C_{2v}^2	(26)	68	$P321$	D_3^2	150
29	$P2_1ma$	C_{2v}^2	(26)	69	$P3m1$	C_{3v}^1	156
30	$P2mb$	C_{2v}^4	(28)	70	$P31m$	C_{3v}^2	157
31	$P2_1mn$	C_{2v}^7	(31)	71	$P\bar{3}1m$	D_{3d}^1	162
32	$P2aa$	C_{2v}^3	(27)	72	$P\bar{3}m1$	D_{3d}^3	164
33	$P2_1ab$	C_{2v}^5	(29)	73	$P6$	C_6^1	168
34	$P2an$	C_{2v}^6	(30)	74	$P\bar{6}$	C_{3h}^1	174
35	$C2mm$	C_{2v}^1	(38)	75	$P6/m$	C_{6h}^1	175
36	$C2mb$	C_{2v}^5	(39)	76	$P622$	D_6^1	177
37	$Pmmm$	D_{2h}^1	47	77	$P6mm$	C_{6v}^1	183
38	$Pmam$	D_{2h}^5	(51)	78	$P\bar{6}m2$	D_{3h}^1	187
39	$Pmma$	D_{2h}^5	51	79	$P\bar{6}2m$	D_{3h}^3	189
40	$Pmmn$	D_{2h}^3	59	80	$P6/mmm$	D_{6h}^1	191

A.3 Labels of the rod groups (polymers)

The available rod groups belong to a subset of the 230 space groups; the symmetry operators are generated for the space groups (principal axis z) and then rotated by 90° through y, to have the polymer axis along x (CRYSTAL convention).
In the table, the first column gives the label to be used in the input card for identifying the rod group (IGR variable).
The second column gives the "polymer" symbol, according to the the following convention: x is the first symmetry direction, y the second.
The third column gives the Schoenflies symbol.
The fourth column gives the Hermann-Mauguin symbol (generally the short one; the full symbol is adopted when the same short symbol could refer to different settings) of the corresponding space group (principal axis z).
The fifth column gives the number of the corresponding space group, according to the International Tables for Crystallography; this number is written in parentheses when the orientation of the symmetry operators does not correspond to the first setting in the I. T.

IGR	"Polymer" symbol (x direction)	Schoenflies	Hermann Mauguin (z direction)	Number of space group
1	$P1$	C_1^1	$P1$	1
2	$P\bar{1}$	C_i^1	$P\bar{1}$	2
3	$P211$	C_2^1	$P112$	(3)
4	$P2_111$	C_2^2	$P112_1$	(4)
5	$P121$	C_2^1	$P121$	(3)
6	$P112$	C_2^1	$P211$	(3)
7	$Pm11$	C_s^1	$P11m$	(6)
8	$P1m1$	C_s^1	$P1m1$	(6)
9	$P1a1$	C_s^2	$P1c1$	(7)
10	$P11m$	C_s^1	$Pm11$	(6)
11	$P11a$	C_s^2	$Pc11$	(7)
12	$P2/m11$	C_{2h}^1	$P112/m$	(10)
13	$P2_1/m11$	C_{2h}^2	$P112_1/m$	(11)
14	$P12/m1$	C_{2h}^1	$P12/m1$	(10)
15	$P12/a1$	C_{2h}^4	$P12/c1$	(13)
16	$P112/m$	C_{2h}^1	$P2/m11$	(10)
17	$P112/a$	C_{2h}^4	$P2/c11$	(13)
18	$P222$	D_2^1	$P222$	16
19	$P2_122$	D_2^2	$P222_1$	17
20	$P2mm$	C_{2v}^1	$Pmm2$	25
21	$P2_1am$	C_{2v}^2	$Pmc2_1$	26
22	$P2_1ma$	C_{2v}^2	$Pcm2_1$	(26)
23	$P2aa$	C_{2v}^3	$Pcc2$	27
24	$Pm2m$	C_{2v}^1	$Pm2m$	(25)
25	$Pm2a$	C_{2v}^4	$Pc2m$	(28)
26	$Pmm2$	C_{2v}^1	$P2mm$	(25)
27	$Pma2$	C_{2v}^4	$P2cm$	(28)
28	$Pmmm$	D_{2h}^1	$Pmmm$	47
29	$P2/m2/a2/a$	D_{2h}^3	$Pccm$	49
30	$P2_1/m2/m2/a$	D_{2h}^5	$Pcmm$	(51)
31	$P2_1/m2/a2/m$	D_{2h}^5	$Pmcm$	(51)

IGR	"Polymer" symbol (x direction)	Schoenflies	Hermann Mauguin (z direction)	Number of space group
32	$P4$	C_4^1	$P4$	75
33	$P4_1$	C_4^2	$P4_1$	76
34	$P4_2$	C_4^3	$P4_2$	77
35	$P4_3$	C_4^4	$P4_3$	78
36	$P\bar{4}$	S_4^1	$P\bar{4}$	81
37	$P4/m$	C_{4h}^1	$P4/m$	83
38	$P4_2/m$	C_{4h}^2	$P4_2/m$	84
39	$P422$	D_4^1	$P422$	89
40	$P4_122$	D_4^3	$P4_122$	91
41	$P4_222$	D_4^5	$P4_222$	93
42	$P4_322$	D_4^7	$P4_322$	95
43	$P4mm$	C_{4v}^1	$P4mm$	99
44	$P4_2am$	C_{4v}^3	$P4_2cm$	101
45	$P4aa$	C_{4v}^5	$P4cc$	103
46	$P4_2ma$	C_{4v}^7	$P4_2mc$	105
47	$P\bar{4}2m$	D_{2d}^1	$P\bar{4}2m$	111
48	$P\bar{4}2a$	D_{2d}^2	$P\bar{4}2c$	112
49	$P\bar{4}m2$	D_{2d}^5	$P\bar{4}m2$	115
50	$P\bar{4}a2$	D_{2d}^6	$P\bar{4}c2$	116
51	$P4/mmm$	D_{4h}^1	$P4/mmm$	123
52	$P4/m2/a2/a$	D_{4h}^2	$P4/mcc$	124
53	$P4_2/m2/m2/a$	D_{4h}^9	$P4_2/mmc$	131
54	$P4_2/m2/a2/m$	D_{4h}^{10}	$P4_2/mcm$	132
55	$P3$	C_3^1	$P3$	143
56	$P3_1$	C_3^2	$P3_1$	144
57	$P3_2$	C_3^3	$P3_2$	145
58	$P\bar{3}$	C_{3i}^1	$P\bar{3}$	147
59	$P312$	D_3^1	$P312$	149
60	$P3_112$	D_3^3	$P3_112$	151
61	$P3_212$	D_3^5	$P3_212$	153
62	$P321$	D_3^2	$P321$	150
63	$P3_121$	D_3^4	$P3_121$	152
64	$P3_221$	D_3^6	$P3_221$	154
65	$P3m1$	C_{3v}^1	$P3m1$	156
66	$P3a1$	C_{3v}^3	$P3c1$	158
67	$P31m$	C_{3v}^2	$P31m$	157
68	$P31a$	C_{3v}^4	$P31c$	159
69	$P\bar{3}1m$	D_{3d}^1	$P\bar{3}1m$	162
70	$P\bar{3}1a$	D_{3d}^2	$P\bar{3}1c$	163
71	$P\bar{3}m1$	D_{3d}^3	$P\bar{3}m1$	164
72	$P\bar{3}a1$	D_{3d}^4	$P\bar{3}c1$	165

IGR	"Polymer" symbol (x direction)	Schoenflies	Hermann Mauguin (z direction)	Number of space group
73	$P6$	C_6^1	$P6$	168
74	$P6_1$	C_6^2	$P6_1$	169
75	$P6_5$	C_6^3	$P6_5$	170
76	$P6_2$	C_6^4	$P6_2$	171
77	$P6_4$	C_6^5	$P6_4$	172
78	$P6_3$	C_6^6	$P6_6$	173
79	$P\bar{6}$	C_{3h}^1	$P\bar{6}$	174
80	$P6/m$	C_{6h}^1	$P6/m$	175
81	$P6_3/m$	C_{6h}^2	$P6_3/m$	176
82	$P622$	D_6^1	$P622$	177
83	$P6_122$	D_6^2	$P6_122$	178
84	$P6_522$	D_6^3	$P6_522$	179
85	$P6_222$	D_6^4	$P6_222$	180
86	$P6_422$	D_6^5	$P6_422$	181
87	$P6_322$	D_6^6	$P6_322$	182
88	$P6mm$	C_{6v}^1	$P6mm$	183
89	$P6aa$	C_{6v}^2	$P6cc$	184
90	$P6_3am$	C_{6v}^3	$P6_3cm$	185
91	$P6_3ma$	C_{6v}^4	$P6_3mc$	186
92	$P\bar{6}m2$	D_{3h}^1	$P\bar{6}m2$	187
93	$P\bar{6}a2$	D_{3h}^2	$P\bar{6}c2$	188
94	$P\bar{6}2m$	D_{3h}^3	$P\bar{6}2m$	189
95	$P\bar{6}2a$	D_{3h}^4	$P\bar{6}2c$	190
96	$P6/mmm$	D_{6h}^1	$P6/mmm$	191
97	$P6/m2/a2/a$	D_{6h}^2	$P6/mcc$	192
98	$P6_3/m2/a2/m$	D_{6h}^3	$P6_3/mcm$	193
99	$P6_3/m2/m2/a$	D_{6h}^4	$P6_3/mmc$	194

A.4 Labels of the point groups (molecules)

The centre of symmetry is supposed to be at the origin; for the rotation groups the principal axis is z.
The first column gives the label to be used in the input card for identifying the point group (IGR variable). The second column gives the short Hermann-Mauguin symbol. The third column gives the Schoenflies symbol; for the C_2 , C_{2h} and C_s groups the C_2 direction or the direction orthogonal to the plane is indicated. The fourth column gives the number of pure rotations for molecules (σ).

IGR	Hermann Mauguin	Schoenflies		σ
1	1	C_1		1
2	$\bar{1}$	C_i		1
3	2 (x)	C_2 (x)		2
4	2 (y)	C_2 (y)		2
5	2 (z)	C_2 (z)		2
6	m (x)	C_s (x)		1
7	m (y)	C_s (y)		1
8	m (z)	C_s (z)		1
9	2/m (x)	C_{2h} (x)		2
10	2/m (y)	C_{2h} (y)		2
11	2/m (z)	C_{2h} (z)		2
12	222	D_2		4
13	2mm	C_{2v} (x)		2
14	m2m	C_{2v} (y)		2
15	mm2	C_{2v} (z)		2
16	mmm	D_{2h}		4
17	4	C_4		4
18	$\bar{4}$	S_4		2
19	4/m	C_{4h}		4
20	422	D_4		8
21	4mm	C_{4v}		4
22	$\bar{4}$2m	D_{2d}	(σ_v planes along x+y and x-y)	4
23	$\bar{4}$m2	D_{2d}	(σ_v planes along x and y)	4
24	4/mmm	D_{4h}		8
25	3	C_3		3
26	$\bar{3}$	C_{3i}		3
27	321	D_3	(one C_2 axis along y)	6
28	312	D_3	(one C_2 axis along x)	6
29	3m1	C_{3v}	(one σ_v plane along x)	3
30	31m	C_{3v}	(one σ_v plane along y)	3
31	$\bar{3}$m1	D_{3d}	(one σ_d plane along x)	6
32	$\bar{3}$1m	D_{3d}	(one σ_d plane along y)	6
33	6	C_6		6
34	$\bar{6}$	C_{3h}		3
35	6/m	C_{6h}		6
36	622	D_6		12
37	6mm	C_{6v}		6
38	$\bar{6}$m2	D_{3h}	(one C_2 axis along x)	6
39	$\bar{6}$2m	D_{3h}	(one C_2 axis along y)	6
40	6/mmm	D_{6h}		12
41	23	T		12
42	m$\bar{3}$	T_h		12
43	432	O		24
44	$\bar{4}$3m	T_d		12
45	m$\bar{3}$m	O_h		24

A.5 From conventional to primitive cells: transforming matrices

The matrices describing the transformations from conventional (given as input) to primitive (internally used by CRYSTAL) cells of Bravais lattices are coded in CRYSTAL. A point called $\mathbf{x}$ in the *direct lattice* has $\mathbf{x}_P$ coordinates in a primitive cell and $\mathbf{x}_C$ coordinates in a conventional cell. The relation between $\mathbf{x}_P$ and $\mathbf{x}_C$ is the following:

$$W\mathbf{x}_P = \mathbf{x}_C \tag{A.1}$$

Likewise, for a point in the *reciprocal space* the following equation holds:

$$\tilde{W}^{-1}\mathbf{x}^*_P = \mathbf{x}^*_C \tag{A.2}$$

The W transforming matrices adopted in CRYSTAL, and reported below, satisfy the following relation between the two metric tensors $\mathbf{G}_P$ and $\mathbf{G}_C$:

$$\mathbf{G}_P = W\mathbf{G}_C\tilde{W} \tag{A.3}$$

The values of the elements of the metric tensors $\mathbf{G}_P$ and $\mathbf{G}_C$ agree with those displayed in Table 5.1 of the International Tables of Crystallography (1992 edition).

$$
P \rightarrow A \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\bar{1}}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}
\quad P \rightarrow B \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{\bar{1}}{2} & 0 & \frac{1}{2} \end{pmatrix}
\quad A \rightarrow P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}
\quad B \rightarrow P \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}
$$

$$
P \rightarrow C \begin{pmatrix} \frac{1}{2} & \frac{\bar{1}}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}
\quad P \rightarrow F \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}
\quad C \rightarrow P \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\quad F \rightarrow P \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}
$$

$$
P \rightarrow I \begin{pmatrix} \frac{\bar{1}}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{\bar{1}}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{\bar{1}}{2} \end{pmatrix}
\quad R \rightarrow H \begin{pmatrix} \frac{2}{3} & \frac{\bar{1}}{3} & \frac{\bar{1}}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}
\quad I \rightarrow P \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}
\quad H \rightarrow R \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}
$$

Table A.1: W matrices for the transformation from conventional to primitive and from primitive to conventional cells. P stands for primitive, A, B and C for A-, B- and C-face centred, I for body centred, F for all-face centred, R for primitive rhombohedral ('rhombohedral axes') and H for rhombohedrally centred ('hexagonal axes') cell (Table 5.1, ref. [13]).

Appendix B

Summary of input keywords

All the keywords are entered with an A format; the keywords must be typed left-justified, with no leading blanks. The input is not case sensitive.

Geometry (Input block 1)

Symmetry information			
ATOMSYMM	printing of point symmetry at the atomic positions	28	–
MAKESAED	printing of symmetry allowed elastic distortions (SAED)	36	–
PRSYMDIR	printing of displacement directions allowed by symmetry.	40	–
SYMMDIR	printing of symmetry allowed geom opt directions	45	–
SYMMOPS	printing of point symmetry operators	46	–
TENSOR	tensor of physical properties	46	I

Symmetry information and control			
BREAKSYM	allow symmetry reduction following geometry modifications	29	–
KEEPSYMM	maintain symmetry following geometry modifications	36	–
MODISYMM	removal of selected symmetry operators	36	I
PURIFY	cleans atomic positions so that they are fully consistent with the group	40	–
SYMMREMO	removal of all symmetry operators	46	–
TRASREMO	removal of symmetry operators with translational components	46	–

Modifications without reduction of symmetry			
ATOMORDE	reordering of atoms in molecular crystals	26	–
NOSHIFT	no shift of the origin to minimize the number of symmops with translational components before generating supercell	39	–
ORIGIN	shift of the origin to minimize the number of symmetry operators with translational components	39	–
PRIMITIV	crystallographic cell forced to be the primitive cell	40	–
SLABINFO	definition of a new cell, with $xy \parallel$ to a given plane	43	I

Atoms and cell manipulation (possible symmetry reduction (BREAKSYMM)

ATOMDISP	displacement of atoms	26	I
ATOMINSE	addition of atoms	26	I
ATOMREMO	removal of atoms	27	I
ATOMROT	rotation of groups of atoms	27	I
ATOMSUBS	substitution of atoms	28	I
ELASTIC	distortion of the lattice	31	I
POINTCHG	point charges input	39	I
USESAED	given symmetry allowed elastic distortions, reads δ	46	I
SUPERCEL	generation of supercell - input refers to primitive cell	44	I
SUPERCON	generation of supercell - input refers to conventional cell	44	I

From crystals to slabs

SLABCUT	generation of a slab parallel to a given plane (3D→2D)	42	I

From periodic structure to clusters

CLUSTER	cutting of a cluster from a periodic structure (3D→0D)	29	I
HYDROSUB	border atoms substituted with hydrogens (0D→0D)	35	I

Molecular crystals

MOLECULE	extraction of a set of molecules from a molecular crystal (3D→0D)	37	I
MOLEXP	variation of lattice parameters at constant symmetry and molecular geometry (3D→3D)	38	I
MOLSPLIT	periodic structure of non interacting molecules (3D→3D)	38	–
RAYCOV	modification of atomic covalent radii	40	I

BSSE correction

MOLEBSSE	counterpoise method for molecules (molecular crystals only) (3D→0D)	36	I
ATOMBSSE	counterpoise method for atoms (3D→0D)	26	I

Auxiliary and control keywords

ANGSTROM	sets inputs unit to Ångstrom	25	–
BOHR	sets input units to bohr	28	–
BOHRANGS	input bohr to Å conversion factor (0.5291772083 default value)	28	I
BOHRCR98	bohr to Å conversion factor is set to 0.529177 (CRYSTAL98 value)		–
END/ENDG	terminate processing of geometry input		–
FRACTION	sets input unit to fractional	35	–
NEIGHBOR	number of neighbours in geometry analysis	38	I
PARAMPRT	printing of parameters controlling dimensions of static allocation arrays	39	–
PRINTCHG	printing of point charges coordinates in geometry output	39	
PRINTOUT	setting of printing options by keywords	40	–
SETINF	setting of inf array options	42	I
SETPRINT	setting of printing options	42	I
STOP	execution stops immediately	43	–
TESTGEOM	stop after checking the geometry input	46	–

Output of data on external units

COORPRT	coordinates of all the atoms in the cell	30	–
EXTPRT	generation of file as CRYSTAL input	32	–
FINDSYM	generation of file as FINDSYM input	35	–
MOLDRAW	generation of file for the program MOLDRAW	36	–
STRUCPRT	cell parameters and coordinates of all the atoms in the cell	43	–

External electric field - modified Hamiltonian

FIELD	electric field applied along a periodic direction	32	I
FIELDCON	electric field applied along a non periodic direction	34	I

Geometry optimization

OPTGEOM	Geometry optimization	82	I

Type of optimization (default: atom coordinates)

FULLOPTG	full geometry optimization	–
CELLONLY	cell parameters optimization	–
INTREDUN	optimization in redundant internal coordinates	–
ITATOCEL	iterative optimization (atom/cell)	–
CVOLOPT	full geometry optimization at constant volume	–

Initial Hessian

HESGUESS	initial guess for the Hessian	I
HESSIDEN	initial guess for the Hessian - identity matrix	–
HESSMOD1	initial guess for the Hessian - model 1 (default)	–
HESSMOD2	initial guess for the Hessian - model 2	–

Convergence criteria modification

TOLDEG	RMS of the gradient [0.0003]	I
TOLDEX	RMS of the displacement [0.0012]	I
TOLDEE	energy difference between two steps [10^{-7}]	I
MAXCYCLE	max number of optimization steps	I

Optimization control

FRAGMENT	partial geometry optimization	I
RESTART	data from previous run	–
FINALRUN	Wf single point with optimized geometry	I

Gradient calculation control

NUMGRAD	numerical first derivatives	–

Printing options

PRINTFORCES	atomic gradients	–
PRINTHESS	Hessian	–
PRINTOPT	optimization procedure	–
PRINT	verbose printing	–

Frequencies at Γ

FREQCALC	Frequency at Γ - Harmonic calculation 4- [default]	98	I
ANALYSIS			–
[NOANALYSIS]			–
DIELISO			I
DIELTENS			I
FRAGMENT			I
INTENS			–
[NOINTENS]			–
ISOTOPES			I
[MODES]			–
NOMODES			–
NORMBORN			–
NUMDERIV			I
PRESSURE			I
PRINT			–
RESTART			–
SCANMODE			I
STEPSIZE			I
TEMPERAT			I
TESTFREQ			–
[USESYMM]			–
NOUSESYMM			–
END[FREQ]			–
ANHARM	Frequency at Γ - Anharmonic calculation	106	I
TESTANHA			–
KEEPSYMM			–
ISOTOPES			I
NOGUESS			–
END[ANHA]			–

Basis set input (Input block 2)

Symmetry control

ATOMSYMM	printing of point symmetry at the atomic positions	28	–

Basis set modification

CHEMOD	modification of the electronic configuration	47	I
GHOSTS	eliminates nuclei and electrons, leaving BS	49	I

Auxiliary and control keywords

CHARGED	allows non-neutral cell	47	–
NOPRINT	printing of basis set removed	49	–
PARAMPRT	printing of parameters controlling code dimensions	39	–
PRINTOUT	setting of printing options	40	I
SETINF	setting of inf array options	42	I
SETPRINT	setting of printing options	42	I
STOP	execution stops immediately	43	–
SYMMOPS	printing of point symmetry operators	46	–
END/ENDB	terminate processing of basis set definition keywords		–

Output of data on external units

GAUSS98	printing of an input file for the GAUSS94/98 package	49	–

General information, hamiltonian, SCF (Input block 3)

All DFT related keyword are collected under the heading "**DFT**", closed b **END[DFT]**

Single particle Hamiltonian

RHF			Restricted Closed Shell	74	–
UHF			Unrestricted Open Shell	80	–
DFT	DFT Hamiltonian			80	–
	SPIN		spin-polarized solution	59	–
	Choice of the exchange-correlation functionals				
	EXCHANGE		exchange functional	59	I
		LDA	Dirac-Slater [41] (LDA)		
		VBH	von Barth-Hedin [42] (LDA)		
		BECKE	Becke [43] (GGA)		
		PWGGA	Perdew-Wang 91 (GGA)		
		PBE	Perdew-Becke-Ernzerhof [44] (GGA)		
	CORRELAT		correlation functional	59	I
		VBH	von Barth-Hedin [42] (LDA)		
		PWGGA	Perdew-Wang 91 (GGA)		
		PBE	Perdew-Becke-Ernzerhof [44] (GGA)		
		PZ	Perdew-Zunger [45] (LDA)		
		PWLSD	Perdew-Wang 92 [46, 47, 48] (GGA)		
		VWN	Vosko,-Wilk-Nusair [49] (LDA)		
		P86	Perdew 86 [50] (LDA)		
		LYP	Lee-Yang-Parr [51] (GGA)		
	HYBRID		hybrid mixing	60	I
	NONLOCAL		local term parameterization	60	I
	B3PW		B3PW parameterization	60	–
	B3LYP		B3LYP parameterization	60	–
	Numerical accuracy control				
	[BECKE]		selection of Becke weights (default)		–
	SAVIN		selection of Savin weights		–
	RADIAL		definition of radial grid		I
	ANGULAR		definition of angular grid		I
	LGRID		"large" predefined grid		I
	XLGRID		"extra large" predefined grid		I
	TOLLDENS		density contribution screening 6		I
	TOLLGRID		grid points screening 14		I
	RADSAFE		safety radius for grid point screening		I
	BATCHPNT		grid point grouping for integration		I
	Atomic parameters control				
	RADIUS		customized atomic radius	65	I
	FCHARGE		customized formal atomic charge	65	I
	Auxiliary				
	PRINT		extended printing		
	END		close DFT input block		

Numerical accuracy and computational parameters control

BIPOLAR	Bipolar expansion of bielectronic integrals	58	I
BIPOSIZE	size of coulomb bipolar expansion buffer	58	I
EXCHSIZE	size of exchange bipolar expansion buffer	58	I
INTGPACK	Choice of integrals package [0]	70	I
NOBIPOLA	All bielectronic integrals computed exactly	73	–
POLEORDR	Maximum order of multipolar expansion [4]	73	I
TOLINTEG	Truncation criteria for bielectronic integrals [6 6 6 6 12]	80	I
TOLPSEUD	Pseudopotential tolerance [6]	80	I

Type of run

ATOMHF	Atomic wave functions	56	I
MPP	MPP execution (programmers only)	72	I
SCFDIR	SCF direct (mono+biel int computed)	74	–
NOMONDIR	SCF semidirect (mono on disk, biel computed)	73	–
EIGS	S(**k**) eigenvalues - basis set linear dependence check	65	–
FIXINDEX	Reference geometry to classify integrals	67	–

Integral file distribution

BIESPLIT	writing of bielectronic integrals in n files [$n = 1$],max=10	57	I
MONSPLIT	writing of mono-electronic integrals in n file [$n = 1$], max=10	72	I

Numerical accuracy control and convergence tools

ANDERSON	Fock matrix mixing	56	I
BROYDEN	Fock matrix mixing	58	I
FMIXING	Fock/KS matrix (cycle i and i-1) mixing [0]	69	I
LEVSHIFT	level shifter [no]	71	I
MAXCYCLE	maximum number of cycles [50]	72	I
SMEAR	Finite temperature smearing of the Fermi surface [no]	77	I
TOLDEE	convergence on total energy [5]	79	I
TOLDEP	convergence on density matrix [16]	79	I

Initial guess

EIGSHIFT	alteration of orbital occupation before SCF [no]	66	I
GUESSF	Fock/KS matrix from previous run	69	–
GUESSP	density matrix from a previous run	70	–
[**GUESSPAT**]	superposition of atomic densities	70	–

Spin-polarized system

ATOMSPIN	setting of atomic spin to compute atomic densities	57	I
BETALOCK	beta electrons locking	57	I
SPINLOCK	spin difference locking	78	I
SPINEDIT	editing of the spin density matrix	78	I

Auxiliary and control keywords

END	terminate processing of block3 input		–
KSYMMPRT	printing of Bloch functions symmetry analysis	71	–
NEIGHBOR	number of neighbours to analyse in **PPAN**	38	I
PARAMPRT	output of parameters controlling code dimensions	39	–
PRINTOUT	setting of printing options	40	I
NOSYMADA	No Symmetry Adapted Bloch Functions	73	–
SYMADAPT	Symmetry Adapted Bloch Functions (default)	79	–
SETINF	setting of inf array options	42	I
SETPRINT	setting of printing options	42	I
STOP	execution stops immediately	43	–
TESTPDIM	stop after symmetry analysis	79	–
TESTRUN	stop after integrals classification and disk storage estimate	79	–

Output of data on external units

NOFMWF	wave function formatted output not written in file fort.98.	73	–
SAVEWF	wave function data written every two SCF cycles	74	–

Post SCF calculations

POSTSCF	post-scf calculations when convergence criteria not satisfied	74	–
EXCHGENE	exchange energy evaluation (spin polarized only)	67	–
GRADCAL	analytical gradient of the energy	69	–
PPAN	population analysis at the end of the SCF [no]	74	

Properties

RDFMWF wave function data conversion formatted-binary (fort.98 → fort.9)

Preliminary calculations

NEWK	Eigenvectors calculation	133	I
NOSYMADA	No symmetry Adapted Bloch Functions	73	–
PATO	Density matrix as superposition of atomic (ionic) densities	134	I
PBAN	Band(s) projected density matrix (preliminary **NEWK**)	134	I
PGEOMW	Density matrix from geometrical weights (preliminary **NEWK**)	135	I
PDIDE	Energy range projected density matrix (preliminary **NEWK**)	135	I
PSCF	Restore SCF density matrix	140	–

Properties computed from the density matrix

ADFT	Atomic density functional correlation energy	111	I
BAND	Band structure	112	I
CLAS	Electrostatic potential maps (point multipoles approximation)	115	I
ECHG	Charge density maps and charge density gradient	119	I
ECH3	Charge density 3D maps	119	I
EDFT	Density functional correlation energy (HF wave function only)	120	I
POLI	Atom and shell multipoles evaluation	135	I
POTM	Electrostatic potential maps	138	I
POTC	Electrostatic properties	137	I
PPAN	Mulliken population analysis	74	
XFAC	X-ray structure factors	141	I

Properties computed from the density matrix (spin-polarized systems)

ANISOTRO	Hyperfine electron-nuclear spin tensor	111	I
ISOTROPIC	Hyperfine electron-nuclear spin interaction - Fermi contact	123	I
POLSPIN	Atomic spin density multipoles	136	I

Properties computed from eigenvectors (after keyword **NEWK**)

ANBD	Printing of principal AO component of selected CO	110	I
BWIDTH	Printing of bandwidth	114	I
DOSS	Density of states	118	I
EMDL	Electron momentum distribution - line	121	I
EMDP	Electron momentum distribution - plane maps	121	I
PROF	Compton profiles and related quantities	139	I

New properties

POLARI	Berry phase calculations	143	I
SPOLBP	Spontaneous polarization (Berry phase approach)	145	–
SPOLWF	Spontaneous polarization (localized CO approach)	146	–
PIEZOBP	Piezoelectricity (Berry phase approach) preliminary	142	–
PIEZOWF	Piezoelectricity (localized CO approach) - preliminary	142	–
LOCALWF	Localization of Wannier functions	124	I
DIEL	Optical dielectric constant	116	I

Auxiliary and control keywords

ANGSTROM	Set input unit of measure to Ångstrom	25	–
BASISSET	Printing of basis set, Fock/KS, overlap and density matrices	113	–
BOHR	Set input unit of measure to bohr	28	–
CHARGED	Non-neutral cell allowed (**PATO**)	47	–
END	Terminate processing of properties input keywords		–
FRACTION	Set input unit of measure to fractional	35	–
MAPNET	Generation of coordinates of grid points on a plane	130	I
NEIGHBOR	Number of neighbours to analyse in **PPAN**	38	I
PRINTOUT	Setting of printing options	40	I
RAYCOV	Modification of atomic covalent radii	40	I
SETINF	Setting of inf array options	42	I
SETPRINT	Setting of printing options	42	I
STOP	Execution stops immediately	43	–
SYMMOPS	Printing of point symmetry operators	46	–

Output of data on external units

ATOMIR	Coordinates of the irreducible atoms in the cell	112	–
ATOMSYMM	Printing of point symmetry at the atomic positions	28	–
COORPRT	Coordinates of all the atoms in the cell	30	–
CRYAPI_OUT	geometry, BS, direct lattice information	116	–
KNETOUT	Reciprocal lattice information, eigenvalues, eigenvectors	124 obsolete	- –
EXTPRT	Explicit structural/symmetry information	32	–
FMWF	Wave function formatted output. Section 5.2	122	–
INFOGUI	Generation of file with wf information for visualization	123	–
KNETOUT	Reciprocal lattice information + eigenvalues	124	–
MOLDRAW	generation of input file for the program MOLDRAW	36	–

Appendix C

Reciprocal lattice sampling

The keyword **KNETOUT** entered in the program ***properties*** writes the formatted file KIBZ.DAT. The structure of the file is as follows:

rec	data type	n. data	content
1	3I, F	3+9	ndf, nkf, iuhf, reciprocal lattice vectors cartesian components (a.u.)
2	I	3*nkf	oblique coordinates of the points in reciprocal lattice
3	I	nkf	k points flag: 0 (complex); 1 (real)
4	I	3x3x48	symmetry operators matrices
5	F	nkf	geometrical weight of k points
6	F	ntot	eigenvalues
7	F	ntot	weight eigenvalues at each k point

where:

ndf		number of basis set functions
nkf		number of k points (Monkhorst sampling)
iuhf		0 (Restricted calculation); 1 (Unrestricted calculation)
ntot	(nkf*ndf*(iuhf+1))	number of eigenvalues

The eigenvectors (in the AO basis) computed by ***properties*** (keyword **NEWK**, page 133), corresponding to the eigenvalues written in KIBZ.DAT, are written in fortran unit 8.
The reciprocal lattice vectors cartesian components and the oblique coordinates of the points in reciprocal lattice are printed when the input block 3, SCF input is processed. Printing of the other data may be obtained by setting the appropriate printing options (see keyword **PRINTOUT**, page 207):

keyword	input	information
KNETCOOR		reciprocal lattice sampling points coordinates
KWEIGHTS		geometrical weight of k points
EIGENVAL	n	eigenvalues at the first n k points
EIGENVEC	n	eigenvectors at the first n k points
EIGENALL		eigenvalues at all k points

Appendix D

Printing options

Extended printing can be obtained by entering the keywords **PRINTOUT** (page 40) or **SET-PRINT** (page 42).
In the **scf** (or **scfdir**) program the printing of quantities computed is done at each cycle if the corresponding LPRINT value is positive, only at the last cycle if the LPRINT value is negative. The LPRINT options to obtain intermediate information can be grouped as follows. The following table gives the correspondence between position number, quantity printed, and keyword.

crystal — Keyword — inp

- direct lattice - geometry information: 1 — *GLATTICE* –
- symmetry operators : 4, 2 — *SYMMOPS* –
- atomic functions basis set : 72 — *BASISSET* –
- DF auxiliary basis set for the fitting: 79 — *DFTBASIS* –
- scale factors and atomic configuration: 75 — *SCALEFAC* –
- k-points geometrical wheight: 53 — *KWEIGHTS* –
- shell symmetry analysis : 5, 6, 7, 8, 9
- Madelung parameters: 28
- multipole integrals: 20
- Fock/KS matrix building - direct lattice: 63, 64, 74 — *FGRED FGIRR* N
- Total energy contributions: 69 — *ENECYCLE* –

crystal - ***properties***

- shell and atom multipoles: 68 — *MULTIPOLE* N
- reciprocal space integration to compute Fermi energy: 51, 52, 53, 54, 55, 78
- density matrix - direct lattice: irreducible (58); reducible (59) (reducible P matrix in ***crystal*** if PPAN requested only) — *PGRED PGIRR* N
- Fock/KS eigenvalues : 66 — *EIGENVAL* N
  EIGENALL –
- Fock/KS eigenvectors : 67 — *EIGENVEC* N
- symmetry adapted functions : 47 — *KSYMMPRT* –

- Population analysis: 70, 73, 77 *MULLIKEN* N
- Atomic wave-function: 71

properties

- overlap matrix S(g) - direct lattice: 60 (keyword **PSIINF**) *OVERLAP* N
- Densities of states: 105, 107 *DOSS* –
- Projected DOSS for embedding: 36, 37, 38
- DF correlation correction to total energy: 106
- Compton profile and related quantities: 116, 117, 118
- Fermi contact tensor : 18 *FTENSOR* –
- rotated eigenvectors (keyword ROTREF): 67 *EIGENVEC* –
- Charge density and electrostatic potential maps: 119 *MAPVALUES* –

Example

To print the eigenvalues at each scf cycle enter:

```
PRINTOUT
EIGENALL
END
```

To print the eigenvalues at the first 5 k points at the end of scf only, enter in any input block:

```
SETPRINT
1
66 -5
```

Eigenvectors printed by default are from the first valence eigenvector up to the first 6 virtual ones. Core and virtual eigenvectors are printed by "adding" 500 to the selected value of LPRINT(67). To obtain print all the eigenvectors at the end of scf insert in any input block:

```
SETPRINT
1
66 -505
```

Printing options LPRINT array values

	subroutine	value	printed information	keyword	input
1	GCALCO	N	up N=6 stars of direct lattice vectors	GLATTICE	
2	CRYSTA	$\neq 0$	crystal symmetry operators	SYMMOPS	
3	EQUPOS	$\neq 0$	equivalent positions in the reference cell	EQUIVAT	
4	CRYSTA	$\neq 0$	crystal symmops after geometry editing		
5	GILDA1	N>0	g vector irr- first n set type of couples		
		N<0	g vector irr- n-th set type of couples		
6	GROTA1	$\neq 0$	information on shells symmetry related		
7	GV	N>0	stars of g associated to the first n couples		
7		N<0	stars of g associated to the n-th couple		
8	GORDSH	$\neq 0$	information on couples of shells symmetry related		
9	GSYM11	$\neq 0$	intermediates for symmetrized quantities		
10	GMFCAL	$\neq 0$	nstatg, idime, idimf, idimcou		
11	MAIN2U	$\neq 0$	exchange energy	EXCHGENE	
	MAIND			EXCHGENE	
12	IRRPR	$\neq 0$	symmops (reciprocal lattice)	SYMMOPSR	
13	MATVIC	N	n stars of neighbours in cluster definition		
14	GSLAB	$\neq 0$	coordinates of the atoms in the slab		
15	symdir	$\neq 0$	print symmetry allowed directions	PRSYMDIR	
18	TENSOR	$\neq 0$	extended printing for hyperfine coupling cost	FTENSOR	
19					
20	MONIRR	N	multipole integrals up to pole l=n		
21					
24	POINTCH		printing of point charges coordinates		
28	MADEL2	$\neq 0$	Madelung parameters		
29					
30	CRYSTA	$\neq 0$	write file FINDSYM.DAT		
31		$\neq 0$	values of the dimension parameters	PARAMETERS	
32		$N > 0$	printing of ccartesian coordinates of the atoms		
33	COOPRT	$N > 0$	cartesian coordinates of atoms in file fort.33	ATCOORDS	
34	FINE2	$N > 0$		KNETOUT	
	READ2		output of reciprocal space information	KNETOUT	
35		$N > 0$	printing of symmops in short fomr		
36	XCBD	$\neq$	properties - exchange correlation printing		
37					
38					
39					
40					
41	SHELL*	$\neq 0$	printing of bipolar expansion parameters		
47	KSYMBA	n	Symmetry Adapted Bloch Functions printing level		
48	KSYMBA	$\neq 0$	Symmetry Adapted Bloch Functions printing active	KSYMMPRT	
51	AB	$\neq 0$	B functions orthonormality check		
52	DIF	> 0	Fermi energy - Warning !!!! Huge printout !!!		
53	SCFPRT	$\neq 0$	k points geometrical weights	KWEIGHTS	
54	CALPES	> 0	k points weights- Fermi energy		
55	OMEGA	> 0	f0 coefficients for each band		
56					
57	PDIG	N	p(g) matrices-first n g vectors	PGIRR	N
58	PROT1	$\neq 0$	mvlu, ksh, idp4		
59	RROTA	$N > 0$	P(g) matrices - first N vectors at the end of SCF, if PPAN present	PGRED	N
	NEWK	N	P(g) matrices - first N vectors	PGRED	N
	PSIINF	> 0	P(g) matrices - first N vectors	PGRED	N
		N <0	P(g) matrix for g=N	PGRED	N
60	PSIINF	> 0	overlap matrix S(g) - first N vectors	OVERLAP	N
		N <0	overlap matrix S(g) for g = N		N
61					
63	TOTENY	$\neq 0$	bielectronic contribution to irred. F(g) matrix		
64	FROTA	N	F(g) matrix - first N g vectors	FGRED	N
	PSIINF	N> 0		FGRED	N
		N <0	f(g) matrix - for g = N (N-th g vector only)	FGRED	N

	subroutine	value	printed information	keyword	input
65					
66	AOFK	N	e(k)- fock eigenvalues- first N k vectors	EIGENVAL	N
	ADIK	N			
	BANDE	N			
	DIAG	N		EIGENALL	
	FDIK	N			
	FINE2	N			
	NEWK	N			
67	AOFK	N	a(k) - fock eigenvectors - first N k vectors	EIGENVEC	N
	ADIK	N			
	DIAG	N			
	FINE2	N			
	NEWK	N			
68	POLGEN	N <0	shell and atom multipoles up to pole l=N	MULTIPOL	N
	POLGEN	N >0	atom multipoles up to pole l=N	MULTIPOL	N
	QGAMMA	N	shell multipoles up to pole l=N	MULTIPOL	N
69	TOTENY	$\neq 0$	contributions to total energy at each cycle	ENECYCLE	
70	FINE2	$\neq 0$	Mulliken population analysis		
	NEIGHB		at the end of scf cycles		
	POPAN		calls PPBOND, to perform Mulliken analysis		
	PDIBAN				
71	PATIRR	$\neq 0$	atomic wave function	ATOMICWF	
	PATIR1	$\neq 0$	" "	ATOMICWF	
72	INPBAS	$\neq 0$	basis set	BASISSET	
	INPUT2	$\neq 0$	basis set	BASISSET	
	READFG	SET = 1			
73	POPAN	$\neq 0$	Mulliken matrix up to N direct lattice vector	MULLIKEN	N
	PPBOND				
	PDIBAN	N			
74	TOTENY	N	f(g) irreducible up to g=N	FGIRR	N
	DFTTT2	N		FGIRR	N
75	INPBAS	$\neq 0$	printing of scale factor and	SCALEFAC	
			atomic configuration	CONFIGAT	
76					
77	PPBOND	0	printing of neighbouring relationship		
		$\neq 0$	no printing of neighbours relationship		
78	FERMI	$\neq 0$	informations on Fermi energy calculation		
	EMIMAN	$\neq 0$			
79	DFGPRT	$\neq 0$	dft auxiliary basis set - default no printing	DFTBASIS	
80	ROTOP	> 0	printing of atoms coord. in rotated ref. frame	ROTREF	
92	INPBAS		G94 deck on ft92	GAUSS94	
93	MOLDRW		input deck to MOLDRAW		
105	DENSIM	< 0	DOSS along energy points	DOSS	
106	DFFIT3	> 0	DFT intermediate printout		
			(keyword PRINT in dft input)		
107	STARIN	$\neq 0$	DOSS information		
112	PROFCA	$\neq 0$	projected DOSS coefficients		
116	PROFI	$\neq 0$	Compton profile information		
117	PROFI	$\neq 0$			
118	PROFI	N			
119	INTEG	$\neq 0$	charge density at grid points	MAPVALUES	
	JJTEG	$\neq 0$	charge density at grid points	MAPVALUES	
	MAPNET	$\neq 0$	electrostatic potential at grid points	MAPVALUES	
	NAPNET	$\neq 0$	charge density gradient components	MAPVALUES	
120	LIBPHD	$\neq 0$	extended printing in berny optimizer		
121			reserved for geometry optimizer		
122			reserved for geometry optimizer		
123			reserved for geometry optimizer		
124			reserved for geometry optimizer		
125			reserved for geometry optimizer		

Appendix E

External format

Formatted data are written in files according to the following table:

program	keyword	ftn	filename		pg
	OPTGEOM	34	optaxxx	Geometry input - opt atoms coord. only - See **EXTPRT**	32
	OPTGEOM	34	optcxxx	Geometry input - opt cell [atoms] - See **EXTPRT**	32
		66	OPTHESS.DAT	Hessian - to restart optimization	
		68	OPTINFO.DAT	Information to restart optimization	
crystal	GAUSS98	92	GAUSSIAN.DAT	Input for GAUSS98	49
	FINDSYM	26	FINDSYM.DAT	data in crystallographic format - read by program findsym(IUCR)	30
	STRUCPRT	33	STRUC.INCOOR	Cell parameters, coordinates of atoms	43
	COORPRT	33	fort.33	Coordinates of the atoms in the cell	30
crystal	EXTPRT	34	fort.34	Geometry input	32
&	MOLDRAW	93	MOLDRAW.DAT	Input for MOLDRAW	36
properties	PPAN	24	PPAN.DAT	Mulliken population analysis	74
properties	BAND	25	fort.25	Bands (Crgra2006)	112
		24	BAND.DAT	Bands data	112
	CLAS	25	fort.25	Classical potential	115
	DIEL	24	DIEL.DAT	Dielectric constant	116
	DOSS	25	fort.25	Density of states (IPLOT=1)	118
		24	DOSS.DAT	Density of states (IPLOT=2)	118
	ECHG	25	fort.25	Electronic charge density - 2D grid	119
		25	RHOLINE.DAT	Electronic charge density - 1 grid	119
	ECH3	31	—	Electronic charge density - 3D grid	119
	EMDL	25	fort.25	EMD line (IPLOT=1)	121
		24	EMDL.DAT	EMD line(IPLOT=2)	121
	EMDP	25		EMD - 2D grid	214
	INFOGUI	32		Data for the graphical user interface	123
	POTC	24	POTC.DAT	Electrostatic potential V, Electric field, Electric field gradient	137
	POTM	25	fort.25	Electrostatic potential - 2D grid	138
	PROF	25	fort.25	Compton profile and related quantities (IPLOT=1)	139
		24	PROF.DAT	Compton profile and related quantities (IPLOT=2)	139

Data in file fort.25 are read by the programs **maps06, doss06, band06** of the package Crgra2006. In the same run bands, density of states, value of a function in a 2D grid of points can be computed. The appropriate command (**maps06, doss06, band06**) selects and plots the selected data .
The package can be downloaded from:

http://www.crystal.unito.it/Crgra2006/Crgra2006.html

CLAS - ECHG - POTM - Isovalue maps

The value of the function chosen (classic electrostatic potential (CLAS), charge(+spin) density (ECHG), electrostatic potential (POTM)) is computed in a given net of points. The data are written in file fort.25.
If the system is spin polarized, total density data are followed by spin density data.

Structure of the file fort.25

```
1ST RECORD : -%-,IHFERM,TYPE,NROW,NCOL,DX,DY,COSXY format: A3,I1,A4,2I4,3E12.5
2ND RECORD : XA,YA,ZA,XB,YB,ZB                     format: 1P,6E12.5
3RD RECORD : XC,YC,ZC,NAF,LDIM                     format: 1P,3E12.5,4X,2I4
4TH RECORD
AND FOLLOWING : ((RDAT(I,J),I=1,NROW),J=1,NCOL)    format: 1P,6E12.5
```

Meaning of the variables:

```
1 '-%-'     3 character string marks the beginning of a block of data;
1 IHFERM:   0 : closed shell, insulating system
            1 : open shell,   insulating system
            2 : closed shell, conducting system
            3 : open shell,   conducting system

1 TYPE      4 characters string corresponding to the type of data "MAPN"
1 NROW      number of rows of the data matrix RDAT
1 NCOL      number of columns of the data matrix RDAT
1 DX        increment of x (\AA ngstrom) in the plane of the window
1 DY        increment of y (\AA ngstrom) in the plane of the window
1 COSXY     cosine of the angle between x and y axis;
2 XA,YA,ZA coordinates of the  points A,B (see keyword MAPNET) (\AA ngstrom)
2 XB,YB,ZB defining the window where the functions is computed (\AA ngstrom)
3 XC,YC,ZC coordinates of point C (\AA ngstrom)
3 NAF       number of atoms in the cell
3 LDIM      dimensionality (0 molecule; 1 polymer, 2 slab, 3 bulk)
4->         ncol*nrow values of the function (a.u.) at the nodes of the grid
```

naf records follow, with atomic number, symbol, coordinates (Ångstrom) of the atoms in the cell:

```
NAT,SYMBAT,XA,YA,ZA                              format: I4,1X,A,1P,3E20.12

NAT        atomic number
SYMBAT     Mendeleev symbol
XA,YA,ZA   cartesian coordinates of the atoms in the cell (\AA ngstrom)
```

Cartesian components of cell parameters follow (Ångstrom)

```
AX, AY, AZ  cartesian component of vector a      format: 3E20.12
BX, BY, BZ  cartesian component of vector b      format: 3E20.12
CX, CY, CZ  cartesian component of vector c      format: 3E20.12
```

The program **maps06** looks for the atoms lying in the windows used to compute the function, and it can draw the symbol of the atoms, the van der Waals sphere, or the bonds between atoms closer than the sum of their vdW radii.

ECHG Charge (spin) density - 1D profile

When points B and C coincides in **ECHG** 5.2 input, coordinates relative to the origin of the segment and charge density value [coordinate along the line, charge density: charge density derivative x,y,z components] are written with format (2E20.12:3E20.12) in file RHOLINE.DAT. A second set of data, spin density, is written for spin polarized systems, after a blank line.

BAND - Band structure

Hamiltonian eigenvalues are computed at k points corresponding to a given path in the Brillouin zone. Data are written in file BAND.DAT and processed by DLV; see http://www.cse.clrc.ac.uk/cmg/DLV) and in file fort.25 (processed by Crgra2006/band06)

Structure of the file fort.25

One block is written for each segment of the path in k reciprocal space: the segment is defined by two k points, whose crystallographic coordinates (I1,I2,I3) and (J1,J2,J3) are given as integers in ISS units (see keyword BAND).
If the system is spin polarized, α electrons bands are followed by β electrons bands.
For each segment:

```
1ST RECORD : -%-,IHFERM,TYPE,NBAND,NKP,DUM,,DK,EF  format: A3,I1,A4,2I4,3E12.5
2ND RECORD : EMIN,EMAX                             format: 1P,6E12.5
3RD RECORD : I1,I2,I3,J1,J2,J3                     format: 6I3
4TH RECORD
AND FOLLOWING : ((RDAT(I,J),I=1,NROW),J=1,NCOL)    format: 1P,6E12.5
```

Meaning of the variables:

```
1 '-%-'    3 character string marks the beginning of a block of data;
1 IHFERM:  0 : closed shell, insulating system
           1 : open shell,   insulating system
           2 : closed shell, conducting system
           3 : open shell,   conducting system

1 TYPE     4 characters string corresponding to the type of data "BAND"
1 NBAND    number of bands
  NKP      number of k points along the segment
  DUM      not used
  DK       distance in k space between two adjacent sampling points
           along the segment
  EF       Fermi energy (hartree)
2 EMIN     minimum energy of the bands in the explored path (hartree)
  EMAX     maximum energy (hartree)
3 I1,I2,I3,J1,J2,J3 : coordinates of the segment extremes in iunit of ISS
4 EPS(I,J) eigenvalues (hartree): eps(i,j) corresponds to the i-th
..          band, and the j-th k point of the segment.
```

DIEL

The data computed are written in file DIEL.DAT according to the following format:

```
#
@ XAXIS LABEL "DISTANCE(BOHR)"
@ YAXIS LABEL "MACRORHO  MACROE MACROV RHOPLANE"
5 columns  - format(08E15.7)
last record is blank
```

DOSS Density of states

Total and projected density of states are written in file DOSS.DAT (processed by DLV; see http://www.cse.clrc.ac.uk/cmg/DLV) and in file fort.25 (processed by Crgra2006).
One block is written for each projected density of states, including the total one: so NPRO (number pf projections) +1 blocks are written per each run.
If the system is spin polarized, α electrons bands are followed by β electrons bands.

Structure of the file written in file fort.25

```
1ST RECORD : -%-,IHFERM,TYPE,NROW,NCOL,DX,DY,COSXY
    format : A3,I1,A4,2I5,1P,(3E12.5)
2ND RECORD : X0,Y0                              format : 1P,6E12.5
3RD RECORD : I1,I2,I3,I4,I5,I6                  format : 6I3
4TH RECORD
AND FOLLOWING : ((RDAT(I,J),I=1,NROW),J=1,NCOL) format :  1P,6E12.5

Meaning of the variables:

1 NROW                1 (DOSS are written one projection at a time)
  NCOL                number of energy points in which the DOS is calculated
```

```
  DX                  energy increment (hartree)
  DY                  not used
  COSXY               Fermi energy (hartree)
2 X0                  energy corresponding to the first point
  Y0                  not used
3 I1                  number of the projection;
  I2                  number of atomic orbitals of the projection;
  I3,I4,I5,I6         not used
4 RO(J),J=1,NCOL      DOS: density of states ro(eps(j)) (atomic units).
```

Structure of the file written in file DOSS.DAT

Data written in file DOSS.DAT:

```
1ST RECORD : NPUNTI, NPRO1, IUHF
    format : '# NEPTS',1X,I5,1X,'NPROJ',1X,I5,1X,'NSPIN',1X,I5
2ND RECORD : '#'
3RD RECORD : '@ YAXIS LABEL "DENSITY OF STATES (STATES/HARTREE/CELL)"'
4TH RECORD : (ENE(I),DOSS(IPR,I),IPR=1,NPRO1)
AND FOLLOWING :
    format : 1P,15E12.4
```

PROF

The computed quantities are written following the same sequence of the printout. Each record contains:

```
4F    coordinate,  all electron, core,  valence contribution
```

EMDL

The computed quantities are written following the same sequence of the printout. Each record contains:

```
coordinate/ band projections / orbital projections / total
```

DOSS

```
1   #
2   @ XAXIS LABEL "ENERGY (HARTREE)"'
3   @ YAXIS LABEL "DENSITY OF STATES (STATES/HARTREE/CELL)"

NPUNTI records follow (all electron DOSS for restricted calculations,
alpha electrons DOSS for unrestricted calculations. Each records contains:

energy, total DOSS, doss projections 1, 2, 3....

Unrestricted calculation only: beta electrons DOSS, preceded by a delimiter:
&
NPUNTI records follow, with the same format as the previous set:

energy, total DOSS, doss projections 1, 2, 3....
```

EMDP

```
1ST RECORD : -%-,IHFERM,TYPE,NMAX1,NMAX2,PMAX1,PMAX2,COS12
    format : A3,I1,A4,2I5,1P,(3E12.5)
2ND RECORD : XDUM,YDUM                               format : 1P,6E12.5
3RD RECORD : I11,I12,I13,I21,I22,I23                   format : 6I3
```

```
4TH RECORD
AND FOLLOWING : ((RDAT(I,J),I=1,NMAX1),J=1,NMAX2) format :  1P,6E12.5

Meaning of the variables:
1 '-%-'        3 character string marks the beginning of a block of data;
1 IHFERM:      0 : closed shell, insulating system
               1 : open shell,   insulating system
               2 : closed shell, conducting system
               3 : open shell,   conducting system

1 TYPE         4 characters string corresponding to the type of data "EMDP"
1 NMAX1        number of points in the first direction
  NMAX2        number of points in the second direction
  PMAX1        maximum p value along the first direction
  PMAX2        maximum p value along the first direction
  COSXY        angle between the two vectors defining the plane
2 X0           not used
  Y0           not used
3 I11,I12,I13  fractional coordinates of the first reciprocal lattice
               vector defining the plane
  I21,I22,I23  fractional coordinates of the second reciprocal lattice
               vector defining the plane
4 RO(J),J=1,NMAX1*NMAX2  electron momentum density at the grid points
              (atomic units).
```

POTC

When ICA $\neq 0$; NPU $\neq 0$ (2D or 3D systems) the data computed are written in file POTC.DAT according to the following format:

```
#
@ XAXIS LABEL "Z (AU)"
@ YAXIS LABEL "ELECTROSTATIC PROPERTIES (AU)"
@ TITLE "String in the first record in crystal input                        "
@ SUBTITLE "ELECTRIC FIELD INTENSITY:    0.100 AU" ! if external field applied
@ LEGEND ON
@ LEGEND LENGTH 3
@ LEGEND X1 0.87
@ LEGEND Y1 0.8
@ LEGEND STRING 0 "V"
@ LEGEND STRING 1 "E"
@ LEGEND STRING 2 "DE/DZ"
@ LEGEND STRING 3 "RHO"
@ LEGEND STRING 4 "V+VEXT"                          ! if external field applied
@ LEGEND STRING 5 "VEXT"                            ! if external field applied
NPU records of 5 (7 when external field applied) columns - format 08E15.7
```

COORPRT

The keyword **COORPRT**, entered in geometry input or in ***properties*** writes in file fort.33 (append mode) the following data:

record #	data type	content
1	I	number of atoms (NAF)
2	A	Title - If written after an SCF calculation, on the same line; totalenergy, convergence on energy, number of cycles
3	A,3F	Mendeleev symbol of the atom; *x, y, z* cartesian coordinates (Å)
......		
NAF+2	A,3F	Mendeleev symbol of the atom; *x, y, z* cartesian coordinates (Å)

The coordinates of the atoms are written at each geometry optimization cycle (keyword **OPTGEOM**
The file "fort.33" is read by the program **MOLDEN** [23] which can be downloaded from: `www.cmbi.kun.nl/ schaft/molden/molden.html`

STRUCPRT

The file STRUCPRT.DAT is written according to the format given in the example (output for bulk MgO, 2 atoms per cell).

```
$cell vectors                cartesian components of cell parameters (bohr)
    0.00000000000000    3.97787351190423    3.97787351190423
    3.97787351190423    0.00000000000000    3.97787351190423
    3.97787351190423    3.97787351190423    0.00000000000000
$coordinates                 cartesian coordinates of atoms (bohr)
MG  0.00000000000000    0.00000000000000    0.00000000000000    12
O   0.00000000000000    0.00000000000000   -3.97787351190423     8
$END
```

PPAN

```
# Mulliken Populations:
# NSPIN,NATOM          n. determinants, number of atoms
  ----  for each atom
# IAT,NSHELL              atomic number, number o shells
# Xiat,Yiat,Ziat (AU)     cartesian coordinates (bohr)
# QTOT, QSHELL,I=1,NSHELL atom total electronic charge, (shell charges)
# NORB, QORB, I=1,NORB    number of orbitals, (orbital electronic charges)

Example:

graphite STO-3G basis set, RHF (1 eterminant)
2 atoms, 2 shells per atom, 5 AO per atom

      1          2                 |      1 determinant, 2 atoms
      6          2                 |  1st atom: atomic number 6, 2 shells
 -1.320 -2.287  0.000              |        cartesian coordinates 1st atom
  6.000  1.993  4.007              |     6, electronic charge of 1st atom
                                   |  1.993 electronic charge of 1st shell (1s)
                                   |  4.007 electronic charge of 2nd shell (2sp)
      5                            |      5 atomic orbitals
  1.993  1.096  0.956  0.956  1.000|  1.993 electronic charge of 1st AO (1s)
                                   |  1.096 electronic charge of 2nd AO (2s)
                                   |  0.956 electronic charge of 3rd AO (px)
                                   |  0.956 electronic charge of 4th AO (py)
                                   |  1.000 electronic charge of 5th AO (pz)
      6          2                 |  2nd atom: atomic number 6, 2 shells
```

```
 -2.640  0.000  0.000             |        cartesian coordinates 2nd atom
  6.000  1.993  4.007             |     6, electronic charge of 1st atom
                                  |  1.993 electronic charge of 1st shell (1s)
                                  |  4.007 electronic charge of 2nd shell (2sp)
     5                            |      5 atomic orbitals
  1.993  1.096  0.956  0.956  1.000|  1.993 electronic charge of 1st AO (1s)
                                  |  1.096 electronic charge of 2nd AO (2s)
                                  |  0.956 electronic charge of 3rd AO (px)
                                  |  0.956 electronic charge of 4th AO (py)
                                  |  1.000 electronic charge of 5th AO (pz)
     6         2                  |   second atom: atomic number 6, 2 shells
 -2.640  0.000  0.000             |   cartesian coordinates 2nd atom
```

EXTPRT / EXTERNAL - file fort.34

Geometry information can be read from an external file, fort.34, by entering the keyword **EXTERNAL**. The system can be a molecule, a polymer, a slab or a crystal. The file is written by entering the keyword **EXTPRT** in the input block 1. The file is written at the end of successful geometry optimization. The "history" of the optimization process is written in files optaxxx (xxx number of optimization cycle) or optcxxx. //[0.2cm] The structure of the file is as follow:

rec #	data type	contents
2	3I	dimensionality, centring and crystal type
3	3F	cartesian components of the direct lattice vectors
4	3F	
5	3F	
6	1I	number of symmetry operators
		For each symmetry operator 4 records:
7	3F	symmetry operators matrices in cartesian coordinates
8	3F	
9	3F	
10	3F	cartesian components of the translation
n	1I	number of atoms in the primitive cell
		For each atom, 1 record:
n+1	I,3F	conventional atomic number, cartesian coordinates of the atoms

The keyword **EXTERNAL** and **END** must be inserted at the top and bottom of the deck to use it as CRYSTAL geometry input.

Example - Test05 - Graphite 2D - standard geometry input

```
SLAB              dimensionality
77                layer group number
2.42              lattice parameter
1                 number of irreducible atoms in the cell
6  -0.33333333333 0.33333333333 0.  coordinates of the atoms
EXTPRT
TESTGEOM
END
```

Data written in file fort.34 (Ångstrom):

```
   2   1   5                                  ! dimensionality, centring and crystal type
  0.2095781E+01 -0.1210000E+01  0.0000000E+00 ! cartesian components of direct lattice vectors
  0.0000000E+00  0.2420000E+01  0.0000000E+00 !
  0.0000000E+00  0.0000000E+00  0.5000000E+03 ! 2D system - formal value 500. \AA
 12                                           ! number of symmetry operators
  0.1000000E+01  0.0000000E+00  0.0000000E+00 ! 1st symmetry operator - 3x3 transformation matrix
  0.0000000E+00  0.1000000E+01  0.0000000E+00 !
  0.0000000E+00  0.0000000E+00  0.1000000E+01 !
  0.0000000E+00  0.0000000E+00  0.0000000E+00 ! 1st symmetry operator - 3x1 translation component
 -0.1000000E+01  0.0000000E+00  0.0000000E+00 ! 2nd symmetry operator
  0.0000000E+00 -0.1000000E+01  0.0000000E+00 !
  0.0000000E+00  0.0000000E+00  0.1000000E+01 !
```

```
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
-0.5000000E+00 -0.8660254E+00  0.0000000E+00 ! 3rd symmetry operator
 0.8660254E+00 -0.5000000E+00  0.0000000E+00 !
 0.0000000E+00  0.0000000E+00  0.1000000E+01 !
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
-0.5000000E+00  0.8660254E+00  0.0000000E+00 ! 4th symmetry operator
-0.8660254E+00 -0.5000000E+00  0.0000000E+00 !
 0.0000000E+00  0.0000000E+00  0.1000000E+01 !
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
 0.5000000E+00 -0.8660254E+00  0.0000000E+00 ! 5th symmetry operator
 0.8660254E+00  0.5000000E+00  0.0000000E+00 !
 0.0000000E+00  0.0000000E+00  0.1000000E+01 !
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
 0.5000000E+00  0.8660254E+00  0.0000000E+00 ! 5th symmetry operator
-0.8660254E+00  0.5000000E+00  0.0000000E+00 !
 0.0000000E+00  0.0000000E+00  0.1000000E+01 !
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
-0.5000000E+00  0.8660254E+00  0.0000000E+00 ! 7th symmetry operator
 0.8660254E+00  0.5000000E+00  0.0000000E+00 !
 0.0000000E+00  0.0000000E+00  0.1000000E+01 !
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
 0.1000000E+01  0.0000000E+00  0.0000000E+00 ! 8th symmetry operator
 0.0000000E+00 -0.1000000E+01  0.0000000E+00 !
 0.0000000E+00  0.0000000E+00  0.1000000E+01 !
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
-0.5000000E+00 -0.8660254E+00  0.0000000E+00 ! 9th symmetry operator
-0.8660254E+00  0.5000000E+00  0.0000000E+00 !
 0.0000000E+00  0.0000000E+00  0.1000000E+01 !
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
 0.5000000E+00  0.8660254E+00  0.0000000E+00 ! 10th symmetry operator
 0.8660254E+00 -0.5000000E+00  0.0000000E+00 !
 0.0000000E+00  0.0000000E+00  0.1000000E+01 !
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
-0.1000000E+01  0.0000000E+00  0.0000000E+00 ! 11th symmetry operator
 0.0000000E+00  0.1000000E+01  0.0000000E+00 !
 0.0000000E+00  0.0000000E+00  0.1000000E+01 !
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
 0.5000000E+00 -0.8660254E+00  0.0000000E+00 ! 12th symmetry operator
-0.8660254E+00 -0.5000000E+00  0.0000000E+00 !
 0.0000000E+00  0.0000000E+00  0.1000000E+01 !
 0.0000000E+00  0.0000000E+00  0.0000000E+00 !
 1                                           ! number of irreducible atoms in the primitive cell
   6     -0.6985938  -1.2100000    0.0000000 ! conventional atomic number, cartesian coordinate
```

ECH3/POT3/GRID3D

Fortran unit 31 is written through the keyword **ECH3** (page 119), **POT3** (page136), **GRID3D** (page122). All data in atomic units.

rec #	data type	contents
1	A	title: charge density /spin density
2	3I	npa,npb,npc, number of points along the 3 directions
3	3E	x,y,z cartesian coordinates of the point (1,1,1)
4	3E	dxa, dya, dza cartesian components of the step along a
5	3E	dxb, dyb, dzb cartesian components of the step along b
6	3E	dxc, dyc, dzc cartesian components of the step along c
7 ...	5E	npa*npb*npc floating point data, 5/record

INFOGUI

Fortran unit 32 is written through the keyword **INFOGUI** (page 123). The format is almost self-explaining. The following data are written for MgO bulk (test11).

```
  2 atom(s) per cell
  6 shells
 18 atomic orbitals
 20 electrons per cell
 12 core electrons per cell
 No eigenvalue level shifting
 No Alpha-Beta Spin locking
```

```
No N. Beta Spin locking
Type of Calculation: RESTRICTED CLOSED SHELL
Total Energy = -0.27466415E+03H
Fermi Energy = -0.31018989E+00H
1 -0.31018989E+00
6    18    20    12 |   # shells, # AO, # electrons, # core electrons
2                   |   # atoms
1   12    1   0.000000  0.000000  0.000000 | sequence number, atomic number,?,cartesian coor(bohr)
3   # shells attributed to the first atom
0   shell type (s) of the 1st shell
1   shell type (sp) of the 2nd shell
1   shell type (sp) of the 3rd shell
2    8    2   3.977874  3.977874  3.977874 | sequence number, atomic number,?,cartesian coor(bohr)
3   # shells attributed to the second atom
0   shell type (s) of the 1st shell
1   shell type (sp) of the 2nd shell
1   shell type (sp) of the 3rd shell
```

Interface to external programs

The keyword **CRYAPI_OUT**, present into ***properties*** input stream writes formatted wave function information, both in direct and reciprocal space, in file GRED.DAT and KRED.DAT The scripts *runcry06* and *runprop06* save them in the current directory as inpfilename.GRED and inpfilename.KRED .
The program ***cryapi_inp***, written in fortran 90, is distributed as source code (http://www.crystal.unito.it => documentation => utilities). It reads and prints the data, showing the meaning of the variables and the organization of data
cryapi_inp should be compiled by any fortran 90 compiler: comments and request for more information are welcome (mail to *crystal@unito.it*).

GRED.DAT

The file GRED.DAT contains:

- Geometry, symmetry operators;
- Local functions basis set (including ECP)
- Overlap matrix in direct lattice
- Hamiltonian matrix in direct lattice
- Density matrix in direct lattice
- Wannier functions (if file fort.80, written by **LOCALWF** when localization is successful, is present)

Overlap, hamiltonian, density matrices are written as arrays of non-zero elements. GRED.DAT contains the information to build full matrices.
All data are printed executing ***cryapi_inp***

KRED.DAT

The file KRED.DAT is written if eigenvectors have been computed (keyword **NEWK** 5.2) by ***properties***.
CRYSTAL works in the irreducible Brillouin (IBZ) zone only: eigenvectors in the full Brillouin zone (BZ) are computed by rotation, and by time reversal symmetry, when necessary. The file KRED.DAT contains:

- Coordinates of k points in irreducible Brillouin zone, according to Pack-Monkhorst net
- Symmetry operators in reciprocal lattice
- Geometrical weight of k points
- Hamiltonian eigenvalues
- Weight of k points for each band (computed by Fermi energy calculation)
- Eigenvectors in full Brillouin zone

Structure of matrices in direct lattice

Overlap, hamiltonian, and density matrices in direct lattice are arrays of non-zero elements: ***cryapi_inp*** prints the matrices as triangular (hamiltonian) or square matrices of size (local BS x local BS), for a limited number of direct lattice vectors, to show the structure of the arrays.

From IBZ to BZ

CRYSTAL works on irreducible Brillouin zone (IBZ), full information is generated by applying rotation operators.
Time reversal symmetry is exploited in reciprocal lattice: the inversion symmetry is always present, even if the inversion operator is not present in direct lattice.
Given a shrinking factor according to Pack-Monkhorst sampling, to total number of k points is for instance:

System	n. symmops	shrink factors	IBZ	NOSYMM	BZ
graphite (2D)	12	3	3	5	9
SiC (3D)	24	4	8	36	64
MgO (3D)	48	4	8	36	64

IBZ	number of points in IBZ
NOSYMM	number of points removing direct lattice symmetry
BZ	number of points in Brillouin zone

Appendix F

Normalization coefficients

A. Bert - Thesis 2002

The aim of this appendix is to show how normalization coefficients of the basis functions are defined in CRYSTAL and to describe how they are stored in the program.

Basic Definitions

Let us consider a function, $f(\mathbf{r})$; we have in general:

$$\int d\mathbf{r}\ |f(\mathbf{r})|^2 \neq 1; \tag{F.1}$$

however, we can always define a related $f'(\mathbf{r})$, multiplying $f(\mathbf{r})$ by a constant N:

$$f'(\mathbf{r}) = N f(\mathbf{r}), \tag{F.2}$$

such that:

$$\int d\mathbf{r}\ |f'(\mathbf{r})|^2 = 1. \tag{F.3}$$

$f'(\mathbf{r})$ is said to be a *normalized* function and N is its *Normalization Coefficient* (NC). Substituting eq. F.2 in F.3, we have:

$$N = \left(\int d\mathbf{r}\ |f(\mathbf{r})|^2\right)^{-1/2}. \tag{F.4}$$

Gaussians: Product Theorem and Normalization

Let us define Gaussian functions as:

$$G(\alpha_i; \mathbf{r}-\mathbf{A}) = \exp(-\alpha_i(\mathbf{r}-\mathbf{A})^2), \tag{F.5}$$

where $\mathbf{A}$ is the *centroid* of the function.
The Gaussian product theorem states that the product of two Gaussians, is still a Gaussian function:[1]

$$G(\alpha; \mathbf{r}-\mathbf{A})G(\beta; \mathbf{r}-\mathbf{B}) = \exp\left(-\frac{\alpha\beta}{\xi}|\mathbf{R}|^2\right) G(\xi; \mathbf{r}-\mathbf{P}); \tag{F.8}$$

[1] Let us prove the Gaussian product theorem:

$$\begin{aligned} G(\alpha; \mathbf{r}-\mathbf{A})G(\beta; \mathbf{r}-\mathbf{B}) &= \exp(-\alpha_i(\mathbf{r}-\mathbf{A})^2)\exp(-\alpha_j(\mathbf{r}-\mathbf{B})^2 \\ &= \exp\left(-\alpha(\mathbf{r}^2+\mathbf{A}^2+2\mathbf{r}\mathbf{A}) - \beta(\mathbf{r}^2+\mathbf{B}^2+2\mathbf{r}\mathbf{B})\right) \\ &= \exp\left[-\xi\left((\mathbf{r}-\mathbf{P})^2+\mathbf{P}^2-\frac{\alpha\mathbf{A}^2+\beta\mathbf{B}^2}{\xi}\right)\right]. \end{aligned} \tag{F.6}$$

with:

$$\xi = \alpha + \beta, \tag{F.9}$$

$$\mathbf{P} = \frac{\alpha \mathbf{A} + \beta \mathbf{B}}{\xi}, \tag{F.10}$$

$$\mathbf{R} = \mathbf{A} - \mathbf{B}. \tag{F.11}$$

From eq. F.4, the NC of Gaussian functions, g_i, can be written as:

$$\begin{aligned} g_i &= \left(\int d\mathbf{r} \; (G(\alpha_i; \mathbf{r}))^2 \right)^{-1/2} \\ &= \left(\int d\mathbf{r} \, G(2\alpha_i; \mathbf{r}) \right)^{-1/2} \\ &= \left(\frac{\pi}{2\alpha_i} \right)^{-3/4}, \end{aligned} \tag{F.12}$$

where the Gaussian product theorem and the Gaussian integral [119] have been used. $G'(\alpha_i; \mathbf{r})$, defined as:

$$G'(\alpha_i; \mathbf{r}) = g_i \, G(\alpha_i; \mathbf{r}), \tag{F.13}$$

is a normalized function.

Harmonic Gaussians

The Definition

The *Solid Harmonic Functions*, Y_ℓ^m, [120] are defined as:

$$Y_\ell^m(\mathbf{r}) = r^\ell P_\ell^{|m|}(cos\vartheta) e^{im\phi}, \tag{F.14}$$

where P_ℓ^m is the *Legendre Polynomial Function* characterized by the integers ℓ and m, such that: $\ell \geq 0$ and $-\ell \leq m \leq \ell$. [121]
Starting from Y_ℓ^m, the *Real Solid Harmonic*, X_ℓ^m, can be defined:

$$X_\ell^{|m|}(\mathbf{r}) \quad = \Re(Y_\ell^{|m|}) = \quad \frac{Y_\ell^{|m|}(\mathbf{r}) + Y_\ell^{-|m|}(\mathbf{r})}{2}, \tag{F.15}$$

$$X_\ell^{-|m|}(\mathbf{r}) \quad = \Im(Y_\ell^{|m|}) = \quad \frac{Y_\ell^{|m|}(\mathbf{r}) - Y_\ell^{-|m|}(\mathbf{r})}{2i}. \tag{F.16}$$

We report some examples of X functions.
$\ell = 0$:

$$X_0^0(\mathbf{r}) = 1; \tag{F.17}$$

$\ell = 1$:

$$X_1^0(\mathbf{r}) = z, \quad X_1^1(\mathbf{r}) = x, \quad X_1^{-1}(\mathbf{r}) = y; \tag{F.18}$$

$\ell = 2$:

$$X_2^0(\mathbf{r}) = z^2 - 0.5(x^2 - y^2), \quad X_2^1(\mathbf{r}) = 3zx, \quad X_2^{-1}(\mathbf{r}) = 3zy, \tag{F.19}$$

$$X_2^2(\mathbf{r}) = 3(x^2 + y^2), \quad X_2^{-2}(\mathbf{r}) = 3xy. \tag{F.20}$$

Using eqs. F.9, F.10 and F.11, eq. F.6 can be rewritten as:

$$G(\alpha; \mathbf{r} - \mathbf{A}) G(\beta; \mathbf{r} - \mathbf{B}) = \exp\left(-\frac{\alpha\beta}{\xi} |\mathbf{R}|^2 \right) G(\xi; \mathbf{r} - \mathbf{P}). \tag{F.7}$$

We have now the tools required to define the *Solid Harmonic Gaussian*, [120] ξ:

$$\xi^{n\ell m}(\alpha_i;\mathbf{r}) = |\mathbf{r}|^{2n} Y_\ell^m(\mathbf{r}) G_i(\alpha_i;\mathbf{r}), \tag{F.21}$$

where n is a non-negative integer number ($n \geq 0$). We are interested here only in $n = 0$ harmonic Gaussians (that is, $\xi^{0\ell m}$), so we shall simply write (omitting the $n = 0$ index):

$$\xi^{\ell m}(\alpha_i;\mathbf{r}) = Y_\ell^m(\mathbf{r}) G(\alpha_i;\mathbf{r}). \tag{F.22}$$

Substituting Y with X (eqs. F.15 and F.16) in eq. F.22, *Real Harmonic Gaussians*, γ, can be defined:

$$\gamma^{\ell m}(\alpha_i;\mathbf{r}) = X_\ell^m(\mathbf{r}) G(\alpha_i;\mathbf{r}). \tag{F.23}$$

γ are used as basis functions in the CRYSTAL program and are related to the ξ ones by followings relations:

$$\gamma^{\ell|m|} = \frac{\xi^{\ell|m|} + \xi^{\ell\,-|m|}}{2}, \tag{F.24}$$

$$\gamma^{\ell-|m|} = \frac{\xi^{\ell|m|} - \xi^{\ell\,-|m|}}{2i}, \tag{F.25}$$

where eqs. F.15 and F.16 have been used.
Note that, when ℓ is equal to 0, ξ and γ functions degenerate to simple Gaussians:

$$\xi^{00} = \gamma^{00} = G, \tag{F.26}$$

where eq. F.17 has been used and ξ degenerates to γ when $m = 0$:

$$\xi^{\ell 0} = \gamma^{\ell 0}, \tag{F.27}$$

where eqs. F.24 and F.25 have been used.

The Normalization Coefficient

Let us consider now ξ and γ's normalization coefficients (b and c, respectively), from eq. F.4, follows:

$$b_i^{\ell m} = (\Xi)^{-1/2} \tag{F.28}$$

$$c_i^{\ell m} = (\Upsilon)^{-1/2}, \tag{F.29}$$

where

$$\Xi = \int d\mathbf{r}\ \left|\xi^{\ell m}(\alpha_i;\mathbf{r})\right|^2 \tag{F.30}$$

$$\Upsilon = \int d\mathbf{r}\ \left(\gamma^{\ell m}(\alpha_i;\mathbf{r})\right)^2. \tag{F.31}$$

Using eqs. F.5, F.8, F.14, F.22 and a spherical polar coordinate system,[2] the Ξ integral can be factorized as:

$$\begin{aligned}\Xi &= \int d\mathbf{r}\ \left[Y_\ell^m(\mathbf{r})G(\alpha_i;\mathbf{r})\right]^* Y_\ell^m(\mathbf{r})G(\alpha_i;\mathbf{r}) \\ &= \int d\mathbf{r}\, Y_\ell^{-m}(\mathbf{r})\, Y_\ell^m(\mathbf{r})\, G(2\alpha_i;\mathbf{r}) \\ &= \Xi_r\ \Xi_\vartheta\ \Xi_\phi,\end{aligned} \tag{F.32}$$

[2] $d\mathbf{r} = r^2 sin\vartheta\, dr\, d\vartheta\, d\phi$

with:

$$\begin{aligned}\Xi_r &= \int_0^\infty dr\ \exp(-2\alpha_i r^2) r^{2\ell+2} \\ &= \frac{\Gamma(\ell+3/2)}{2(2\alpha_i)^{\ell+3/2}} \\ &= \frac{\pi^{1/2}(2\ell+1)!!}{2^{\ell+2}(2\alpha_i)^{\ell+3/2}},\end{aligned} \tag{F.33}$$

where we have used the Γ function's properties; [121]

$$\begin{aligned}\Xi_\vartheta &= \int_0^\pi d\vartheta\ \left(P_\ell^{|m|}(\cos\vartheta)\right)^2 \sin\vartheta \\ &= \frac{2(\ell+|m|)!}{(2\ell+1)(\ell-|m|)!},\end{aligned} \tag{F.34}$$

where the Legendre polynomials' properties have been used, [121] and

$$\Xi_\phi = \int_0^{2\pi} d\phi = 2\pi. \tag{F.35}$$

Substituting eqs. F.32, F.33, F.34 and F.35 in the b definition (eq. F.28) we obtain:

$$\begin{aligned}b_i^{\ell m} &= \frac{\pi^{1/2}(2\ell+1)!!}{2^{\ell+2}(2\alpha_i)^{\ell+3/2}}\ \frac{2(\ell+|m|)!}{(2\ell+1)(\ell-|m|)!}\ 2\pi \\ &= \left(\frac{\pi^{3/2}\ (2\ell-1)!!\ (\ell+|m|)!}{2^{2\ell+3/2}\ \alpha_i^{\ell+3/2}\ (\ell-|m|)!}\right)^{-1/2}.\end{aligned} \tag{F.36}$$

Note that b is independent from the sign of m (as Ξ is), that is:

$$b_i^{\ell|m|} = b_i^{\ell\,-|m|}. \tag{F.37}$$

In order to deduce the explicit expression for c, we are interested now in solving the integral of eq. F.31:

$$\Upsilon = \int d\mathbf{r}\ \left(X_\ell^m(\mathbf{r})\right)^2\ G(2\alpha_i;\mathbf{r}), \tag{F.38}$$

where eqs. F.8 and F.23 have been used. Substituting eq. F.24 (γ functions with $m \geq 0$) in previous equation, we have:

$$\begin{aligned}\Upsilon^{m\geq 0} = \frac{1}{4}\Bigg(&\int d\mathbf{r}\ \left|\xi^{\ell|m|}(\alpha_i;\mathbf{r})\right|^2 + \int d\mathbf{r}\ \left|\xi^{\ell\,-|m|}(\alpha_i;\mathbf{r})\right|^2 + \\ &+ 2\int d\mathbf{r}\,\xi^{\ell|m|}(\alpha_i;\mathbf{r})\xi^{\ell\,-|m|}(\alpha_i;\mathbf{r})\Bigg).\end{aligned} \tag{F.39}$$

The first two integrals in eq. F.39 can be recognized as Ξ (eq. F.32, reminding that Ξ is independent from the m sign); the last one, if $m \neq 0$, is null for the orthogonality properties of the Harmonic functions, [121] therefore:

$$\Upsilon^{m>0} = \frac{\Xi}{2}. \tag{F.40}$$

The same result is found for negative m, substituting eq. F.25 (instead of eq. F.24, as done) in eq. F.31:

$$\Upsilon^{m<0} = \frac{\Xi}{2}, \tag{F.41}$$

so Υ (as Ξ is) is independent from the m sign. If $m=0$, the last integral in eq. F.39 is equal to Ξ, as the first two ones:

$$\Upsilon^{m=0}=\Xi; \tag{F.42}$$

the previous equation can be deduced also from eq. F.27.
Summarizing, from eqs. F.40, F.41 and F.42, we get:

$$\Upsilon=\frac{\Xi}{2-\delta_{m0}} \tag{F.43}$$

and, finally, substituting eqs. F.32 and F.43 in eq. F.29, we obtain:

$$c_i^{\ell m}=\left(\frac{\pi^{3/2}\,(2\ell-1)!!\,(\ell+|m|)!}{2^{2\ell+3/2}(2-\delta_{m0})\,\alpha_i^{\ell+3/2}\,(\ell-|m|)!}\right)^{-1/2}. \tag{F.44}$$

The c expression (eq. F.44) can be reorganized in a two factors formula:

$$c_i^{\ell m}=a_i^{\ell}\,f^{\ell m}, \tag{F.45}$$

with:

$$a_i^{\ell}=\left(\frac{\pi^{3/2}}{(2\alpha_i)^{\ell+3/2}}\right)^{-1/2}, \tag{F.46}$$

the α-dependent term, and

$$f^{\ell m}=\left(\frac{(2\ell-1)!!\,(\ell+|m|)!}{2^{\ell}(2-\delta_{m,0})(\ell-|m|)!}\right)^{-1/2}, \tag{F.47}$$

the m dependent term.
Note that,

- If $\ell=0$, γ degenerates in a simple Gaussian (eq. F.26),

$$f^{00}=1 \quad \text{and} \quad c_i^{00}=a_i^0=g_i, \tag{F.48}$$

  where g_i is the G's NC (eq. F.12).

- If $\ell=1$, $f^{1m}=1/2$ for the three m-values:

$$f^{1m}=1/2 \quad \text{and} \quad c_i^{1m}=\frac{a_i^1}{2}=\frac{\alpha_i^{5/4}2^{7/4}}{\pi^{3/4}},\ \forall\ m=-1,0,1. \tag{F.49}$$

- If $\ell=2$, we have:

$$c_i^{20}=\frac{\alpha_i^{7/4}2^{11/4}}{\pi^{3/4}\sqrt{3}};\quad c_i^{21}=c_i^{2-1}=\frac{\alpha_i^{7/4}2^{11/4}}{\pi^{3/4}3};\quad c_i^{22}=c_i^{2-2}=\frac{\alpha_i^{7/4}2^{7/4}}{\pi^{3/4}3}. \tag{F.50}$$

Let us verify, for two examples, that

$$\gamma'=c\,\gamma \tag{F.51}$$

is a normalized function, proving that the following integral, I, is equal to 1,

$$I_i^{\ell m}=\int d\mathbf{r}\,\left(c_i^{\ell m}\gamma^{\ell m}(\alpha_i;\mathbf{r})\right)^2. \tag{F.52}$$

The s Case ($\ell = 0$, $m = 0$)

$$\begin{aligned} I_i^{00} &= \left(c_i^{00}\right)^2 \int d\mathbf{r}\ \left(\gamma^{00}(\alpha_i;\mathbf{r})\right)^2 \\ &= (g_i)^2 \int d\mathbf{r}\ \left(G(\alpha_i;\mathbf{r})\right)^2 \\ &= \left(\int \left(G(\alpha_i;\mathbf{r})\right)^2 d\mathbf{r}\right)^{-1} \int \left(G(\alpha_i;\mathbf{r})\right)^2 d\mathbf{r} = 1, \end{aligned} \tag{F.53}$$

where eqs. F.48, F.27 and F.12 have been used.

A d Case ($\ell = 2$, $m = 1$)

$$I_i^{21} = \int d\mathbf{r}\ \left(c_i^{21}\gamma^{21}(\alpha_i;\mathbf{r})\right)^2 = \left(c_i^{21}\right)^2 J, \tag{F.54}$$

with:

$$J = \int d\mathbf{r}\ \left(3zxG(2\alpha_i;\mathbf{r})\right)^2, \tag{F.55}$$

where eqs. F.23 and F.19 have been used.
Gaussians are separable functions, that is:

$$G(\alpha_i;\mathbf{r}) = G_x(\alpha_i;x)\, G_y(\alpha_i;y)\, G_z(\alpha_i;z), \tag{F.56}$$

with:

$$G_x(\alpha_i;x) = \exp(-\alpha_i x^2) \tag{F.57}$$

and similarly for y and z. Substituting eq. F.56 in eq. F.55, we have:

$$J = 9J_xJ_yJ_z, \tag{F.58}$$

with:

$$J_x = \int x^2 G_x(2\alpha_i;x)dx = \frac{\sqrt{\pi}}{2}(2\alpha_i)^{-3/2}, \tag{F.59}$$

$$J_y = \int G_y(2\alpha_i;y)dy = \left(\frac{\pi}{2\alpha_i}\right)^{1/2}, \tag{F.60}$$

$$J_z = \int z^2 G_z(2\alpha_i;z)dz = \frac{\sqrt{\pi}}{2}(2\alpha_i)^{-3/2}, \tag{F.61}$$

where ref. [119] has been used in solving the integrals. Substituting now eqs. F.49 and F.58 in eq. F.54, we obtain:

$$I_i^{21} = \frac{\alpha_i^{7/2}\,2^{11/2}}{\pi^{3/2}\,9}\ 9\left(\frac{\sqrt{\pi}}{2}(2\alpha_i)^{-3/2}\right)^2\left(\frac{\pi}{2\alpha_i}\right)^{1/2} = 1. \tag{F.62}$$

Atomic Orbitals Normalization

The variational basis functions of the CRYSTAL program (AOs), φ_μ, are normalized *contractions* (fixed linear combinations) of normalized real solid harmonic Gaussian type functions (*primitive functions*), γ' (eq. F.51). The AOs are organized in *shells*, φ_μ belonging to the same shell, λ, have same radial part, that is, same contraction coefficients, d_i^λ, same Gaussian exponents, α_i^λ and different angular part, X_ℓ^m:

$$\varphi_\lambda^{\ell m} = N_\lambda \sum_i d_i^\lambda\, c_i^{\ell m}\, \gamma^{\ell m}(\alpha_i^\lambda;\mathbf{r}) = N_\lambda \sum_i d_i^\lambda\, c_i^{\ell m}\, X_\ell^m(\mathbf{r})\, G(\alpha_i^\lambda;\mathbf{r}). \tag{F.63}$$

The index i runs over the primitive functions of the contraction, d_i^λ is the contraction coefficient of the i-th primitive in shell λ and, as we have seen, it is the same for all the AOs of λ, that is, it does not depend on ℓ or m. γ and c are the primitive function and its NC (eq. F.29), respectively. N_λ is the NC of AOs belonging to λ and is defined as:

$$N_\lambda = \left(\int d\mathbf{r} \left(\sum_i d_i^\lambda \, c_i^{\ell m} \, \gamma^{\ell m}(\alpha_i^\lambda; \mathbf{r}) \right)^2 \right)^{-1/2} , \tag{F.64}$$

in the following will be shown that N depends only on the shell, λ.
We report, as an example, the three AOs of a p-type shell ($\ell = 1$), supposing that λ is classified as the fourth shell of the unitary cell and each AO is a contraction of two primitives.

$$p_z = \varphi_4^{10} = N^4 \left(d_1^4 \, c_1^{10} \, \gamma^{10}(\alpha_1^4; \mathbf{r}) + d_2^4 \, c_2^{10} \, \gamma^{10}(\alpha_2^4; \mathbf{r}) \right) , \tag{F.65}$$

$$p_x = \varphi_4^{11} = N^4 \left(d_1^4 \, c_1^{11} \, \gamma^{11}(\alpha_1^4; \mathbf{r}) + d_2^4 \, c_2^{11} \, \gamma^{11}(\alpha_2^4; \mathbf{r}) \right) , \tag{F.66}$$

$$p_y = \varphi_4^{1-1} = N^4 \left(d_1^4 \, c_1^{1-1} \, \gamma^{1-1}(\alpha_1^4; \mathbf{r}) + d_2^4 \, c_2^{1-1} \, \gamma^{1-1}(\alpha_2^4; \mathbf{r}) \right) . \tag{F.67}$$

Let us put our attention on N_λ. Eq. F.64 can be rewritten as:

$$N^\lambda = \left(\sum_{i,j} d_i^\lambda \, d_j^\lambda \, c_i^{\ell m} c_j^{\ell m} \Upsilon' \right)^{-1/2} , \tag{F.68}$$

with:

$$\Upsilon' = \int d\mathbf{r} \, \gamma^{\ell m}(\alpha_i; \mathbf{r}) \, \gamma^{\ell m}(\alpha_j; \mathbf{r}), \tag{F.69}$$

where the shell index on α has been omitted for simplicity. Substituting eq. F.23 in eq. F.69, we have:

$$\Upsilon' = \int X_\ell^m(\mathbf{r}) \, G(\alpha_i; \mathbf{r}) \, X_\ell^m(\mathbf{r}) \, G(\alpha_j; \mathbf{r}) d\mathbf{r} = \int \left(X_\ell^m(\mathbf{r}) \right)^2 G[(\alpha_i + \alpha_j); \mathbf{r}] d\mathbf{r}, \tag{F.70}$$

where the Gaussian product theorem (eq. F.8) has been used.
From eq. F.31, it can be seen that Υ' differs from Υ only in the Gaussian exponent ($\alpha_i + \alpha_j$ instead of $2\alpha_i$), using then eqs. F.43, F.32, F.34 and F.35, Υ' is rewritten as:

$$\Upsilon' = \frac{\Upsilon_r' \, \Xi_\vartheta \, \Xi_\varphi}{2 - \delta_{m0}}, \tag{F.71}$$

with:

$$\begin{aligned} \Upsilon_r' &= \int_0^\infty dr \, \exp[-(\alpha_i + \alpha_j) r^2] r^{2\ell+2} \\ &= \frac{\Gamma(\ell + 3/2)}{2(\alpha_i + \alpha_j)^{\ell+3/2}} \\ &= \frac{\pi^{1/2} (2\ell + 1)!!}{2^{\ell+2} (\alpha_i + \alpha_j)^{\ell+3/2}} . \end{aligned} \tag{F.72}$$

Substituting eqs. F.44, F.71 and F.72 in eq. F.68, we obtain:

$$N_\lambda = \left(\sum_{i,j} d_i^\lambda \, d_j^\lambda \left(\frac{2\sqrt{\alpha_i^\lambda \alpha_j^\lambda}}{\alpha_i^\lambda + \alpha_j^\lambda} \right)^{\ell+3/2} \right)^{-1/2} , \tag{F.73}$$

where it is clear that N depends only on λ.

The Code

In order to explain easily the organization of NCs in CRYSTAL, eq. F.63 is reorganized as follows:

$$\varphi_\lambda^{\ell m} = \sum_i n_{\lambda,i}^{\ell m} \, \gamma^{\ell m}(\alpha_i^\lambda; \mathbf{r}), \tag{F.74}$$

with:

$$n_{\lambda,i}^{\ell m} = N_\lambda \, d_i^\lambda \, c_i^{\ell m}. \tag{F.75}$$

Note that, while the AO is normalized, the function $\gamma'' = n\,\gamma$ is not; in fact n is not a normalization factor, and it will be referred as the *pre-Gaussian factor*.

At the moment the CRYSTAL code is able to treat four type of shells: s, sp, p and d.[3] An s shell has only an AO, that is a contraction of simple Gaussians ($\ell = 0$); in a p one there are three AOs (different for the m value, p_x, p_y and p_z) with $\ell = 1$ primitives; d shells are obviously formed by five $\ell = 2$ AOs. The three basis functions of a sp shell are contractions of one s primitive function and several ps'.

In the calculation of the integrals required in the SCF process, n must be very often multiplied by the constant factor $\pi^{5/8}\,2^{1/4}$; [22] therefore, in the code, pre-Gaussian factors are not stored, but the following quantities, that we shall call *code pre-Gaussian constants*:

$$\mathsf{S}_i^\lambda = \pi^{5/8}\,2^{1/4}\,n_{\lambda,i}^{00} \tag{F.76}$$

$$\mathsf{P}_i^\lambda = \pi^{5/8}\,2^{1/4}\,n_{\lambda,i}^{1m} \;\forall\; m = 0, 1, -1 \tag{F.77}$$

$$\mathsf{D}_i^\lambda = \pi^{5/8}\,2^{1/4}\sqrt{\frac{(2+|m|)!}{(2-\delta_{m0})(2-|m|)!}}\;n_{\lambda,i}^{2m} \;\forall\; m = 0, 1, -1, 2, -2. \tag{F.78}$$

Note that the square root in eq. F.78 (the inverse of the m-dependent part of c, eq. F.44) makes D independent from the m value, whereas $n_{\lambda,i}^{2m}$ depends from it. In such a way, S, P and D are m-independent

In the `inpbas` routine, contraction coefficients (as defined in input), d_i^λ, related to s, p and d AOs, are loaded in the two dimension packed arrays `c1`, `c2` and `c3`, respectively (they are stored in the module `basato_module`). Their length corresponds to the total number of primitives in the unit cell and is the same for the three arrays. The first elements are the contraction coefficients for the first shell (d_i^1), then the d_i^2s (second shell) follows, and so on; the contraction index, i, is the internal one. For an s shell, for example, the elements of `c2` and `c3` are null, of course.

In the `gaunov` routine, `c1`, `c2` and `c3` are redefined and loaded with the code pre-Gaussian constants S, P and D, respectively; naturally they maintain the described organization and module `basato_module` is overwritten.

In `gaunov` two further arrays, `c2w` and `c3w` (that follow the convention used in the ATMOL program) are also defined and loaded in `basato_module`. They have the same organization as `c1`, `c2` and `c3` and contain $\mathsf{P}_i'^\lambda$ and $\mathsf{D}_i'^\lambda$ coefficients, respectively:

$$\mathsf{P}_i'^\lambda = \frac{\pi^{5/8}\,2^{1/4}}{2\alpha_i}\,n_{\lambda,i}^{1m} \;\forall\; m = 0, 1, -1 \tag{F.79}$$

$$\mathsf{D}_i'^\lambda = \frac{\pi^{5/8}\,2^{1/4}}{(2\alpha_i)^2}\sqrt{\frac{(2+|m|)!}{(2-\delta_{m0})(2-|m|)!}}\;n_{\lambda,i}^{2m} \;\forall\; m = 0, 1, -1, 2, -2. \tag{F.80}$$

[3] The implementation of higher polynomial functions is now in progress.

We give an example of evaluation of an overlap integral $S_{\mu\nu}$ over an s and a $m = 0$ d AO ($\varphi_\mu \equiv \varphi^\lambda_{00}$, $\varphi_\nu \equiv \varphi^\sigma_{20}$) sitting in the reference cell:

$$S_{\mu\nu} = \int d\mathbf{r}\, \varphi^\lambda_{00}(\mathbf{r})\varphi^\sigma_{20}(\mathbf{r}). \tag{F.81}$$

Substituting eq. F.74 in the previous equation, we have:

$$S_{\mu\nu} = \sum_{ij} n^{00}_{\lambda,i}\, n^{20}_{\sigma,j} \int d\mathbf{r}\, \gamma^{00}(\alpha^\lambda_i;\mathbf{r})\gamma^{20}(\alpha^\sigma_j;\mathbf{r}). \tag{F.82}$$

Since in the code, S and D are available (but not the n coefficients), we express n as a function of code pre-Gaussian constants, using eqs. F.76 and F.78, and we rewrite the overlap integral as:

$$S_{\mu\nu} = \left(\pi^{5/8}\, 2^{1/4}\right)^{-2} \sqrt{\frac{(2-\delta_{m0})(2-|m|)!}{(2+|m|)!}} \sum_{ij} \mathsf{S}^\lambda_i \mathsf{D}^\sigma_i \int d\mathbf{r}\, \gamma^{00}(\alpha^\lambda_i;\mathbf{r})\gamma^{20}(\alpha^\sigma_j;\mathbf{r}). \tag{F.83}$$

Note that the m-dependent term contained in n, for d shells, must be multiplied *a posteriori*, because is not included in D. This operation is performed in the `dfac3` routine, that provides McMurchie-Davidson coefficients multiplied by code pre-Gaussian constants and, when λ is a d shell, by the m-dependent part of $n^{2m}_{\lambda,i}$.

Appendix G

Incompatibility

Crystal

ANDERSON and **LEVSHIFT** are not compatible.

Properties

When symmetry adaption of Bloch Functions is active, the symmetry analysis in reciprocal space is performed and stored at each **k** point.
NEWK computes the eigenvectors of the Hamiltonian (Hartree-Fock or Kohn-Sham) at the **k** points defined in the Monkhorst net.
BAND computes the eigenvectors of the Hamiltonian (Hartree-Fock or Kohn-Sham) at the **k** points defined by the path chosen in input.
If properties are computed from the eigenvectors, the sequence:

```
NEWK
...
BANDE
....
newprop from eigenvectors
```

is forbidden, if symmetry adaption of the Bloch functions is active (default choice).
The keyword **NOSYMADA** removes symmetry adaption of the Bloch functions, and the sequence is allowed.

Appendix H

CRYSTAL06 versus CRYSTAL03

Input

CRYSTAL06 wave function calculation (program ***crystal***) input now consists of 3 blocks, instead of 4):

1. geometry input:
   geometry optimization new package and keyword **OPTGEOM**;

2. basis set input;

3. hamiltonian & SCF The shrinking factors to define Monkhorst (IS) and Gilat (ISP) net are entered after a new keyword **SHRINK** (3rd input block) (page 75):

   ```
   SHRINK
   IS ISP
   ```

CRYSTAL06 properties calculation (program ***properties***) **NEWK** input is modified:

```
NEWK
IS ISP
IFE NUMPRT (see User's Manual)
```

Output

Wave function data written in file fort.9 and fort.98 are different.

Algorithms and numerical implementation

The most important improvements in algorithms and numerical implementation are:

- Screening of the bielectronic integrals
  A few modifications in the screening of terms of infinite sums may affect the value of total energy slightly in some cases. In the calculation of the Coulomb and exchange series the main selection criteria for the integrals to be computed are based on the overlap between shells of a pair. However, a shell consists of atomic orbitals resulting from a contraction of gaussian functions and some of these integrals may give negligible contributions to the sum. For this reason, an additional screening operates a finer selection of the integrals to be computed which is based on every single gaussian exponent. With the default setting

of tolerances (TOLINTEG) a change in the geometry of a system may affect this latter selection and introduce small artificial discontinuities in the potential energy surface. This happened in CRYSTAL03, even if the FIXINDEX option, aiming at smoothing potential energy surfaces, was active. This problem has been fixed in CRYSTAL06. That implies that this new release may produce slightly different results at some geometry of a system when using FIXINDEX, for example during geometry optimizations.

- DFT grids for numerical integration
  In the generation of the grids used for the numerical integration of the exchange-correlation density functional a screening of the grid points based on their geometrical weight and the parameter TOLGRID leads to a beneficial reduction of the grid size. These geometrical weights include a multiplicity factor associated with the symmetry properties of each point. This factor was erroneously taken into account during the selection of the grid points in CRYSTAL03. Such inconsistency has been removed from CRYSTAL06. This may result in slight changes in the total energy per cell.

New features keywords are entered in geometry input block: full optimization (**OPTGEOM**, page 82) and frequency calculation (**FREQCALC**, page 98). They are presented in new sections.

Test cases
crystal03 versus crystal06 total energies (hartree)

	crystal03	crystal06	diff
TEST 0	-110.7649354548	-110.7649354541	-7.6E-10
TEST 1	-39.7267242377	-39.7267242374	-2.8E-10
TEST 2	-223.7874757199	-223.7874756819	-3.8E-08
TEST 3	-893.8746580039	-893.8746580004	-3.4E-09
TEST 4	-1400.1776585267	-1400.1776585535	2.7E-08
TEST 5	-74.8333583547	-74.8333583570	2.3E-09
TEST 6	-58.4208255976	-58.4208255980	3.6E-10
TEST 7	-2800.7355953670	-2800.7355953744	7.4E-09
TEST 8	-571.3207540598	-571.3207540594	-4.2E-10
TEST 9	-29.2566111179	-29.2566111179	0.0E+00
TEST10	-577.8265583366	-577.8265583285	-8.0E-09
TEST11	-274.6641530285	-274.6641530559	2.7E-08
TEST12	-447.6810664605	-447.6810664796	1.9E-08
TEST13	-23.9856901336	-23.9856901143	-1.9E-08
TEST14	-159.6970601311	-159.6970601598	2.9E-08
TEST15	-5229.8366028287	-5229.8366027787	-5.0E-08
TEST16	-2995.2869386925	-2995.2869386582	-3.4E-08
TEST17	-2674.3752958292	-2674.3752958032	-2.6E-08
TEST18	-679.2766564865	-679.2766564082	-7.8E-08
TEST19	-223.8070778082	-223.8070777860	-2.2E-08
TEST20	-89.9552982254	-89.9552981100	-1.2E-07
TEST21	-447.5749512396	-447.5749511978	-4.2E-08
TEST22	-460.7186326899	-460.7186326563	-3.4E-08
TEST23	-8.0613160104	-8.0613160317	2.1E-08
TEST24	-1400.1776187881	-1400.1776188146	2.6E-08
TEST25	-74.8442039897	-74.8442039913	1.5E-09
TEST26	-58.4208255857	-58.4208255860	3.1E-10
TEST27	-2800.7355409765	-2800.7355409839	7.4E-09
TEST28	-8.0429325843	-8.0429330764	4.9E-07
TEST29	-2047.6430863277	-2047.6430862965	-3.1E-08
TEST30	-109.0441465735	-109.0441458665	-7.1E-07
TEST31	-4095.2867581984	-4095.2867581744	-2.4E-08
TEST32	-92.1408103960	-92.1408103960	0.0E+00
TEST33	-92.1416132011	-92.1416129818	-2.2E-07
TEST34	-1117.5230436301	-1117.5230436113	-1.9E-08
TEST35	-936.5017524997	-936.5017511475	-1.4E-06
TEST36	-112.5648953627	-112.5648952230	-1.4E-07
TEST37	-3028.3682393194	-3028.3682392875	-3.2E-08
TEST38	-2279.1395902355	-2279.1395902381	2.6E-09

Appendix I

Relevant strings

Selected information can be extracted from CRYSTAL output referring to some strings of characters uniquely linked to the requested information.

```
TOTAL ENERGY(        final SCF energy
TOTAL ENERGY(HF      Hartree-Fock
TOTAL ENERGY(DFT     DFT
TTT END              final elapsed and CPU time (crystal/properties)
 OPT END             energy after geometry optimization
 OPT END - FAILED    failed opt only
 OPT END - CONVERGED successful opt only
TTT BERNY            cpu time for each opt cycle
GEOMETRY FOR WAVE    printing of the geometry used for wf calculation
                     (after editing)
FINAL OPTIMIZED      printing of geometry at the end of optimization
```

Appendix J

Acronyms

AFM – Anti ferromagnetic
AO – Atomic Orbital
APW – Augmented Plane Wave
a.u. – atomic units
BF – Bloch Function
BS – Basis set
BSSE – Basis Set Superposition Error
BZ – Brillouin Zone (first)
B3PW – Becke Perdew Wang
B3LYP – Becke - Lee - Yang - Parr
CO – Crystalline Orbital
CPU – Central Processing Unit
DF(T) – Density Functional (Theory)
DM – Dipole Moment (see Wannier Functions)
DOS – Density of States
ECP – Effective Core Potentials
EFG – Electric Field Gradient
EMD – Electron Momentum Density
FM – Ferromagnetic
GC – Gradient-Corrected
GGA – Generalised Gradient Approximation
GS(ES) – Ground State (Electronic Structure)
GT(O) – Gaussian Type (Orbital)
GT(F) – Gaussian Type (Function)
GUI – Graphical User Interface
KS – Kohn and Sham
HF – Hartree-Fock
IBZ – Irreducible Brillouin zone
IR – Irreducible Representation
LAPW – Linearized Augmented Plane Wave
LCAO – Linear Combination of Atomic Orbitals
LDA – Local Density Approximation
LP – Local Potential
LSDA – Local Spin Density Approximation
LYP – GGA Lee-Yang-Parr
MO – Molecular Orbital

MPP – Massive Parallel Processor
MSI – Molecular Simulation Inc.
NLP – Non-local potential (correlation)
PBE – GGA Perdew-Burke-Ernzerhof
PDOS – Projected Density of States
PP – Pseudopotential
PVM – Parallel Virtual Machine
PW – Plane Wave
PWGGA – GGA. Perdew-Wang
PWLSD – LSD Perdew-Wang
PZ – Perdew-Zunger
P86 – GGA Perdew 86
P91 – Perdew 91
QM – Quantum Mechanics
RCEP – Relativistic Compact Effective Potential
RHF – Restricted Hartree-Fock
ROHF – Restricted Open-shell Hartree-Fock
SAED – Symmetry Allowed Elastic Distortions
SABF – Symmetry Adapted Bloch Functions SC – Supercell
SCF – Self-Consistent-Field
STO – Slater Type Orbital
UHF – Unrestricted Hartree-Fock
VBH – von Barth-Hedin
VWN – Vosko-Wilk-Nusair
WnF – Wannier Functions 0D – no translational symmetry
1D – translational symmetry in 1 direction (x, CRYSTAL convention)
2D – translational symmetry in 2 directions (x,y, CRYSTAL convention)
3D – translational symmetry in 3 directions (x,y,z CRYSTAL convention)

Bibliography

[1] R. Dovesi, C. Pisani, C. Roetti, M. Causà and V.R. Saunders, *CRYSTAL88, An ab initio all-electron LCAO-Hartree-Fock program for periodic systems. QCPE Pgm N.***577**, Quantum Chemistry Program Exchange, Indiana University, Bloomington, Indiana (1989).

[2] R. Dovesi, C. Roetti and V.R. Saunders, *CRYSTAL92 User's Manual*, Università di Torino and SERC Daresbury Laboratory (1992).

[3] R. Dovesi, V.R. Saunders, C. Roetti, M. Causà, N. M. Harrison, R. Orlando and E. Aprà, *CRYSTAL95 User's Manual*, Università di Torino (1996).

[4] V.R. Saunders, R. Dovesi, C. Roetti, M. Causà, N. M. Harrison, R. Orlando and C.M. Zicovich-Wilson, *CRYSTAL98 User's Manual*, Università di Torino (Torino,1998).

[5] V.R. Saunders, R. Dovesi, C. Roetti, R. Orlando, C.M. Zicovich-Wilson, N. M. Harrison, K. Doll, B. Civalleri, I. Bush, Ph. D'Arco, M. Llunell, *CRYSTAL03 User's Manual*, Università di Torino (Torino,2003).

[6] B. Roos, C. Salez, A. Veillard and E. Clementi, *A general program for calculation of atomic SCF orbitals by the expansion method*, Tech. Rep. RJ518 IBM Res. (1968).

[7] D.J. Hehre, W.A. Lathan, M.D. Newton, R. Ditchfield and J.A. Pople, *GAUSSIAN70 Program number 236, QCPE.* Indiana University,Bloomington, Indiana (1972).

[8] H. Burzlaff and A. Hountas, *J. Appl. Cryst.* **15**, 464 (1982).

[9] D. Moncrieff and V.R. Saunders, *ATMOL*. University of Manchester Regional Computing Centre, Doc NAT648 (, 1986).

[10] B.G. Searle, "DL Visualize", *Comp. Phys. Communications* **137**, 25–32 (2001).

[11] C. Giacovazzo, *Fundamentals of Crystallography*, IUC-Oxford University Press, Oxford (1992).

[12] C. Pisani, *Quantum-Mechanical Ab-Initio Calculation of the Properties of Crystalline Materials*, *Lecture Notes in Chemistry*, volume 67, Springer Verlag, Berlin (1996).

[13] T. Hahn, *International Tables for Crystallography*, Reidel Publishing Company (1987).

[14] T. Asthalter, W. Weyrich, A.H. Harker, A.B. Kunz, R. Orlando and C. Pisani, "Comparison of quasi-Hartree-Fock wave-functions for lithium hydride", *Solid State Comm.* **83**, 725 (1992).

[15] D.J. Hehre, L. Radom, P.v.R. Schleyer and J.A. Pople, *Ab initio molecular orbital theory*, Wiley, New York (1986).

[16] M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T.A. Keith, G.A. Pettersson, J.A. Montgomery, K. Raghavachari, M.A. Al-Laham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, J. Cioslowski, B.B. Stefanov, A. Nanayakkara, M. Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogle, R. Gomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. Defrees, J. Baker, J.J.P. Stewart, M. Head-Gordon, C. Gonzalez, and J.A. Pople, *Gaussian 94 (Revision A.I)*, Gaussian Inc., Pittsburgh PA (1995).

[17] R. Dovesi and R. Orlando, "Convergence properties of the supercell approach in the study of local defects in solids", *Phase Trans.* **52**, 151 (1994).

[18] M. Causà, R. Dovesi, C. Pisani and C. Roetti, "Electronic structure and stability of different crystal phases of magnesium oxide", *Phys. Rev. B* **33**, 1308 (1986).

[19] H.J. Monkhorst and J.D. Pack, *Phys. Rev.* **13**, 5188 (1976).

[20] G. Gilat and L.J. Raubenheimer, "Accurate numerical method for calculating frequency-distribution in solids", *Phys. Rev.* **144**, 390 (1966).

[21] G. Gilat, *J. Comp. Phys.* **10**, 432 (1972).

[22] C. Pisani, R. Dovesi and C. Roetti, *Hartree-Fock ab initio Treatment of Crystalline Systems*, *Lecture Notes in Chemistry*, volume 48, Springer Verlag, Heidelberg (1988).

[23] G. Schaftenaar, *MOLDEN a pre- and post processing program of molecular and electronic structure*, CMBI, the Netherlands.

[24] R. Dovesi, C. Roetti, C. Freyria Fava, M. Prencipe and V.R. Saunders, "On the elastic properties of lithium, sodium an potassium oxide. An ab initio study", *Chem. Phys.* **156**, 11 (1991).

[25] L. Fu , E. Yaschenko, L. Resca and R. Resta, "Hartree–Fock approach to macroscopic polarisation: Dielectric constant and dynamical charge of $KNbO_3$", *Phys. Rev. B* **57**, 6967–6971 (1988).

[26] C. Darrigan, M. Rerat, G. Mallia, R. Dovesi, "Implementation of the finite field perturbation method in the CRYSTAL program for calculating the dielectric constant of periodic systems", *J. Comp. Chem.* **24**, 1305–1312 (2003).

[27] P. Ugliengo, D. Viterbo and G. Chiari, "MOLDRAW: Molecular Graphics on a Personal Computer", *Zeitschrift fur Kristallographie* **207**, 9 (1993).

[28] S.F. Boys and F. Bernardi, "The calculation of small molecular interaction by the difference of separate total energies. Some procedures with reduced errors", *Mol. Phys.* **19**, 553 (1970).

[29] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, *Gaussian 98, Revision A.3.* Pittsburgh PA (1998).

[30] P.J. Hay and W.R. Wadt, "*Ab initio* effective core potentials for molecular calculations. Potentials for transition metal atoms Sc to Hg.", *J. Chem. Phys.* **82**, 270 (1985).

[31] P.J. Hay and W.R. Wadt, "*Ab initio* effective core potentials for molecular calculations. Potentials for K to Au including the outermost orbitals", *J. Chem. Phys.* **82**, 299 (1985).

[32] J.C. Barthelat. Private communication.

[33] P. Durand and J.C. Barthelat, *Theor. Chim. Acta* **38**, 283 (1975).

[34] J.C. Barthelat and P. Durand, *Gazz. Chim. Ital.* **108**, 225 (1978).

[35] J.C. Barthelat, P. Durand and A. Serafini, *Molec. Phys.* **33**, 159 (1977).

[36] H. Stoll, G. Igel-Mann and H. Preuss. Arbeitsbericht 27 des Instituts für Theoretische Chemie, Universität Stuttgart (, 1990).

[37] P.J. Hay and W.R. Wadt, "*Ab initio* effective core potentials for molecular calculations. Potentials for main group elements Na to Bi", *J. Chem. Phys.* **82**, 284 (1985).

[38] Y. Bouteiller, C. Mijoule, M. Nizam, J.C. Barthelat, J.P. Daudey, M. Pellissier and B. Silvi, *Mol. Phys.* **65**, 295 (1988).

[39] W.J. Stevens, H. Bash, M. Krauss, "Compact effective potentials and efficient shared-exponent basis sets for the first- and second-row atoms.", *J. Chem. Phys* **81**, 6026–6033 (1984).

[40] W.J. Stevens, M. Krauss, H. Basch, P.G. Jasien, "Relativistic compact effective potentials and efficient, shared-exponent basis sets for the third-, fourth-, and fifth-row atoms", *Can. J. Chem* **70**, 612–630 (1992).

[41] P. A. M. Dirac, *Proc. Cambridge Phil. Soc.* **26**, 376 (1930).

[42] U. von Barth and L. Hedin, *J. Phys. C: Solid State Phys.* **5**, 1629 (1972).

[43] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior", *Phys. Rev. A* **38**, 3098–3100 (1988).

[44] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple", *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

[45] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems", *Phys. Rev. B* **23**, 5048–5079 (1981).

[46] J. P. Perdew and Wang Yue, "Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation", *Phys. Rev. B* **33**, 8800–8802 (1986).

[47] J. P. Perdew and Wang Yue, "Erratum: Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation", *Phys. Rev. B* **40**, 3399 (1989).

[48] J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron gas correlation energy", *Phys. Rev. B* **45**, 13244 (1992).

[49] S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis", *Can. J. Phys.* **58**, 1200 (1980).

[50] J. P. Perdew, "Density-functional approximation for the correlation energy of the inhomogeneous electron gas", *Phys. Rev. B* **33**, 8822 (1986).

[51] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", *Phys. Rev. B* **37**, 785–789 (1988).

[52] D. G. Anderson, *J. Assoc. Comput. Mach.* **12**, 547 (1964).

[53] D. R. Hamann, "Semiconductor Charge Densities with Hard-Core and Soft-Core Pseudopotentials", *Phys. Rev. Lett.* **42**, 662–665 (1979).

[54] D. D. Johnson, "Modified Broyden's method for accelerating convergence in self-consistent calculations", *Phys. Rev B* **38**, 12807–12813 (1988).

[55] C. G. Broyden, *Math. Comput.* **19**, 577 (1965).

[56] W. Kohn and L.J. Sham, "Self-consistent equations including exchange and correlation effects", *Phys. Rev.* **140**, A1133 (1965).

[57] P Hohenberg and W. Kohn, "Inhomogeneous electron gas", *Phys. Rev.* **136**, B864 (1964).

[58] Zhigang Wu, R.E. Cohen, "More accurate generalized gradient approximation for solids", *Phys. Rev. B* **73**, Art. n. 235116 (2006).

[59] J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh and C. Fiolhais, "Atoms, molecules, solids and surfaces: applications of the generalized gradient approximation for exchange and correlation", *Phys. Rev. B* **46**, 6671 (1992).

[60] A. D. Becke, “Density-functional thermochemistry. III The role of exact exchange”, *J. Chem. Phys.* **98**, 5648 (1993).

[61] J. P. Perdew, *Electronic Structure of Solids 1991*, Akademie Verlag, Berlin (1991).

[62] C. Adamo and V. Barone, “Toward reliable density functional methods without adjustable parameters: the PBE0 model”, *J. Chem. Phys.* **110**, 6158–6170 (1999).

[63] A. D. Becke, *J. Chem. Phys.* **88**, 2547 (1988).

[64] M.D. Towler, A. Zupan and M. Causà, “Density Functional theory in periodic systems using Local Gaussian basis sets”, *Comp. Phys. Commun.* **98**, 181–205 (1996).

[65] A. D. Becke, *J. Chem. Phys.* **88**, 1053 (1988).

[66] A. Savin, *Int. J. Quantum Chem.* **S22**, 457 (1988).

[67] P.M.W. Gill, B.G. Johnson, J.A. Pople, “A standard grid for density function calculations”, *Chem. Phys. Lett.* **209**, 506–512 (1993).

[68] O. Treutler, R. Ahlrichs, “Efficient molecular numerical integration schemes”, *J. Chem. Phys.* **102**, 346–354 (1995).

[69] L. Kronsjo, *Algorithms, their complexity and efficiency*, volume 2nd ed., Wiley, New York (1986).

[70] C. Zicovich-Wilson and R. Dovesi, “On the use of Symmetry Adapted Crystalline Orbitals in SCF-LCAO periodic calculations.
I. The construction of the Symmetrized Orbitals”, *Int. J. Quantum Chem* **67**, 299–310 (1998).

[71] C. Zicovich-Wilson and R. Dovesi, “On the use of Symmetry Adapted Crystalline Orbitals in SCF-LCAO p eriodic calculations.
II. Implementation of the Self-Consistent-Field scheme and examples”, *Int. J. Quantum Chem.* **67**, 311–320 (1998).

[72] M.F. Guest and V.R. Saunders, “On methods for converging open-shell Hartree-Fock wave-functions”, *Mol. Phys.* **28**, 819–828 (1974).

[73] E. Clementi, *MOTECC 91*, ESCON, Leiden (1991).

[74] C.C.J. Roothaan, “Self-Consistent Field Theory for Open Shells of Electronic Systems”, *Rev. Mod. Phys.* **32**, 179–185 (1960).

[75] N.D. Mermin, “Thermal properties of the inhomogeneous electron gas”, *Phys. Rev. A* **137**, 1441 (1965).

[76] E. Aprà, *Metodi quanto-meccanici per lo studio di solidi cristallini*, tesi di dottorato, University of Torino, Torino (1993).

[77] J.A Pople and R.K. Nesbet, “Self-consistent orbitals for radicals”, *J. Chem. Phys.* **22**, 571 (1954).

[78] C.G. Broyden, *J. Inst. Math. Appl.* **6**, 76 (1970).

[79] R. Fletcher, *Comput. J* **13**, 317 (1970).

[80] D. Goldfarb, *Math. Comput.* **24**, 23 (1970).

[81] D.F. Shanno, *Math. Comput.* **24**, 647 (1970).

[82] H. B. Schlegel, "Optimization of equilibrium and transition structures", *J. Comp. Chem.* **3**, 214–218 (1982).

[83] R. Lindh, A. Bernhardsson, G. Katlström, P.A. Malmqvist, "On the use of a hessian model function in molecular geometry optimizations", *Chem. Phys. Letters* **241**, 423–428 (1995).

[84] H.B. Schlegel, "Estimating the Hessian for gradient-type geometry optimizations", *Theoret. Chim. Acta* **66**, 333–340 (1984).

[85] J.M. Wittbrodt and H.B. Schlegel, "Estimating stretching force constants for geometry optimization", *J. Mol. Structure - Theochem* **398-399**, 55–61 (1997).

[86] M.J.D Powell, Academic Press, New York (1982).

[87] B. Lindberg, *J. Chem. Phys.* **88**, 3805–xxxx (1988).

[88] J.A. Weil, J.R. Bolton and E. Wertz, *Electron ParamagneticResonance - Elementary Theory and Practical Applications*, John Wiley, New York (1994).

[89] C. Pisani, E. Aprà, and M. Causà, "Density matrix of crystalline systems. I. Long range behavior and related computational problems", *Int. J. Quantum Chem.* **38**, 395 (1990).

[90] C. Pisani, E. Aprà, M. Causà and R. Orlando, "Density matrix of crystalline systems.
II. The influence of structural and computational parameters", *Int. J. Quantum Chem.* **38**, 419 (1990).

[91] C.M. Zicovich-Wilson, R. Dovesi and V.R. Saunders, "A general method to obtain well localized Wannier functions for composite energy bands in LCAO periodic calculations", *J. Chem. Phys.* **115**, 9708–9718 (2001).

[92] J. Pipek and P. G. Mezey, *J. Chem. Phys.*, 4916 (1989).

[93] V. Magnasco and A. Perico, *J. Chem. Phys.* **47**, 971 (1967).

[94] S. Casassa, C. M. Zicovich-Wilson, C. Pisani, "Symmetry-adapted localized Wannier functions suitable for periodic calculations", *Theor. Chem. Acc.* **116**, 726–733 (2006).

[95] K.H. Lee, M. Causà and S.S. Park, "Ab initio periodic Hartree-Fock calculations for interpretation of the scanning tunneling microscope (STM) images of graphite", *J. Phys. Chem. B* **102**, 6020 (1998).

[96] J. Tersoff and D.R. Hamann, "Theory of scanning tunneling microscope", *Phys. Rev. B* **31**, 805 (1985).

[97] V.R. Saunders, C. Freyria-Fava, R. Dovesi, L. Salasco and C. Roetti, "On the electrostatic potential in crystalline systems where the charge density is expanded in Gaussian functions", *Mol. Phys.* **77**, 629 (1992).

[98] V.R. Saunders, C. Freyria-Fava, R. Dovesi and C. Roetti, "On the electrostatic potential in linear periodic polymers", *Comp. Phys. Comm.* **84**, 156 (1993).

[99] R. Resta, "The geometric phase approach to macroscopic polarization in crystalline dielectrics", *Rev. Mod. Phys.* **66**, 809 (1994).

[100] S. Dall'Olio, R. Dovesi, R. Resta, "Spontaneous polarization as a Berry phase of the Hartree-Fock wave function: the case of $KNbO_3$", *Phys. Rev.* **B56**, 10105–10114 (1997).

[101] W.J. Hehre, R.F. Stewart and J.A. Pople, *J. Chem. Phys.* **51**, 2657 (1969).

[102] W.J. Hehre, R. Ditchfield, R.F. Stewart and J.A. Pople, *J. Chem. Phys.* **52**, 2769 (1970).

[103] W.J. Pietro, B.A. Levi, W.J. Hehre, R.F. Stewart, *Inorg. Chem.* **19**, 2225 (1980).

[104] W.J. Pietro, E.S. Blurock, R.F. Hout, W.J. Hehre, W.J. DeFrees and R.F. Stewart, *Inorg. Chem.* **20**, 3650 (1981).

[105] J.S. Binkley, J.A. Pople and W.J. Hehre, *J. Am. Chem. Soc.* **102**, 939 (1980).

[106] M.S. Gordon, J.S. Binkley, J.A. Pople, W.J. Pietro and W.J. Hehre, *J. Am. Chem. Soc.* **104**, 2797 (1982).

[107] R. Dovesi, M. Causà, R. Orlando, C. Roetti and V.R. Saunders, "Ab initio approach to molecular crystals: a periodic Hartree-Fock study of crystalline urea", *J. Chem. Phys.* **92**, 7402 (1990).

[108] P.C. Hariharan and J.A. Pople, *Theor. Chim. Acta* **28**, 213 (1973).

[109] R. Orlando, R. Dovesi, C. Roetti and V.R. Saunders, "Ab initio Hartree-Fock calculations of periodic compounds: application to semiconductors", *J. Phys. Condens. Matter* **2**, 7769 (1990).

[110] M. Causà, R. Dovesi and C. Roetti, "Pseudopotential Hartree-Fock study of seventeen III-V and IV-IV semiconductors", *Phys. Rev. B* **43**, 11937 (1991).

[111] L. Salasco, R. Dovesi, R. Orlando, C. Pisani, M. Causà and V.R. Saunders, "A Periodic ab initio extended basis set study of α-Al_2O_3", *Mol. Phys.* **72**, 267 (1992).

[112] R. Dovesi, E. Ferrero, C. Pisani and C. Roetti, "Ab initio study of the electron momentum distribution of metallic lithium", *Z. Phys. B* **51**, 195 (1983).

[113] V.R. Saunders, "*Ab initio* Hartree-Fock calculations for periodic systems", *Faraday Symp. Chem. Soc.* **19**, 79–84 (1984).

[114] R. Dovesi, C. Pisani, C. Roetti and V.R. Saunders, "Treatment of Coulomb interactions in Hartree-Fock calculations of periodic systems", *Phys. Rev. B* **28**, 5781 (1983).

[115] M. Causà, R. Dovesi, R. Orlando, C. Pisani and V.R. Saunders, "Treatment of the exchange interactions in Hartree-Fock LCAO calculations of periodic systems", *J. Phys. Chem.* **92**, 909 (1988).

[116] R. Dovesi, "On the role of symmetry in the ab initio Hartree-Fock linear combination of atomic orbitals treatment of periodic systems", *Int. J. Quantum Chem.* **29**, 1755 (1986).

[117] V.R. Saunders et al., *Electronic structure theory: from molecules to crystals*, private communication (1999).

[118] M. Catti, R. Dovesi, A. Pavese and V. R. Saunders, "Elastic constants and electronic structure of fluorite (CaF_2): an ab initio Hartree-Fock study", *J. Phys.: Condens. Matter* **3**, 4151 (1991).

[119] R. C. Weast and M. J. Astle and W. H. Beyer. *CRC Handbook of Chemistry and Physics*, A–61 and A–62. CRC Press, Inc., Boca Raton, Florida (, 1986-87).

[120] V. R. Saunders, "From molecules to solids", Book in preparation.

[121] M. Abramovitz and I. Segun, *Handbook of Mathematical Functions*, Dover, New York (1965).

Index

Keyword Index

Subject Index