CRYSTAL23

CRYSTAL23

User's Manual

December 1, 2022

R. Dovesi,[1] V.R. Saunders,[1] C. Roetti,[1] R. Orlando,[1] C. M. Zicovich-Wilson,[2]
F. Pascale,[3] B. Civalleri,[1] K. Doll,[4] N.M. Harrison,[5,6] I.J. Bush,[7]
Ph. D'Arco,[8] M. Llunell,[9] M. Causà,[10] Y. Noël,[8] L. Maschio,[1]
A. Erba,[1] M. Rérat,[11] S. Casassa,[1] B.G. Searle,[5] J.K. Desmarais[1]

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

[2] Centro de Investigación en Ciencias, Universidad Autónoma del Estado de Morelos,
Av. Universidad 1001, Col. Chamilpa, 62209 Cuernavaca (Morelos) Mexico

[3] Faculté des Sciences et Technologies, Université de Lorraine
BP 70239, Boulevard des Aiguillettes 54506 Vandoeuvre-lés-Nancy Cedex, France

[4] Institut für Elektrochemie, Universität Ulm
Albert-Einstein-Allee 47, 89081 Ulm, Germany

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

[6] Department of Chemistry, Imperial College
South Kensington Campus, London, U.K.

[7] STFC Rutherford Appleton Laboratory, Chilton Didcot, Oxfordshire, OX11 0QX, UK

[8] Institut des Sciences de la Terre de Paris (UMR 7193 UPMC-CNRS),
UPMC, Sorbonne Universités, 4 Place Jussieu, 75232 Paris CEDEX 05, France

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

[10] Dipartimento di Ingegneria Chimica, dei Materiali e della Produzione industriale,
Università di Napoli "Federico I"
Via Cintia (Complesso di Monte S. Angelo) 21, Napoli - Italy

[11] Equipe de Chimie Physique, IPREM UMR5254,
Université de Pau et des Pays de l'Adour, 64000 Pau, France

Contents

Chapter 1

Introductory Remarks

1.1 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 18.

The local functions are, in turn, linear combinations of Gaussian type functions (GTF) whose exponents and coefficients are defined by input (section 2.2). Functions of symmetry *s*, *p*, *d*, *f*, *g* can be used (see page 22). 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 can treat helical structures (translation followed by a rotation around the periodic axis).

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

Input tools allow the generation of slabs (2D system), nano-rods (1D 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 3.1

Specific input options allow generation of special 1D (single and multi-wall nanotubes) and 0D (fullerenes) structures from 2D ones.

Previous releases of the software in 1988 (CRYSTAL88, [77]), 1992 (CRYSTAL92, [80]), 1996 (CRYSTAL95, [81]), 1998 (CRYSTAL98, [247]), 2003 (CRYSTAL03, [248]), 2006 (CRYSTAL06, [82]), 2010 (CRYSTAL09, [75]), 2013 (CRYSTAL14, [76]) and 2017 (CRYSTAL17, [83]) have been used in a wide variety of research applications.

The required citation for this work is:

A. Erba, J. K. Desmarais, S. Casassa, B. Civalleri, L. Donà, I. J. Bush, B. Searle, L. Maschio, L. Edith-Daga, A. Cossard, C. Ribaldone, E. Ascrizzi, N. L. Marana, J.-P. Flament and B. Kirtman, *J. Chem. Theory Comput.*, DOI: 10.1021/acs.jctc.2c00958 (2022) *CRYSTAL23: A Program for Computational Solid State Physics and Chemistry*

CRYSTAL23 output will display the references relevant to the property computed, when citation is required.

Updated information on the CRYSTAL code as well as tutorials to learn basic and advanced CRYSTAL usage are in:

http://www.crystal.unito.it

1.2 CRYSTAL23 Program Features

New features with respect to CRYSTAL17 are in ***bold-italics***.

Hamiltonian

- **Hartree-Fock Theory**
  - Restricted (RHF)
  - Unrestricted (UHF)
  - Restricted-Open (ROHF)
  - ***Generalized (GHF), i.e. for a two-component spinor basis***
- **Density Functional Theory**
  - Semilocal functionals: local [L], gradient-corrected [G] and meta-GGA (tau-dependent) [T]
  - Collinear Spin DFT
  - ***Non-Collinear Spin DFT***
  - ***Spin-Current DFT (SCDFT)***
- Hybrid HF-DFT functionals
  - Global Hybrids: B3PW, B3LYP (using the VWN5 functional), PBE0, PBESOL0, B1WC, WC1LYP, B97H
  - Range-Separated Hybrids:
    - Screened-Coulomb (SC): HSE06, HSEsol, SC-BLYP
    - Middle-range Corrected (MC): HISS
    - Long-range Corrected (LC): LC-ωPBE, LC-ωPBEsol, LC-ωBLYP, ωB97, ωB97-X, RSHXLDA, LC-BLYP, CAM-B3LYP
  - Self-consistent Global Hybrid functionals (sc-hyb)
- Minnesota semilocal and hybrid functionals (mGGA):
  - M05 family: M05, M05-2X
  - M06 family: M06, M06-2X, M06-HF, M06-L
  - ***revised M06 functionals: revM06, revM06-L***
  - ***MN15 family: MN15, MN15-L***
  - ***SCAN and r^{2}SCAN functionals***
  - ***B95-based hybrid functionals: B1B95, mPW1B95, mPW1B1K, PWB6K, PW6B95***
- User-defined hybrid functionals
- London-type empirical correction for dispersive interactions (DFT-D2 scheme)
- DFT-D3 correction for dispersive interactions. Automated, parameter-free implementation
- Grimme's geometrical CounterPoise (gCP) empirical correction to remove the BSSE
- Composite methods for molecular crystals (HF3c, PBEh3c, HSE-3c, B97-3c)
- ***Revised composite methods for solid state calculations (HFsol-3c, PBEsol0-3c, HSEsol-3c)***

Analytical Derivatives

- **Analytical first energy derivatives with respect to the nuclear coordinates and cell parameters**
- ***New Integral Engine for Faster Calculation of Analytical Energy Derivatives***
  - Hartree-Fock and Density Functional methods
  - All-electron Basis Sets and Effective Core Pseudo-potentials
- **Analytical energy derivatives, up to fourth order, with respect to an applied electric field (CPHF/CPKS)**
  - Dielectric tensor, polarizability (linear-optical properties)
  - First Hyper-polarizability (non linear-optical property)
  - Second-Harmonic Generation
  - Pockels Effect
  - Second Hyper-polarizability (non linear-optical property).
  - ***Extended to HJS-based range-separated hybrid functionals (e.g. HSE06, HISS, LC-ωPBE)***
- **Mixed analytical derivatives with respect to an applied electric field and either a nuclear displacement or a cell distortion (CPHF/CPKS)**
  - Born-charge tensor
  - Raman polarizability tensors
  - Direct Piezoelectric tensor (electronic term)
- ***Derivatives of the Electron Density, up to fourth order, for f- and g-type AOs***

Type of Calculation

- **One component single-point energy calculation**
  - Broyden Convergence Accelerator
  - Anderson Convergence Accelerator
  - Fock Matrix-mixing Scheme
  - DIIS Convergence Accelerator
- ***Two-component single-point energy calculation***
  - ***Self-consistent Treatment of Spin-Orbit Coupling (SOC)***
  - ***Non-Collinear Initial Guess for the Magnetization***
  - ***Fock Matrix-mixing Scheme***
- **Geometry optimizations**
  - Uses a quasi-Newton algorithm
  - ***Extension of Model Initial Hessian to Lanthanides and Actinides***
  - Optimizes in symmetry-adapted Cartesian coordinates
  - Optimizes in redundant coordinates
    - New internal coordinates handling and algorithm for back-transformation

  - Full geometry optimization (cell parameters and atom coordinates)
  - Freezes atoms during optimization
  - Constant volume or pressure constrained geometry optimizations (3D only)
  - Transition state search

- **Harmonic Vibrational Frequencies**
  - Harmonic vibrational frequencies at Gamma point
  - Phonon dispersion using a direct approach (efficient supercell scheme)
  - Phonon band structure
  - Calculation of Atomic Displacement Parameters and Debye-Waller factors
  - IR intensities through localized Wannier functions and Berry phase
  - IR and Raman intensities through CPHF/CPKS analytical approach
  - Simulated reflectance, IR and Raman spectra
  - Vibrational contribution to dielectric tensor
  - Vibrational contribution to first-hyper-polarizability
  - Exploration of the energy and geometry along selected normal modes
  - Total and Projected Vibrational Density-of-States (VDOS)
  - ***Neutron-weighted VDOS for Inelastic Neutron Scattering - extended to heavy elements***

- **Anharmonic frequencies for X-H bonds**

- ***Anharmonic Vibrational Frequencies***
  - ***Development of the potential energy surface including up to fourth-order force constants***
  - ***Vibrational Self-Consistent Field (VSCF) and Vibrational Configuration Interaction (VCI) Treatments***

- **Automated calculation of the elastic tensor of crystalline systems**
  - Generalized to 1D and 2D systems
  - Calculation of directional elastic wave velocities
  - Calculation of isotropic polycrystalline aggregates elastic properties via Voigt-Reuss-Hill scheme
  - Elastic Tensor under Pressure
  - Complete Analysis of Elastic wave velocities through AWESOME Code
  - Nuclear-relaxation Term through Internal-strain Tensor
  - ***Thermo-Elasticity with quasi-static and quasi-harmonic schemes***

- **Automated E vs V calculation for Equation of State (EOS)**
  - Murnaghan, Birch-Murnaghan, Vinet, Poirer-Tarantola and polynomial
  - Automated calculation of pressure dependence of volume and bulk modulus

- **Automated Quasi-harmonic Approximation (QHA) for Thermal Properties**
  - Volume-dependent Thermodynamic properties
  - Lattice Thermal Expansion (anisotropic)
  - P-V-T Equation-of-State

  - Constant-pressure thermodynamic properties
  - Temperature dependence of Bulk modulus (isothermal and adiabatic)
  - Grüneisen Parameters
  - ***QHA generalized to 1D and 2D systems***
  - ***Thermo-Elasticity with quasi-static and quasi-harmonic schemes***

- **Automated calculation of piezoelectric and photoelastic tensors**
  - Direct and converse piezoelectricity (using the Berry phase approach)
  - Elasto-optic tensor through the CPHF/CPKS scheme
  - Electric field frequency dependence of photoelastic properties
  - Nuclear-relaxation Term of Piezoelectric Tensor through Internal-strain Tensor
  - Piezo-optic fourth-rank Tensor
  - Analytical Piezoelectric Tensor through CPHF/KS Scheme

- **Improved tools to model solid solutions**
  - Automated Generation of Symmetry-independent Configurations
  - Automated algorithm for computing the energy (with or without geometry optimization) of selected configurations

Basis set

- **Gaussian type functions basis sets**
  - s, p, d, f and GTFs
  - ***Extension of the LCAO Approach to g-type AOs***
  - Standard Pople Basis Sets
    - STO-nG n=2-6 (H-Xe), 3-21G (H-Xe), 6-21G (H-Ar)
    - polarization and diffuse function extensions
  - Internal library of basis sets with simplified input
  - ***Internal Libraries for POB-TZVP Consistent Basis Sets for Most Elements of the Periodic Table***
  - ***New Basis Sets for Lanthanides and Actinides with f Electrons in the valence***
  - User-defined External Library supported
  - User-specified basis sets supported
  - ***Internal Basis Set Optimizer***
  - ***Perturbation theory enrichment of the basis set***

- **Pseudopotential Basis Sets**
  - Internal Libraries for AREP only:
    - Hay-Wadt large core
    - Hay-Wadt small core
    - Durand and Barthelat
  - ***Internal Libraries for AREP and SOREP:***
    - ***Columbus large core***
    - ***Columbus small core***
    - ***Stuttgart-Cologne large core***
    - ***Stuttgart-Cologne small core***
  - User-defined pseudopotential basis sets supported

Periodic systems

- **Periodicity**
  - Consistent treatment of all periodic systems
  - 3D - Crystalline solids (230 space groups)
  - 2D - Films and surfaces (80 layer groups)
  - 1D - Polymers
    - space group derived symmetry (75 rod groups)
    - helical symmetry (up to order 48)
  - 1D - Nanotubes (with any number of symmetry operators)
  - ***1D - Multi-wall Nanotubes***
  - 0D - Molecules (32 point groups)
  - 0D - Fullerenes
- **Automated geometry editing**
  - 3D to 2D - slab parallel to a selected crystalline face (hkl)
  - 3D to 0D - cluster from a perfect crystal (H saturated)
  - 3D to 0D - extraction of molecules from a molecular crystal
  - 3D to n3D - supercell creation
  - ***2D to 1D - building single and multi-wall nanotubes from a single-layer slab model***
  - 2D to 0D - building fullerenes from a single-layer slab model
  - 3D to 1D, 0D - building nanorods and nanoparticles from a perfect crystal
  - 2D to 0D - construction of Wulff's polyhedron from surface energies
  - Several geometry manipulations (reduction of symmetry; insertion, displacement, substitution, deletion of atoms)

Wave function analysis and properties

- **Band structure**
- **Density of states**
  - Band projected DOSS
  - AO projected DOSS
- **All Electron Charge Density - Spin Density**
  - Density maps
  - Mulliken population analysis
  - Density analytical derivatives
  - Hirshfeld-I Partitioning Scheme
  - ***Extended to f and g GTFs***
  - ***Extension of the TOPOND module for QTAIM to f and g AOs***
- ***Collinear and Non-collinear Magnetization density maps***
- ***Orbital-current density maps***
- ***Spin-current density maps***

- **Electronic Transport Properties**
- **Atomic multipoles**
- **Electric field**
- **Electric field gradient**
- **Static structure factors and dynamic structure factors including Debye-Waller factor**
  - XRD Spectrum
- **Electron Momentum Density and Compton profiles**
  - Electron momentum density maps
  - Automated anisotropy maps
  - Partitioning according to Wannier functions
- **Electrostatic potential and its derivatives**
  - Quantum and classical electrostatic potential and its derivatives
  - Electrostatic potential maps
- **Fermi contact**
- **Localized Wannier Functions (Boys method)**
- **Mossbauer effect (isotropic effect and quadrupolar interaction)**
- **Topological analysis of the electron charge density via the TOPOND package, fully integrated in the program**
  - ***generalized to f and g GTFs***

Software performance

- **Memory management: dynamic allocation**
- **Full parallelization of the code**
  - Parallel SCF and gradients for both HF and DFT methods
  - Replicated data version (MPI) with parallelization strategy on IRREPs
  - Massive parallel version (MPI) (distributed memory) (Improved version: lower memory usage and better scaling)
  - Parallel (replicated data) version of the "properties" module
  - ***OpenMP+MPI Hybrid Parallelism for SCF and Forces***

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

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 37): `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 not end with 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. See page 21. The same conventional atomic number must be given in geometry input and in basis set input.

Acknowledgments

Embodied in the present code are elements of programs distributed by other groups.
In particular: the atomic SCF package of Roos et al. [240], the GAUSS70 gaussian integral package and STO-nG basis set due to Hehre et al. [143], the code of Burzlaff and Hountas for space group analysis [41], Saunders' ATMOL gaussian integral package [190], the XCFun DFT library of exchange-correlation functionals [86], and the SOREP integral routines of Pitzer and Winter [229], which were also implemented in the EPCISO program [277].
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.

We are in debt with Cesare Pisani, who first conceived the CRYSTAL project in 1976, for his constant support of and interest in the development of the new version of the CRYSTAL program.

It is our pleasure to thank Piero Ugliengo, Marta Corno and Stefano Pantaleone for continuous help, useful suggestions, rigorous testing.

We thank Giuseppe Mallia for useful contribution to test parallel execution and to develop automatic testing procedures.

We kindly acknowledge Jorge Garza-Olguin for his invaluable help in testing and documenting the compilation of parallel executables from object files.

Contribution to the development of the current release has been given by: Matteo Ferrabone, Marco De La Pierre, Mauro Ferrero, Jacopo Baima, Elisa Albanese, Michael F. Peintinger, Radovan Bast, Bernard Kirtman, Valentina Lacivita, Raffaella Demichelis, Jan Gerit Brandenburg, Simone Salustro, Giuseppe Sansone, Micheal T. Ruggiero, Loredana Edith Daga, Lorenzo Doná, Alessandro Cossard, Jean-Pierre Flament.

An invaluable help in testing the code and validating the new features is acknowledged to Lorenzo Doná, Alessandro Cossard, Chiara Ribaldone, Eleonora Ascrizzi, Filippo Bodo, Naiara Marana, Loredana Edith Daga.

Specific contribution to coding is indicated in the banner of the new options.

Getting Started

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

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 CRYSTAL23. 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-support@unito.it).

Chapter 2

Wave-function Calculation: Basic Input Route

2.1 Geometry and symmetry information

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

CRYSTAL	3D system	page 15
SLAB	2D system	page 15
POLYMER	1D system	page 15
HELIX	1D system - roto translational symmetry	page 16
MOLECULE	0D system	page 15
EXTERNAL	geometry from external file	page 16
DLVINPUT	geometry from DLV [257] Graphical User Interface.	page 16

Four input schemes are used.

The first is for crystalline systems (3D), and is specified by the keyword **CRYSTAL**.

The second is for slabs (2D), polymers (1D) and molecules (0D) as specified by the keywords **SLAB**, **POLYMER** or **MOLECULE** respectively.

The third scheme (keyword **HELIX**) defines a 1D system with roto-translational symmetry (helix).

In the fourth scheme, with keyword **EXTERNAL** (page 16) or **DLVINPUT**, the unit cell, atomic positions and symmetry operators may be provided directly from an external file (see Appendix D, page 446). Such an input file can be prepared by the keyword **EXTPRT** (***crystal*** input block 1, page 44; ***properties***).

Sample input decks for a number of structures are provided in section 16.1, page 368.

Geometry input for crystalline compounds. Keyword: CRYSTAL

rec	variable	value	meaning
• ∗	IFLAG		convention for space group identification (Appendix A.1, page 413):
		0	space group sequential number(1-230)
		1	Hermann-Mauguin alphanumeric code
	IFHR		type of cell: for rhombohedral groups, subset of trigonal ones, only (meaningless for non-rhombohedral crystals):
		0	hexagonal cell. Lattice parameters *a,c*
		1	rhombohedral cell. Lattice parameters *a,* α
	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],		**minimal** set of crystallographic cell parameters:
	[α],[β]		translation vector[s] length [Ångstrom],
	[γ]		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. Keywords: SLAB, POLYMER, MOLECULE

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, fractional (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 420) 1D - polymers (Appendix A.3, page 417) 2D - slabs (Appendix A.2, page 416)
			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 3
		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 for polymers with roto translational symmetry. Keyword: HELIX

rec		variable	meaning
•	∗	N1	order of rototranslational axis
	∗	N2	to define the rototranslational vector
•	∗	a_0	lattice parameter of 1D cell - length in Ångstrom
•	∗	NATR	number of non-equivalent atoms in the asymmetric unit
			insert NATR records II
•	∗	NAT	conventional atomic number 3
		X,Y,Z	atoms coordinates. Unit of measure: 1D - polymers : *y,z* in Ångstrom, *x* in fractional units of crystallographic cell translation vector
			optional keywords terminated by **END or STOP** **II**

A helix structure is generated: each atom of the irreducible part is rotated by an angle $\beta = n \cdot 360/\mathrm{N1}$ degrees and translated by a vector $\vec{t} = n \cdot a_0 \frac{N2}{N1}$ with $n = 1, ...(N1 - 1)$.

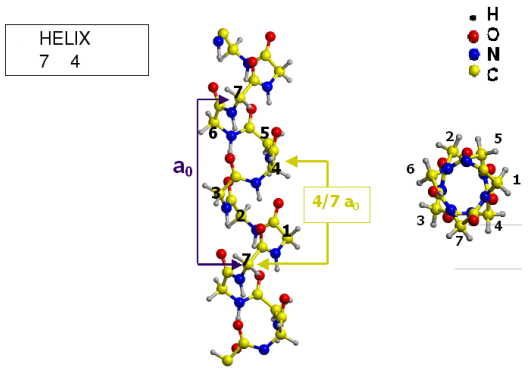
As an example let's consider the α-helix conformer of polyglycine whose structure is sketched in Figure 2.1.

The helix structure is characterized by seven glycine residues per cell. The order of the rototranslational axis is therefore seven, $N1 = 7$. In order to establish the value of $N2$, look for instance at the atom labeled 7 in the Figure. The top view of the helix shows that upon rotation by $\beta = 360/7$ degrees, atom 7 moves to atom 4; the side view clarifies that this movement implies a translational vector $\vec{t} = a_0 \frac{4}{7}$: therefore $N2 = 4$.

Geometry input from external geometry editor. Keywords: EXTERNAL, DLVINPUT

The fourth input scheme works for molecules, polymers, slabs and crystals. The complete geometry 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 D, page

Figure 2.1: Side view (left) and top view (right) of an α-helix conformer of polyglycine



446. 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 44; program ***properties***), or by the the visualization software DLV (*http://www.cse.scitech.ac.uk/cmg/DLV/*).
The geometry defined by **EXTERNAL** can be modified by inserting any geometry editing keyword (page 30) 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 47) or **BOHR** (page 38).

2. The geometry of a system is defined by the crystal structure ([118], Chapter 1 of ref. [220]). Reference is made to the International Tables for Crystallography [133] 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
trigonal	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 421, and are taken from Table 5.1 of the International Tables for Crystallography [133].

   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 70, 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 different settings, the same position can have different multiplicity. For instance, in space group 227 (diamond, silicon) the position (0., 0., 0.) has multiplicity 2 in 1st setting, and multiplicity 4 in 2nd setting.

   Second setting is the default choice in CRYSTAL.

   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 59) 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 $\overline{3}$. See Appendix A.1, page 413.

    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. Rhombohedral space groups are a subset of trigonal ones. The Hermann-Mauguin symbol must begin by 'R'. For instance, space groups 156-159 are trigonal, but not rhombohedral (their Hermann-Mauguin symbols begin by "P"). Rhombohedral space groups (146-148-155-160-161-166-167) may have an hexagonal cell (a=b; c; α, $\beta = 90^0$; $\gamma = 120^0$: input parameters a,c) or a rhombohedral cell (a=b=c; $\alpha = \beta = \gamma$: input parameters = a, α). See input datum IFHR.

14. 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.

15. The conventional atomic number NAT is used to associate a given basis set with an atom (see Basis Set input, Section 2.2, page 20). 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 3.2, page 79).

    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 [10], 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 77).

16. The keyword **SLABCUT** (Geometry editing input, page 66) 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.

17. 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.

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

19. The keywords **MOLECULE** (for molecular crystals only; page 51) and **CLUSTER** (for any n-D structure; page 40) allow the creation of a non-periodic system (molecule(s) or cluster) from a periodic one.

2.2 Basis set

Two different methods are available to input basis set data:

- Standard route
- Basis set input by keywords

2.2.1 Standard route

rec	variable	value	meaning
• ∗	NAT	n	conventional atomic 3 number
		<200> 1000	all-electron basis set (Carbon, all electron BS: 6, 106, 1006)
		>200	valence electron basis set (Carbon, ECP BS: 206, 306) . ECP (Effective Core Pseudopotential) must be defined (page 79)
		=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 79)	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 (F shell)
		5	9 g AOs (G shell) - CRYSTAL22
	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
			LAT=4 → g function coefficient - CRYSTAL22
	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 17. When an effective core pseudo-potential is used, the basis set **must** be optimized with reference to that potential (Section 3.2, page 79).

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 37), the corresponding basis set input can remain in the input stream. The keyword **GHOSTS** (page 77) removes the atom, leaving the associated basis set.

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 18, page 395).

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=Z+1000.

   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 3.2, page 79, for information on core pseudo-potentials.

   Suppose we have four non-equivalent carbon atoms in the unit cell. Conventional atomic numbers 6 106 1006 206 306 mean that carbon atoms (real atomic number 6) unrelated by symmetry are to be associated with different basis sets: the first tree (6, 106, 1006) 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.
   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, g* functions (given in input);
   b) standard Pople basis sets [144] (internally stored as in Gaussian 94 [112]).
      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	$(2z^2 - 3x^2 - 3y^2)z, (4z^2 - x^2 - y^2)x, (4z^2 - x^2 - y^2)y,$ $(x^2 - y^2)z, xyz, (x^2 - 3y^2)x, (3x^2 - y^2)y$
5	G	9	$3x^4 + 6x^2y^2 - 24x^2z^2 + 3y^4 - 24y^2z^2 + 8z^4, (4z^2 - 3x^2 - 3y^2)xz,$ $(4z^2 - 3x^2 - 3y^2)yz, 6x^2z^2 - x^4 + y^4 - 6y^2z^2, (6z^2 - x^2 - y^2)xy$ $(x^2 - 3y^2)xz, (3x^2 - y^2)yz, x^4 - 6x^2y^2 + y^4, (x^2 - y^2)xy$

   G shells can only be used as polarization functions. They cannot be occupied in the guess.

   When symmetry adaptation of Bloch functions is active (default; NOSYMADA in block3 to remove it), if F (and G) functions are used, all lower order functions must be present (F, 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 14: Mulliken population analysis (**PPAN**, page 117), electrostatic multipoles (**POLI**, page 345), projected density of states (**DOSS**,page 315) and to provide an input for some options (**EIGSHIFT**, input block 3, page 99).

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

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

7. Spherical harmonics g-shells consisting of 9 AOs are used.

8. 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$>$2, 2 for s and (CHE$-$2) for p functions, if CHE$\leq$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 may be $\neq 0$ in CRYSTAL09.
5	G	18.	CHE for g functions - it must be 0

9. A maximum of one open shell for each of the s, p, d and/or f 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 $p^m d^n$, $p^m f^l$, or $d^n f^l$. Either m must equal 3 or n must equal 5, or l must equal 7 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.

10. 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.*

11. 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 110). When a different guess is required (**GUESSP**), the shell charges are not used, but checked for electron neutrality when the basis set is entered.

12. 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 [74]. This is obtained by entering the keyword **CHARGED** (page 75) after the standard basis set input. The value of total energy must be carefully checked.

13. 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 75).

14. 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 [48] (keyword **ATOMHF**, input block 3, page 93).

Users that perform calculations with G shells are kindly requested to cite the following paper [67]:

J.K. Desmarais, A. Erba and R. Dovesi, *Theor. Chem. Acc.*, **137**, 2 (2018). *Generalization of the periodic LCAO approach in the CRYSTAL code to g-type orbitals*

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

http://www.crystal.unito.it

2.2.2 Basis set input by keywords

A few predefined basis set data can be retrieved by simply typing a keyword. For the moment being the set of available basis sets includes (available atomic numbers in parentheses):

- Pople's STO-3G minimal basis set (1–53)
- Pople's STO-6G minimal basis set (1–36)
- POB double-ζ valence + polarization basis set for solid state systems (1–35, 49, 74)
- POB double-ζ valence basis set + a double set of polarization functions for solid state systems (1–35, 49, 83)
- POB triple-ζ valence + polarization basis set for solid state systems (1–35, 49, 83)
- POB-REV2 double-ζ valence + polarization basis set for solid state systems (1–35)

- POB-REV2 triple-ζ valence + polarization basis set for solid state systems (1–35, 37–53, 55–56, 72–84)

Features and performance of Peintinger-Oliveira-Bredow (POB) basis sets are illustrated in Ref. [176, 280, 165, 166].

In order to enable basis set input by keywords, the following keyword must replace the final keyword, **END**, of the structure input (input block 1):

BASISSET

This card must be followed by the selection of a basis set type. The following sets are presently available as internal basis sets:

Basis set label	Basis set type
CUSTOM	Standard input basis set: insert cards as illustrated in section 2.2.1
STO-3G	Pople's standard minimal basis set (3 Gaussian function contractions) [144]
STO-6G	Pople's standard minimal basis set (6 Gaussian function contractions) [144]
POB-DZVP	POB Double-ζ + polarization basis set [176]
POB-DZVPP	POB Double-ζ + double set of polarization functions [176]
POB-TZVP	POB Triple-ζ + polarization basis set [176]
POB-DZVP-REV2	POB-REV2 Double-ζ + polarization basis set [280]
POB-TZVP-REV2	POB-REV2 Triple-ζ + polarization basis set [280, 165, 166]

Input example for rock-salt:

```
NaCl Fm-3m ICSD 240598
CRYSTAL
0 0 0
225
5.6401
2
11 0.0 0.0 0.0
17 0.5 0.5 0.5
BASISSET
POB-TZVP
DFT
EXCHANGE
PWGGA
CORRELAT
PWGGA
HYBRID
20
CHUNKS
200
END
TOLINTEG
7 7 7 7 14
SHRINK
8 8
END
```

Basis set input from external file

If the basis set specified after the **BASISSET** keyword is not found in the internal basis set library (*i.e.* it is not one of those listed above), the program will search for the presence of a

file named BASISSETS.DAT in the temporary (scratch) working directory. If the file is found, it will search for a basis with that label in the file.
The BASISSETS.DAT file is an external basis set database file that can contain as many basis set and for as many elements as desired. The format of such file is a number of entries `#BASIS` followed by a basis set name label and the basis set in the usual CRYSTAL format, as it would appear in input. This means the block is closed by a "99 0" line, as in the Crystal input.

Warning: the BASISSETS.DAT *file does not currently support pseudopotentials, only all-electron basis sets can be used.*

Here's an example of one entry in the BASISSETS.DAT file, containing custom basis sets for H and Li atoms

```
#BASIS
mybasis
1 3
0 0 5 1.0 1.0
 120.0 0.000267
 40.0 0.002249
 12.8 0.006389
 4.0 0.032906
 1.2 0.095512
0 0 1 0.0 1.0
 0.5 1.0
0 0 1 0.0 1.0
 0.13 1.0
3 3
0 0 6 2. 1.
 700.0 0.001421
 220.0 0.003973
 70.0 0.01639
 20.0 0.089954
 5.0 0.315646
 1.5 0.494595
0 0 1 1. 1.
 0.5 1.0
0 2 1 0. 1.
 0.6 1.0
 99 0
```

An example .d12 input using this basis set would be:

```
lih bulk
CRYSTAL
0 0 0
225
4.0170
2
 3 0.    0.    0.
 1 0.5   0.5   0.5
BASISSET
mybasis
DFT
PBE0
END
SHRINK
8 8
END
```

In the output the reading from external basis file will be marked by the lines

```
WARNING **** MYBASIS **** SEARCHING FOR REQUESTED BASIS SET IN FILE BASISSETS.DAT
```

and

```
 Loading basis set from BASISSETS.DAT: MYBASIS
```

Dual basis set approaches

Two approaches are available in order to estimate the effect of a basis set enlargement without the need to converge the SCF in the large basis.

- The periodic adaptation of the Dual Basis approach by Martin Head-Gordon[169] (see keyword PROJDUAL, page 117). This approach yields only a correction to the total energy.
- An original perturbative approach that allows an estimate of the correction to the total energy and wavefunction (keywords MK1 and MK2, page114).

In both cases it is *strongly* suggested that the large basis set contains exactly the small one. Otherwise, suboptimal results might be obtained.

2.3 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	page
hamiltonian:	RHF	UHF (SPIN)	126
tolerances for coulomb and exchange sums :	6 6 6 6 12	TOLINTEG	125
Pole order for multipolar expansion:	4	POLEORDR	117
Max number of SCF cycles:	50	MAXCYCLE	114
Convergence on total energy:	10^{-6}	TOLDEE	125

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 [191]. The Hamiltonian matrix computed in direct space, $H_{\mathbf{g}}$, is Fourier transformed for each **k** value, and diagonalized, to obtain eigenvectors and eigenvalues:

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

$$H_k A_k = S k A_k E_k$$

A second shrinking factor, ISP, defines the sampling of k points, "Gilat net" [120, 119], 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 119).
In 3D crystals, the sampling points belong to a lattice (called the Pack-Monkhorst net), with basis vectors:

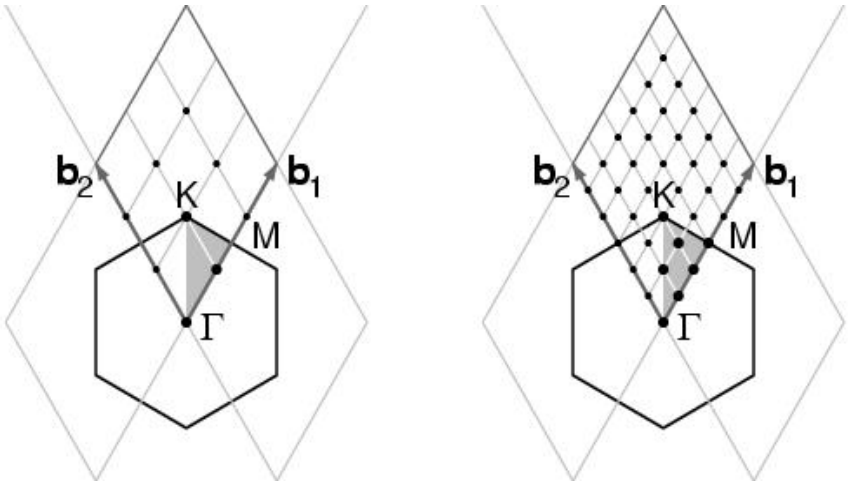
b1/is1, b2/is2, b3/is3 is1=is2=is3=IS, unless otherwise stated

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 18.7, page 401.
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.

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 [224], 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 117.

Some tools for accelerating convergence are given through the keywords **LEVSHIFT** (page 113 and tests 29, 30, 31, 32, 38), **FMIXING** (page 105), **SMEAR** (page 121), **BROYDEN**

(page 95) and **ANDERSON** (page 93).

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 125
TOLDEP (tolerance on SQM in density matrix elements) page **??**
MAXCYCLE (maximum number of cycles) page 114

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 124) and **BETALOCK** (page 94).

Chapter 3

Wave-function Calculation - Advanced Input Route

3.1 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 **TEST[GEOM]**.

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** 3.1 and **BREAKSYM** 3.1, 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	38	–
MAKESAED	printing of symmetry allowed elastic distortions (SAED)	49	–
PRSYMDIR	printing of displacement directions allowed by symmetry.	64	–
SYMMDIR	printing of symmetry allowed geom opt directions	73	–
SYMMOPS	printing of point symmetry operators	73	–
TENSOR	print tensor of physical properties up to order 4	73	I

Symmetry information and control			
BREAKELAS	symmetry breaking according to a general distortion	39	I
BREAKSYM	allow symmetry reduction following geometry modifications	39	–
KEEPSYMM	maintain symmetry following geometry modifications	49	–
MODISYMM	removal of selected symmetry operators	50	I
PURIFY	cleans atomic positions so that they are fully consistent with the group	64	–
SYMMREMO	removal of all symmetry operators	73	–
TRASREMO	removal of symmetry operators with translational components	73	–

Modifications without reduction of symmetry			
ATOMORDE	reordering of atoms in molecular crystals	36	–
NOSHIFT	no shift of the origin to minimize the number of symmops with translational components before generating supercell	59	–
ORIGIN	shift of the origin to minimize the number of symmetry operators with translational components	59	–
PRIMITIV	crystallographic cell forced to be the primitive cell	63	–
ROTCRY	rotation of the crystal with respect to the reference system cell	65	I

Atoms and cell manipulation - possible symmetry reduction (BREAKSYMM)			
ATOMDISP	displacement of atoms	36	I
ATOMINSE	addition of atoms	36	I
ATOMREMO	removal of atoms	37	I
ATOMROT	rotation of groups of atoms	37	I
ATOMSUBS	substitution of atoms	38	I
ELASTIC	distortion of the lattice	42	I
POINTCHG	point charges input	63	I
SCELCONF	generation of supercell for configuration counting	68	I
SCELPHONO	generation of supercell for phonon dispersion	68	I
SUPERCEL	generation of supercell - input refers to primitive cell	70	I
SUPERCON	generation of supercell - input refers to conventional cell	70	I
USESAED	given symmetry allowed elastic distortions, reads δ	73	I

From crystals to slabs (3D→2D)			
SLABINFO	definition of a new cell, with $xy \parallel$ to a given plane	67	I
SLABCUT	generation of a slab parallel to a given plane (3D→2D)	66	I

From slabs to single and multi-wall nanotubes (2D→1D)			
NANOTUBE	building a nanotube from a slab	54	I
SWCNT	building a nanotube from an hexagonal slab	72	I
NANOMULTI	building a multi-wall nanotube from a slab	57	I

From periodic structures to clusters			
CLUSTER	cutting of a cluster from a periodic structure (3D→0D)	40	I
CLUSTSIZE	maximum number of atoms in a cluster	41	I
FULLE	building a fullerene from an hexagonal slab (2D→0D)	47	I
HYDROSUB	border atoms substituted with hydrogens (0D→0D)	49	I

Molecular crystals

MOLECULE	extraction of a set of molecules from a molecular crystal (3D→0D)	51	I
MOLEXP	variation of lattice parameters at constant symmetry and molecular geometry (3D→3D)	51	I
MOLSPLIT	periodic structure of non interacting molecules (3D→3D)	52	–
RAYCOV	modification of atomic covalent radii	64	I
BSSE correction			
MOLEBSSE	counterpoise method for molecules (molecular crystals only) (3D→0D)	50	I
ATOMBSSE	counterpoise method for atoms (3D→0D)	36	I
Systematic analysis of crystal planes			
PLANES	Prints the possible crystal planes	63	I
Gibbs-Wulff construction			
WULFF	Building the Gibbs-Wulff polihedron	73	I
From crystals to nanorods (3D→1D)			
NANORODS	Building a nanorod from a crystal	53	I
From crystals to nanocrystals (3D→0D)			
NANOCRYSTAL	building a nanocrystal from a crystal	52	I
Auxiliary and control keywords			
ANGSTROM	sets input units to Ångstrom	35	–
BOHR	sets input units to bohr	38	–
BOHRANGS	input bohr to Å conversion factor (0.5291772083 default value)	38	I
BOHRCR98	bohr to Å conversion factor is set to 0.529177 (CRY98 value)	–	
END/ENDG	terminate processing of geometry input		–
FRACTION	sets input units to fractional	47	–
LATVEC	maximum number of classified lattice vectors	49	I
MAXNEIGHB	maximum number of equidistant neighbours from an atom	49	I
NEIGHBOR	number of neighbours in geometry analysis	59	I
PRINTCHG	printing of point charges coordinates in geometry output	63	
PRINTOUT	setting of printing options by keywords	64	–
SETINF	setting of inf array options	66	I
SETPRINT	setting of printing options	66	I
STOP	execution stops immediately	68	–
TESTGEOM	stop after checking the geometry input	73	–
Output of data on external units			
COORPRT	coordinates of all the atoms in the cell	42	–
EXTPRT	write file in CRYSTAL geometry input format	44	–
FINDSYM	write file in FINDSYM input format	47	–
STRUCPRT	cell parameters and coordinates of all the atoms in the cell	68	–
External electric field - modified Hamiltonian			
FIELD	electric field applied along a periodic direction	44	I
FIELDCON	electric field applied along a non periodic direction	46	I
Geometry optimization - see index for keywords full list			

OPTGEOM	Geometry optimization input block - closed by **END**	182	I
	Type of optimization (default: atom coordinates)		
FULLOPTG	full geometry optimization		–
CELLONLY	cell parameters optimization		–
INTREDUN	optimization in redundant internal coordinates	192	–
ITATOCEL	iterative optimization (atom/cell)		–
CVOLOPT	full geometry optimization at constant volume	197	–
	Initial Hessian		
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		–
HESSNUM	initial guess for the Hessian - numerical estimate		–
	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	200	I
RESTART	data from previous run		–
FINALRUN	Wf single point with optimized geometry		I
	Gradient calculation control		
NUMGRATO	numerical atoms first derivatives	191	–
NUMGRCEL	numerical cell first derivatives	191	–
NUMGRALL	numerical atoms and cell first derivatives	191	–
	External stress		
EXTPRESS	apply external hydrostatic pressure	202	I
	Printing options		
PRINTFORCES	atomic gradients		–
PRINTHESS	Hessian		–
PRINTOPT	optimization procedure		–
PRINT	verbose printing		–

Vibrational Frequencies - see index for keywords full list

FREQCALC	Harmonic Γ-frequencies calculation input - closed by **END**	207	I
	Normal modes analysis		
ANALYSIS		209	–
COMBMODE	TO combination modes and overtones	210	I

MODES	printing eigenvectors [default]	212	–
SCANMODE	scan geometry along selected modes	220	I
LO/TO splitting			
DIELISO	isotropic dielectric tensor	210	I
DIELTENS	anisotropic dielectric tensor	210	I
Vibrational spectrum simulation			
INTENS	intensities calculation active	215	–
INTCPHF	IR (and Raman) intensities via CPHF	217	I
INTLOC	IR intensities through Wannier functions	216	–
INTPOL	IR intensities through Berry phase [default]	215	–
INTRAMAN	Raman intensities calculation	217	I
IRSPEC	IR spectrum production	223	I
RAMSPEC	Raman spectrum production	225	I
Calculation control			
ECKART	Hessian freed by translations and rotations [default]	210	I
FRAGMENT	partial frequency calculation	211	I
ISOTOPES	isotopic substitution	211	I
NORMBORN	normalized Born tensor	212	–
NUMDERIV	technique to compute numerical 2nd derivatives	213	I
PRINT	verbose printing		–
RESTART	data from previous run		–
STEPSIZE	set size of cartesian displacements [0.003 Å]	213	I
TEST[FREQ]	frequency test run		–
USESYMM	full-symmetry exploitation at each point [default]		–
Phonon dispersion			
DISPERSION	frequencies calculated at $\vec{k} \neq \Gamma$ points	227	–
Thermodynamics			
ADP	anisotropic displacement parameters	213	I
PRESSURE	set pressure range	213	I
TEMPERAT	set temperature range		I

ANHARM	Anharmonic frequencies calculation input block - closed by **END**	232	I
ISOTOPES	isotopic substitution	233	I
KEEPSYMM	displace all symmetry equivalent atoms	**??**	–
NOGUESS		233	–
POINTS26	X-H distance varied 26 times around the equilibrium	233	–
PRINT	verbose printing		–
TEST[ANHA]	test run		–

Configurations counting and characterization

CONFCNT	configurations counting and cluster expansion	272	I
CONFRAND	symmetry-adapted uniform at random Monte Carlo	274	I
RUNCONFS	single-point calculations and geometry optimizations	275	I

CPHF - Coupled Perturbed Hartree-Fock 265

ELASTCON - Second order elastic constants 283

EOS - Equation of state 277

Geometry input optional keywords

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 (FRACTIONAL 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).

ANHARM - Anharmonic calculation of frequencies of X-H (X-D) bond stretching

See Chapter 8, page 232.

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 neighbors, 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.
Use keyword **CLUSTER** (41) to extend the number of atoms in the cluster if requested.
Warning. The system is 0D. No reciprocal lattice information is required in the **scf** input (Section 2.3, page 27).

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 39). To displace all the atoms symmetry related, **KEEPSYMM** must be inserted before **ATOMDISP**.
Increments are in Ångstrom, unless otherwise requested (keyword **BOHR**, **FRACTION**, page 35). 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 35). Remember that the original symmetry of the system is maintained, applying the symmetry operators to the added atoms if the keyword **KEEPSYMM** (page 39) was previously entered. The default is **BREAKSYM** (page 39). Attention should be paid to the neutrality of the cell (see **CHARGED**, page 75). 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 51) and **MOLSPLIT** (page 52). 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 75).
NB. The keyword **GHOSTS** (basis set input, page 77) 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 64).
			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 75.

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 42).

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 (**ATOMSPIN**, block34, page 94) 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 36, keyword **ATOMINSE**).

BREAKELAS (for 3D systems only)

This keyword breaks the symmetry of 3D sysems according to a general distortion (3x3 adimensional matrix, not necessarily symmetric):

rec	variable	value	meaning
•	∗	D11 D12 D13	first row of the matrix.
•	∗	D21 D22 D23	second row of the matrix.
•	∗	D31 D32 D33	third row of the matrix.

BREAKELAS can be used when the symmetry must be reduced to apply an external stress not compatible with the present symmetry.

BREAKELAS reduces the symmetry according to the distortion defined in input, but **does not perform a distortion of the lattice**.

Another possibility is when you compute elastic constants, and you want to fix a reference geometry with FIXINDEX. If your reference geometry has a symmetry higher than the distorted one, then you had to break the symmetry by applying e.g. a tiny elastic distortion with ELASTIC. By using BREAKELAS you can reduce the symetry without distortion of the lattice.

Example - Geometry optimization of MgO bulk, cubic, with an applied uniaxial stress modifying the symmetry of the cell.

```
TEST11 - MGO BULK
CRYSTAL
0 0 0
 225
4.21
2
 12 0.    0.    0.
  8 0.5   0.5   0.5
BREAKELAS    the number of symmops is reduced, from 48 to 16
0.001 0. 0.  the cell has a tetragonal symmetry now
0. 0. 0.
0. 0. 0.
OPTGEOM
FULLOPTG
0.001 0. 0.
0. 0. 0.
```

```
0. 0. 0.
ENDOPT
```

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 ≠ 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 ≠ 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
		$\neq 0$	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
		$\neq 0$	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 $\neq 0$ 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 $\neq 0$, 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.
Use keyword **CLUSTER** (41) to extend the number of atoms in the cluster if requested.
Warning. The system is 0D. No reciprocal lattice information is required in the **scf** input (Section 2.3, page 27). See test 16.

CLUSTSIZE - maximum size of a cluster

rec	variable	meaning
• ∗	NATOCLU	maximum number of atoms allowed in creating a cluster

This keyword sets a new limit to the maximum number of atoms allowed in a cluster created by keywords **CLUSTER**, **ATOMBSSE** and **MOLEBSSE**. Default value is equal to the number of atoms in the unit cell.

CONFCNT - Mapping of CRYSTAL calculations to model Hamiltonians

See Chapter 11.1, page 272.

CIFPRT

This keyword, to be inserted in the geometry input block, generates a file GEOMETRY.CIF where the structure of a 3D system is written in .cif format. Symmetry is not taken into account.

CIFPRTSYM

This keyword, to be inserted in the geometry input block, generates a file GEOMETRY.CIF where the structure of a 3D system is written in .cif format. Symmetry is taken into account.

COORPRT

Geometry information is printed: cell parameters, fractional coordinates of all atoms in the reference cell, symmetry operators.
A formatted file, "fort.33" , is written. See Appendix D, page 444. No input data are required. The file "fort.33" has the right format for the program **MOLDEN** [253] which can be downloaded from:
`www.cmbi.ru.nl/molden/molden.html`

CPHF - performs the Coupled Perturbed HF/KS calculation up to the second order

See Chapter 10.2, page 265.

ELASTCON - Calculation of elastic constants

See Chapter 13, page 283.

ELASTIC

An elastic deformation of the lattice may be defined in terms of the Z or ϵ strain tensors defined in section 18.9, page 404.

rec	variable	value	meaning
• ∗	IDEF	± 1	deformation through equation 18.49, Z matrix.
		± 2	deformation through equation 18.48: ϵ matrix.
		> 0	volume conserving deformation (equation 18.50).
		< 0	not volume conserving (equation 18.49 or 18.48).
• ∗	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 35). The suggested value for IDEF is -2 (deformation through equation 18.48, *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 407) elastic constants of Li_2O [79].

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 18.48
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 38).
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 file 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 "optc(a)xxx", being xxx the number of the cycle, and it is automatically written at each cycle.
Please refer to the standard script for running CRYSTAL09 to handle input/output file names. See Appendix D, page 439.

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::

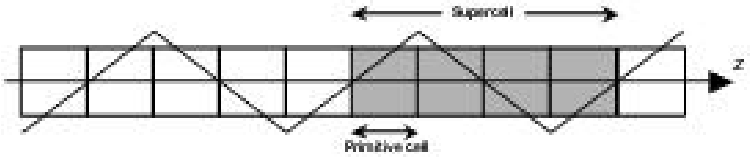
$$\hat{H} = \hat{H}_0 + \hat{H}_1(\vec{E}) \tag{3.1}$$

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 70).
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·$\vec{c}$, see fig. 3.1). 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.

Then, for computational reasons, an automatic rotation of the crystal in the cartesian reference

Figure 3.1: Triangular electric potential ("sawtooth") in a supercell with $SMFACT = 4$.



system is performed by aligning $\vec{C}$ (and therefore $\vec{E}$) along the z cartesian direction (see keyword **ROTCRY**, page 65). After these transformations the field is along the z direction, and the perturbation $\hat{H}_1(\vec{E})$ takes the form:

$$\hat{H}_1^{\pm}(E_z) = V(z) = -qE_0 \cdot f^{\pm}(z) \tag{3.2}$$

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:

$$f^+(z) = \frac{2C}{\pi^2} \sum_{k=0}^{+\infty} \frac{1}{(2k+1)^2} \cos\left(\frac{2\pi(2k+1)z}{C}\right) \tag{3.3}$$

$$f^-(z) = \frac{2C}{\pi^2} \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k+1)^2} \sin\left(\frac{2\pi(2k+1)z}{C}\right) \tag{3.4}$$

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 313). In this way the perturbed wave function is used for the calculation of ϵ, following a macroscopic average scheme, as described in references [113], [59].

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.

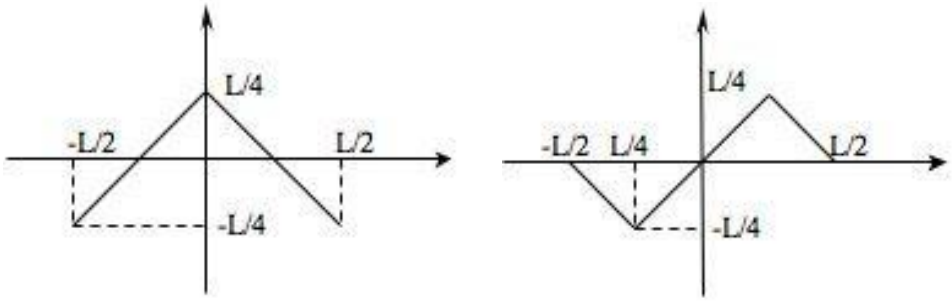
4. In 3D-crystals, the electric potential takes a triangular form to maintain translational symmetry and electric neutrality of cell. The symmetry of triangular potential has two options:

   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.

Figure 3.2: Left: symmetric triangular electric potential (ISIM = 1). Right anti-symmetric triangular electric potential (ISYM=-1).



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 113).

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 119).

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:

$$E(F_0) = E_0 - \mu F_0 - \frac{1}{2!}\alpha F_0^2 - \frac{1}{3!}\beta F_0^3 - \frac{1}{4!}\gamma F_0^4 - \cdots \tag{3.5}$$

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 8, page 207.

FULLE - Building a fullerene from a slab

Fullerenes are molecular cage-like structures. An effective way of constructing them exploiting all the possible symmetry is by starting from a 2D periodic flat structure.
In this flat lattice, the vector $\mathbf{R}=n_1\vec{a_1}+n_2\vec{a_2}$ (where $\vec{a_1}$ and $\vec{a_2}$ are the slab cell vectors) permits to define the side of a triangular face of the fullerene. Then, the indices $(n1, n2)$ completely define the faces of the fullerene, and are used in the literature to characterise this type of systems. The case of a (2,2) carbon fullerene is given in figure 3.3 as an example.
Once the face is defined, the second key information is the type of polyhedron to be constructed, and the corresponding point symmetry. CRYSTAL cuurently permits the automatic construction of the following combinations:

- Icosahedron: I, I_h;
- Octahedron: T, T_d, T_h, O, O_h;
- Tetrahedron: T, T_d.

The **FULLE** keyword can only be used starting from hexagonal 2D lattices.
For further details and explanatory animations, please refer to the tutorial page:
http://www.crystal.unito.it → tutorials → Fullerene systems

Users of the fullerene construction options are kindly requested to cite the following paper[195]:

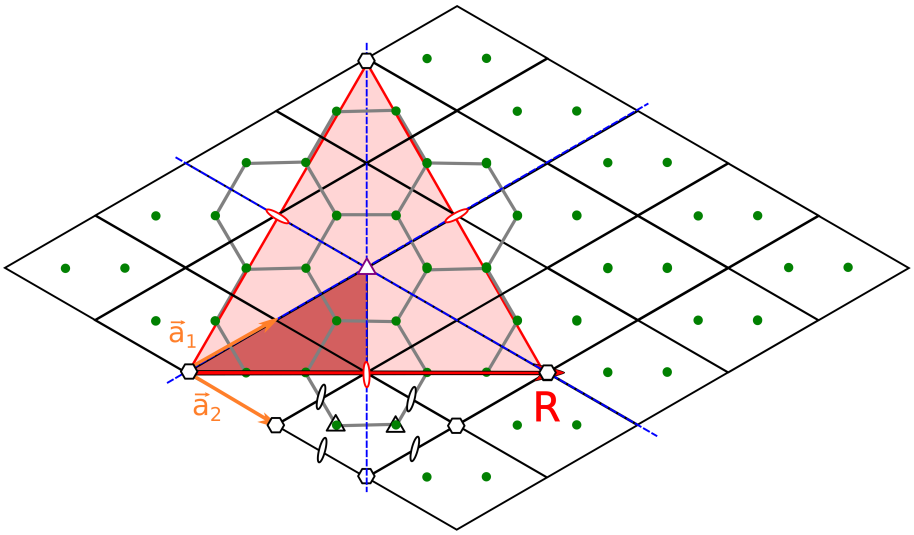


Figure 3.3: Building the (2,2) fullerene from graphene.

Y. Noël, M. De La Pierre, C.M. Zicovich-Wilson, R. Orlando and R. Dovesi, *Phys. Chem. Chem. Phys.*, **16**, 13390 (2014). *Structural, electronic and energetic properties of giant icosahedral fullerenes up to C6000: insights from an ab initio hybrid DFT study.*

rec	variable	value	meaning
•	∗ n_1, n_2		components of the **R** vector in terms of the basis vectors $\vec{a}_1$ and $\vec{a}_2$ of the hexagonal unit cell
•	∗ PG	`T`	point group for the fullerene
		`TD`	
		`TH`	
		`O`	
		`OH`	
		`I`	
		`IH`	
•	∗ POLY	`TETRA`	type of polyhedron
		`OCTA`	
		`ICOSA`	

FULLEJMOL - Graphical visualisation of fullerenes with Jmol

A file named FULLEJMOL.DAT containing the fullerene structure is generated. This file can be directly used by the 3D structure viewer Jmol (www.jmol.org). This keyword must appear in the geometry block ***before*** all the keywords related to fullerene construction.

FULLESPHE - Spherical fullerenes

rec	variable	meaning
•	∗ RSPHE	distance from the origin [Å]

All the atoms of the fullerene are moved to the same radius RSPHE; this is the distance from the origin of the cartesian framework. RSPHE is in Ångstrom, unless otherwise requested (keyword **BOHR**, page 35). The symmetry of the fullerene is preserved. This option permits to construct spherical fullerenes.

This keyword can be used when building fullerenes from a slab (**FULLE** keyword, page 47), when generating a fullerene from molecular point groups with **MOLECULE**, page 15 (additional **RADFULLE** keyword required, page 64), or when recovering a fullerene geometry from an external file with **EXTERNAL**, page 16 (additional **RADFULLE** keyword required). **FULLESPHE** must always be inserted after the **FULLE** or **RADFULLE** keywords.

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 36, 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.

LATVEC - maximum size of the cluster of classified lattice vectors

rec	variable	meaning
• ∗	MAXGSIZE	maximum number of lattice vectors to be classified

A very accurate CRYSTAL calculation may require the evaluation of interactions between atoms in very distant cells (when using very severe computational conditions). If the list of lattice vectors which were classified by default is incomplete with respect to the requirement, an error message appears. In this case it is necessary to extend MAXGSIZE beyond its default value (3500).

MAKESAED

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

MAXNEIGHB - maximum number of equidistant neighbours of an atom

rec	variable	meaning
• ∗	MAXNEIG	maximum number of atoms allowed in creating a cluster

When printing of atom neighbouring relationship is requested (**NEIGHBOR**, page 59) for several orders of vicinity the number of equidistant atoms from a given atom in the unit cell can be large. If such a number exceeds 48 (default value), the program stops with an error message. Use of this keyword allows increasing the value of MAXNEIG.

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

The last version of the program ***MOLDRAW*** reads ***crystal*** standard output, and can generate a movie from an optimization run. See:

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

MOLEBSSE - counterpoise for molecular crystals

rec	variable	meaning
• ∗	NMOL	number of molecules to be isolated
II		insert NMOL records II
• ∗	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 [33] 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 51). The molecule is reconstructed on the basis of the covalent radii reported in Table on page 64. 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.
Use keyword **CLUSTER** (41) to extend the number of atoms in the cluster if requested.

Warning Do not use with ECP

Warning. The system obtained is 0D. No reciprocal lattice information is required in the **scf** input (Section 2.3, page 27). See test 19.

MOLECULE - Extraction of n molecules from a molecular crystal

rec	variable	meaning
• ∗	NMOL	number of molecules to be isolated
II		insert NMOL records II
• ∗	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

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 64). 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 3.1, page 36). 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 2.2, page 20) 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 77)

For molecular crystals only.

Warning. The system is 0D. No reciprocal lattice information is required in the **scf** input (Section 2.3, page 27).
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 17) 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 64)

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 2.3, page 27).
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.

NANOCRYSTAL (NANO)

rec		variable	meaning
•	∗	$h1, k1, l1$	crystallographic (Miller) indices of the first plane parallel to the first nanocrystal face
•	∗	$h2, k2, l2$	crystallographic (Miller) indices of the second plane parallel to the second nanocrystal face
•	∗	$h3, k3, l3$	crystallographic (Miller) indices of the third plane parallel to the third nanocrystal face
•	∗	ISUP1	label of the surface layer of the first nanocrystal face
		NL1	number of atomic layers parallel to the first nanocrystal face
•	∗	ISUP2	label of the surface layer of the second nanocrystal face
		NL2	number of atomic layers parallel to the second nanocrystal face
•	∗	ISUP3	label of the surface layer of the third nanocrystal face
		NL3	number of atomic layers parallel to the third nanocrystal face
•	∗	NCUT	number of further cutting operations

The **NANOCRYSTAL** option is used to create a nanocrystal of given dimension, where three faces are parallel to given planes of the crystal. The other three faces are parallel to the first set.
Before the editing operated by NCUT directives, the nanocrystal is just a supercell of the 3D structure, with faces lying on the three crystallographic planes.
A new 3D crystal unit cell is defined, with three faces parallel to the given planes. The new 3D cell is generally not primitive: the program calculates and prints the multiplicity. Then the atoms of the new 3D cell are classified according their geometric distance from of the three faces parallel to the two given planes.
The thickness of the nanocrystal, the 0D system, is defined by the number of layers parallel to the faces. No reference is made to the chemical units in the nanocrystal. The neutrality is checked by the program.

1. The crystallographic (Miller) indices of the plane refer to the crystallographic cell .

2. A point 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 59).

NANOJMOL - Graphical visualisation of nanotubes with Jmol

Obsolete keyword: Jmol now reads correctly the output of a single and multi-wall nanotube calculation.
A file named NANOJMOL.DAT containing the single or multi-wall nanotube structure is generated. This file can be directly used by the 3D structure viewer Jmol (www.jmol.org). This keyword must appear in the geometry block ***before*** all the keywords related to nanotube construction.

NANORE

Nanotube rebuild: to be used for single-wall tubes built with *NANOTUBE* (page 54). See the *NANOTUBE* keyword for a discussion of nanotube rebuild.

rec		variable	meaning
•	∗	$nold_1, nold_2$	Nanotube rebuild: indices of the starting nanotube.
	∗	n_1, n_2	New indices of the rolling vector.

Consider, for example, the (8,8) and the (10,10) carbon nanotubes (built with *NANOTUBE*, indices refer to the 120° unit cell choice). We have optimised the structure of the former, and we want to build the latter starting from its geometry. With *NANORE* the (8,8) nanotube is unrolled and re-rolled as (10,10). In order to do this, the information on geometry of both the starting slab (graphene) and the (8,8) nanotube is required. The first one is given in input (or read with *EXTERNAL* from file fort.34), the second one is read with an EXTERNAL strategy from file fort.35. The input syntax is then:

```
SLAB
77
2.47
1
6 0.333333  0.666667  0.000000
NANORE
8 8
10 10
```

For nanotubes built with *SWCNT* (page 72) see *SWCNTRE* (page 72).

NANOROD (ROD)

rec		variable	meaning
•	∗	$h1, k1, l1$	crystallographic (Miller) indices of the first plane parallel to the first nanorod face
•	∗	$h2, k2, l2$	crystallographic (Miller) indices of the second plane parallel to the second nanorod face
•	∗	ISUP1	label of the surface layer of the first nanorod face
		NL1	number of atomic layers parallel to the first nanorod face
•	∗	ISUP2	label of the surface layer of the second nanorod face
		NL2	number of atomic layers parallel to the second nanorod face
•	∗	NCUT	number of further cutting operations

The **NANOROD** option is used to create a nanorod of given thickness, where two faces are parallel to given planes of the crystal. The other two faces are parallel to the first pair.
A new 3D crystal unit cell is defined, with two faces parallel to the given planes. The new 3D cell is in general not primitive: the program calculates and prints the multiplicity. Then the atoms of the new 3D cell are classified according to their geometric distance from the two faces parallel to the two given planes.
The thickness of the nanorod, the 1D system, is defined by the number of layers parallel to the two faces. No reference is made to the chemical units in the nanorod. The neutrality is checked by the program.

1. The crystallographic (Miller) indices of the plane refer to the crystallographic cell .
2. A rod 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 59).

NANOTUBE - Building a nanotube from a slab

Nanotubes are cylindrical structures, periodic along one direction. They are therefore characterised by a single lattice vector. However, in order to study their symmetry and orientation, it is easier to start from a 2D lattice where the additional periodicity becomes the cylinder period. This 2D periodic flat structure will be referred in the following, as the nanotube flat lattice.
In this flat lattice, 3 vectors are important for the structure description: the rolling vector, $\vec{R}$, the longitudinal vector, $\vec{L}$ and the helical vector, $\vec{H}$.

- The rolling vector, $\vec{R} = n_1\vec{a_1}+n_2\vec{a_2}$ (where $\vec{a_1}$ and $\vec{a_2}$ are the slab cell vectors), is sufficient to completely define a nanotube. It is used in the literature to characterise the systems (for example the (4,3) nanotube is a nanotube with a rolling vector $\vec{R} = 4\vec{a_1}+3\vec{a_2}$). In the flat nanotube lattice, $\vec{R}$ is a nanotube unit cell vector. Once the nanotube wrapped, the rolling vector becomes a circle normal to the cylinder axis; its norm corresponds to the cylinder perimeter. If $\vec{R}$ passes $\mathcal{N}$ times through the lattice nodes, a $\mathcal{N}$-order rotational axis will exist along the corresponding nanotube axis.
- The longitudinal vector, $\vec{L} = l_1\vec{a_1} + l_2\vec{a_2}$, is the shortest lattice vector normal to $\vec{R}$. In the nanotube, it becomes the 1D lattice parameter and gives the 1D periodicity along the tube axis.
- The helical vector, $\vec{H} = h_1\vec{a_1}+h_2\vec{a_2}$, is a lattice vector defining with $\vec{R}$ an area which is $\mathcal{N}$ times the area of the unit cell of the flat slab. It satisfies, then, the following relationship

$$\frac{S(\vec{R},\vec{H})}{S(\vec{a_1},\vec{a_2})} = |n_1h_2 - n_2h_1| = \mathcal{N} \tag{3.6}$$

  where $S(\vec{v_i},\vec{v_j})$ is the surface defined by the $\vec{v_i}$ and $\vec{v_j}$ vectors. The helical vector defines a correspondence between a translation in the flat slab and a roto-translation in the curved surface; $\vec{H}$ has a rotational component along the circumference vector and a translational component along the lattice parameter.

The direct product between the rotational and roto-translational operations generates the full symmetry of the nanotube.
The three vectors listed previously are represented in figure 3.4 for graphene; the example refers to the construction of a (4,2) single-walled carbon nanotube (SWCNT).
Further information about the implemented method, the computational costs and the symmetry features of nanotubes are found in Ref. [58] and [290, 12].
For further details and explanatory animations, please refer to the tutorial page:
http://www.crystal.unito.it → tutorials → Nanotube systems

Users of the nanotube construction options are kindly requested to cite the following paper[194]:

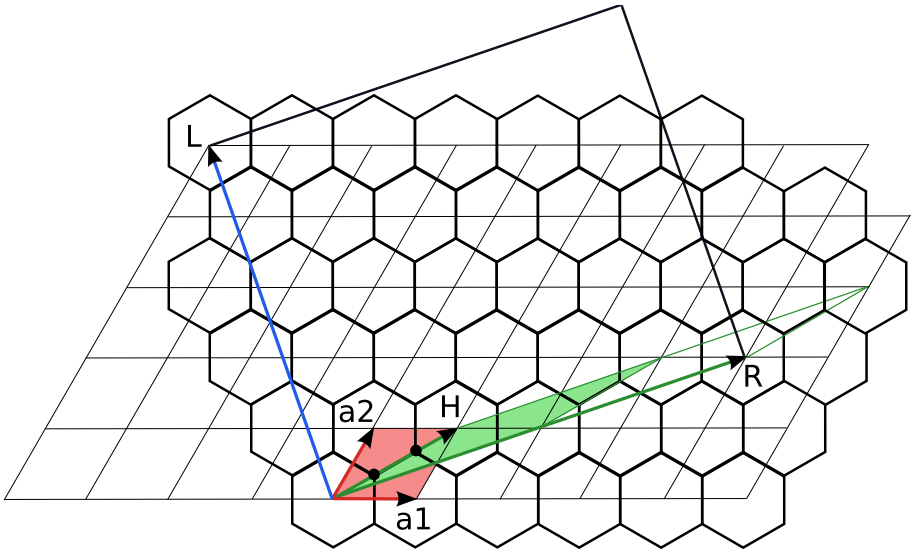


Figure 3.4: Building the (4,2) SWCNT from graphene.

Y. Noël, Ph. D'Arco, R. Demichelis, C.M. Zicovich-Wilson and R. Dovesi, *J. Comput. Chem.*, **31**, 855 (2010). *On the Use of Symmetry in the Ab Initio Quantum Mechanical Simulation of Nanotubes and Related Materials.*

rec	variable	meaning
•	∗ n_1, n_2	components of the roll-up-vector of the nanotube in the slab unit cell basis. The roll-up-vector, whose length corresponds to the tube circumference, is expressed as a linear combination of the unit cell vectors of the slab before rolling up, n_1 and n_2 being integer coefficients.

The convention adopted in CRYSTAL is $n_1 \geq n_2$. In cases when n_2 is required to be larger than n_1, it is sufficient to exchange the x and y coordinates of the reference 2D slab.
NANOTUBE can be used with the following 2D systems:

- square and hexagonal lattices: any (n_1, n_2) combination
- rectangular lattices: any n_1, $n_2 = 0$
- rhombohedral lattices: $n_2 = \pm n_1$

In the other cases, the existence of a lattice vector perpendicular to $\vec{R}$ (so the 1D periodicity along the tube axis) is not guaranteed.
The orthogonality condition between $\vec{R}$ and $\vec{L}$ provides the following equation:

$$\vec{L} \cdot \vec{R} = (l_1\vec{a_1} + l_2\vec{a_2}) \cdot (n_1\vec{a_1} + n_2\vec{a_2}) = n_1 l_1 |\vec{a_1}|^2 + n_2 l_2 |\vec{a_2}|^2 + (n_1 l_2 + n_2 l_1)|\vec{a_1}||\vec{a_2}| cos(\gamma) = 0 \quad (3.7)$$

γ being the angle between $\vec{a_1}$ and $\vec{a_2}$. From Equation 3.7 one obtains:

$$\frac{l_1}{l_2} = -\frac{n_2|a_2|^2 + n_1|a_1||a_2|cos(\gamma)}{n_1|a_1|^2 + n_2|a_1||a_2|cos(\gamma)} \quad (3.8)$$

If the right term is divided and multiplied by $|a_2|^2$, one obtains

$$\frac{l_1}{l_2} = -\frac{n_2 + n_1 a_q cos(\gamma)}{a_q(n_1 a_q + n_2 cos(\gamma))} \quad (3.9)$$

with $a_q = \frac{|a_1|}{|a_2|}$.

The above equation cannot be satisfied for any ($\vec{a_1}$,$\vec{a_2}$,$cos(\gamma)$) combination. This observation is based on the fact that, as l_1 and l_2 are integers, $\frac{l_1}{l_2}$ is a rational number, whereas, in general, $cos(\gamma)$ and a_q are real numbers. In the following the five 2D Bravais lattices are considered separately, in order to show which conditions satisfy the periodicity along the tube axis and which do not.

- Hexagonal lattice: $\vec{a_1} = \vec{a_2}$, $cos(\gamma) = \pm 1/2$. Equation 3.9 becomes:

$$\frac{l_1}{l_2} = -\frac{n_2 + n_1/2}{n_1 + n_2/2} = -\frac{2n_2 + n_1}{2n_1 + n_2} \tag{3.10}$$

  Any roll-up vector is possible.

- Square lattice: $\vec{a_1} = \vec{a_2}$, $cos(\gamma) = 0$.

$$\frac{l_1}{l_2} = -\frac{n_2}{n_1} \tag{3.11}$$

  Any roll-up vector is possible.

- Rectangular lattice: $\vec{a_1} \neq \vec{a_2}$, $cos(\gamma) = 0$.

$$\frac{l_1}{l_2} = -\frac{n_2|a_2|^2}{n_1|a_1|^2} \tag{3.12}$$

  In this case the right term corresponds to a rational number either if $|a_1| = n|a_2|$, with n being a rational number, or if n_2=0. More generally, for rectangular lattices, the periodicity along the tube axis is always satisfied for $(n, 0)$ and $(0, n)$ roll-up vectors.

- Rhombohedral (centred rectangular) lattice: $\vec{a_1} = \vec{a_2}$, any $cos(\gamma)$.

$$\frac{l_1}{l_2} = -\frac{n_2 + n_1 cos(\gamma)}{n_1 + n_2 cos(\gamma)} \tag{3.13}$$

  The right term provides a rational number only when $n_1 = n_2$ or $n_1 = -n_2$, so that Equation 3.13 becomes:

$$\frac{l_1}{l_2} = -\frac{n_1[1 + cos(\gamma)]}{n_1[1 + cos(\gamma)]} = -1 \qquad \frac{l_1}{l_2} = \frac{n_1[1 - cos(\gamma)]}{n_1[1 - cos(\gamma)]} = 1 \tag{3.14}$$

- Oblique lattice: $\vec{a_1} \neq \vec{a_2}$, any $cos(\gamma)$. Equation 3.9 remains as such, and the right term is always an irrational number.

When the above conditions are not satisfied, however, it is possible to manipulate the geometry of the starting slab and force it to assume a suitable form, by building supercells or with minor modifications to the cell parameters.

Note that, in cases of hexagonal lattices, the standard convention adopted in CRYSTAL considers the reference flat lattice cell vectors forming a 120° angle. In the carbon nanotube literature, however, the rolling indices refer to a 60° angle. For this reason, the same input can be obtained with the following choices, where the (6,3) indices in the 60° unit cell become (9,3) in the 120° option:

```
Example I: (6,3) SWCNT from graphene, 60 degrees
SLAB
1
2.47  2.47 60.000
2
6   0.33333333333  0.33333333333 0.000
6  -0.33333333333 -0.33333333333 0.000
NANOTUBE
6   3
```

```
Example II: (6,3) SWCNT from graphene, default choice
SLAB
77
2.47
1
6   0.333333   0.666667   0.000000
NANOTUBE
9   3
```

In order to avoid ambiguity in the choice of the carbon nanotubes indices, use the *SWCNT* (page 72) keyword.

Nanotube rebuild: build a nanotube starting from the structure of another one, with same directions but a different radius.

Two restart keywords, *NANORE* (page 53) and *SWCNTRE*(page 72), allow to build a (n_1, n_2) nanotube by starting from the structure of another one (a previously $(nold_1, nold_2)$ optimised one, read from file fort.35). The "old" nanotube is unrolled and re-rolled according to a "new" $\vec{R}$ vector, with minor modifications to the structure. The rolling direction of the two tubes must be the same.

It is particularly helpful for the geometry optimisation of inorganic nanotubes (thick slabs, large systems, the geometry of the tubes is very different from the one of the slab), as the number of optimisation steps is reduced (see Ref. [61] for documentation of computational gain).

NANOMULTI - Building a multi-wall nanotube from a layer slab

From a 2D periodic flat structure, previously referred as the nanotube flat lattice, multi-wall nanotubes can be wrapped, with $M \geq 2$, adopting the same procedure and conventions outlined for the single-wall materials.3.1
In particular, it is possibile to specify for each wall a couple of indices (n_1, n_2) which univocally determines the rolling vector, $\vec{R}$, the longitudinal vector, $\vec{L}$ and the helical vector, $\vec{H}$, of the tube, that is all its structural features.
The construction of each wall is introduced by the keyword *WALL* and the *OPTWALL* option allows its full geometry optimization, as if it were an isolated nanotube. The final energy of this calculation can be used as a reference for the isolated system and as a first estimate of the inter-wall energy. The density matrices at convergence of the isolated walls, each saved on the file DENSMATM.BIN, can be used as the initial guess for constructing the density matrix of the whole structure. This procedure is activated by the keyword *GUESSMULTI* and can at time facilitate convergence. Then, by means of the *OPTMULTI* keyword the full optimization of the M-wall structure is performed.
Some structural manipulations can be accomplished by using two additional keywords within the *WALL* block: *ROTWALL* and *TRANSWALL*. With the first, the nanotube is rotated around the periodic x-axis of a given angle, between 0° and 360°. It is worth to mention that rotations of a few degrees often already reproduce the same initial configuration. *TRANSWALL* performs a rigid displacement along the x-axis of the atomic position of the selected wall. Both options can be used together and in association with optimization keywords to study the effect of different stacking of tubes on various properties.

rec		variable	meaning
•	∗	NANOMULTI	
•	∗	M	number of walls to wrap starting from the same layer slab
			insert M records
•	∗	WALL	
•	∗	m	number of the wall, $1 < m < M$
•	∗	ROLLINGV	
•	∗	n_1, n_2	components of the roll-up-vector of the nanotube in the slab unit cell basis: $\vec{R} = n_1\vec{a_1} + n_2\vec{a_2}$, being a_1 and a_2 the cell vectors of the original 2D slab.
		I	optional keywords for each *WALL* I
		• OPTWALL	a complete (atomic position, lattice parameter and volume) optimization of the isolated single-wall is performed
		• GUESSP	the density matrix of the single-wall is read on the file DENSMATM.DAT and used as the initial guess
		• TRANSWALL	
		• t	rigid translation of t [Å] of all the atoms of the unit cell of the wall along the periodic x-axis (default: $t = 0.0$)
		• ROTWALL	
		• γ	rotation of γ degrees around the periodic x-axis of all the atoms of the unit cell of the wall (default: $\gamma = 0°$)
		• NOSCFWALL	the wall is build but its energy (and density matrix) is not calculated
		I	end optional keywords for each *WALL* I
			end of M records
		II	global optional keyword II
•	∗	GUESSMULTI	the density matrices at convergence of each single-wall are packed together and used as initial guess for the density matrix of the multi-wall system
•	∗	OPTMULTI	a complete (atomic position, lattice parameter and volume) optimization of the whole multi-wall structure is performed
		II	end global optional keyword II
•	∗	ENDWALL	end of the multi-wall nanotube input cards

```
Input Example: 2-Wall zig-zag Carbon Nanotube
SLAB
77
2.47
1
6   0.333333   0.666667   0.000000
NANOMULTI
2                  <-- double-wall nanotube
WALL               <-- first tube
1
ROLLINGV
6   0
ROTWALL
0.75
OPTWALL
WALL               <-- second tube
2
ROLLINGV
12  0
```

```
TRANSWALL
0.50
OPTWALL           <-- end of the tubes
GUESSMULTI
OPTMULTI          <-- full optimization
ENDWALL
```

For further details, please refer to the tutorial page:
http://www.crystal.unito.it → tutorials → *How to Model* → Multi-Wall Nanotube

Users of the multi-wall nanotube options are kindly requested to cite the following paper[178]:

N.L. Marana, Y. Noël, J.R. Sambrano, C. Ribaldone and S. Casassa, *J. Phys. Chem. A*, **18**, 4003 (2021). *Ab Initio Modeling of MultiWall: A General Algorithm First Applied to Carbon Nanotubes.*

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 7, page 182.

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 66). No input data are required.

OPTBASIS - Basis Set Optimization

rec	variable	meaning
• A	**OPTBASIS**	basis Set Optimization
• A	**ALLBDIIS**	Exponents and coefficients optimization
• A	**EXPONLY**	Optimization of exponents only [default]
• A	**COEFFONLY**	Optimization of coefficients only
• A	**TOLOPTE**	Optimization convergence threshold on Energy
• *	itol_e	Integer defining the threshold[default: $5 \rightarrow 10^{-5}$]
• A	**TOLOPTG**	Convergence threshold on basis set gradient
• *	tol_g	Real [default: $0.002 \rightarrow 0.002 * 1.5 = 3 * 10^{-3}$]
• A	**MAXBASCYC**	Maximum number of basis set optimization cycles
• *	max_bas_cyc	Integer [default: 500]
• A	**REFBAS**	Frequency of integral re-screening
• *	n_ref_bas	Integer defining after how many iterations intgral screening is activated [default: 10]
• A	**ENDBO**	basis Set Optimization END statement

The basis set optimization exploits the BDIIS algorithm[56] in order to minimize the functional Ω:

$$\Omega\left(\{\alpha, d\}\right) = E_{tot}\left(\{\alpha, d\}\right) + \gamma \ln \kappa\left(\{\alpha, d\}\right) \tag{3.15}$$

Ω contains the total energy of the system and a penalty function, namely the condition number (ratio between the highest and lowest eigenvalues of the overlap matrix). The introduction of such penalty function should prevent the onset of harmful linear dependence that can give rise to numerical instabilities and unphysical states.
The optimization of the entire basis set is a formidable task and quite likely to be not very useful. Therefore the users can select from the basis set (the basis set must be written explicitly in the input) which functions they want to optimize
The shells involved in the optimization must be highlighted by a star * in the input basis set. e.g.:

```
...
0 0 1 0 1 *
      1.2429415962           1.0000000
...
```

Suggestions: since the valence shells are usually the ones involved in the chemical bonding, these are the ones usually to be optimized. For less expert users, it is recommended to optimize only exponents and if possible it is also suggested to uncontract the shells if not already. In order to follow the optimization steps, besides checking the minimization of the *Objective Function*, it can be useful to check the stability of the SCF calculations performed in the meanwhile. This information is available in the `SCFOUT.LOG` file where a storing of all the SCF calculations performed are reported.
As an example we report the input for He bulk (molecular def2-QZVP basis set).

```
He-crystal
CRYSTAL
0 0 0
194
2.92661 4.77913
1
2    0.333333333333333   -0.333333333333333   0.250000000000000
OPTBASIS
ENDBO
END
```

```
  2 10
0 0 5 2 1
   1144.6470809              0.35861578618E-03
    171.64596667             0.27725434466E-02
     39.066056254            0.14241892216E-01
     11.051401989            0.55457352277E-01
      3.5725574473           0.16170511810
0 0 1 0 1 *
      1.2429415962           1.0000000
0 0 1 0 1 *
      0.44807668730          1.0000000
0 0 1 0 1 *
      0.16411579128          1.0000000
0 2 1 0 1 *
      5.99400000             1.0000000
0 2 1 0 1 *
      1.74500000             1.0000000
0 2 1 0 1 *
      0.56000000             1.0000000
0 3 1 0 1 *
      4.29900000             1.0000000
0 3 1 0 1 *
      1.22300000             1.0000000
0 4 1 0 1 *
      2.68000000             1.0000000
99 0
END
DFT
PBE
END
TOLINTEG
8 8 8 8 16
SHRINK
8 8
TOLDEE
8
END
```

In this specific case all the uncontracted shells and only the exponents are involved in the optimization. As regard the thresholds, here the default options are kept.
Apart from the usual output, two other files are generated:

- *filename.out2*: this is usually called SCFOUT.LOG. This is a file where all the background SCF calculations are printed out - like in geometry optimization and similar procedures..
- *filename.f53*: a fort.53 unit where the optimized basis sets at each optimization cycle are printed out. To restart an OPTBASIS calculation, the user can simply copy the last basis set printed in this output and paste it into a new input.

PBAND - Phonon Bands -

This keyword, to be inserted in the geometry input block, is specific for phonon dispersion calculations (the use of the **DISPERSI** keyword is then mandatory in the subsequent **FREQCALC** block, similarly to the **SCELPHONO** keyword, see page 68). It can be used for phonon band calculations, and allows for the specification of the path in reciprocal space along which the Hessian matrix is obtained.

rec	variable	meaning
• ∗	ISS	shrinking factor in terms of which the coordinates of the extremes of the segment are expressed. If ISS=0, input by label (see below).
	NK	total number of **k**-points along the line.
	FLAG1	if > 1, activates interpolation (see keyword **INTERPHESS** at page 227)
	FLAG2	coordinates given in terms of primitive (FLAG2=0) or conventional (FLAG2=1) lattice
		if ISS > 0 then
• ∗	I1,I2,I3	integers that define the starting point of the line (I1/ISS $\mathbf{b}_1$+I2/ISS $\mathbf{b}_2$+I3/ISS $\mathbf{b}_3$), with $\mathbf{b}_1$,$\mathbf{b}_2$,$\mathbf{b}_3$ reciprocal lattice vectors.
	J1,J2,J3	integers that define the final point of the line (J1/ISS $\mathbf{b}_1$+J2/ISS $\mathbf{b}_2$+J3/ISS $\mathbf{b}_3$) with $\mathbf{b}_1$,$\mathbf{b}_2$,$\mathbf{b}_3$ reciprocal lattice vectors.
		if ISS $= 0$ then
• ∗	LABELA	label of the the starting point of the line (see tables 14.1 and 14.2 for a legend).
	LABELB	label of the the final point of the line (see tables 14.1 and 14.2 for a legend).

Given two points that define a line in reciprocal space, and the number of sampling points required along this line, a supercell is constructed to allow for the exact calculation of the Hessian matrix at these points. The formal relationship between the direct space supercell and the reciprocal space sampling can be found in the phonon dispersion section (page 227). Multiple runs on supercells constructed in this way allow for the complete calculation and plotting of phonon band structures.

FLAG1 is normally set to 1. If FLAG1$>$ 1 then NK$\times$FLAG1 **k**-points are obtained without building a bigger supercell. The assumptions are the same as those used for the keyword **INTERPHESS** (see page 227). FLAG2 is 1 if the input coordinates refer to the conventional cell, 0 if they refer to the primitive cell.

The computed frequencies at all the requested points are written in output and stored in units fort.25 and PBANDS.DAT (in x-y format) for plotting, in the same formats used for electronic band structures.

An input example is:

```
MGO BULK
CRYSTAL
0 0 0
225
4.16563249
2
12  0.   0.   0.
8   0.5  0.5  0.5
PBAND
0 10 1 0
G
X
FREQCALC
DISPERSI
END
END
```

which, in this particular case, will result in the supercell:

```
         EXPANSION MATRIX OF PRIMITIVE CELL
 E1     1.000     0.000    -1.000
 E2     0.000     1.000     0.000
 E3     0.000     0.000    18.000
```

PLANES

rec	variable	meaning
• ∗	f	minimum of the interplane distance d_{hkl} for the plane families to be printed and indexed, expressed as the fraction f

Crystallographic planes are indexed and printed, grouped in families of symmetry-related planes. All the *hkl* planes are analyzed that have the interplane distance belonging to the interval $[f\, d_{hkl}^{max}, d_{hkl}^{max}]$, where d_{hkl}^{max} is the maximum interplane distance for a given family *hkl* and f a number between 0 and 1. If the crystal structure has a primitive cell different from the crystallographic one, the plane indices are printed relatively to both frames.

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 35). 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**, **ELASTCON**, **EOS**, **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 421) 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 437.
Extended printing request can be entered in any input block. Printing requests are not transferred from wave function to properties calculation.
See Appendix C, page 435.

PRSYMDIR

Printing of the so-called *symmetry allowed directions* adopted in the geometry optimization. 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.

RADFULLE

When printing atomic coordinates of the molecule, an additional column is added that contains the radial distances of the atoms from the origin. The **FULLE** (page 47) keyword constructs fullerenes and automatically displays the radius column. The present option is useful when running a fullerene calculation without using the **FULLE** keyword, i.e. by means of the **MOLECULE** (page 15) and **EXTERNAL** (page 16) keywords.

RADNANO

When printing atomic coordinates of a 1-D system, an additional column is added that contains the radial distances of the atoms from the x axis. The **NANOTUBE** (page 54) and **SWCNT** (page 72) keywords construct nanotubes and automatically display the radius column. The present option is useful when running a nanotube calculation without using the **NANOTUBE/SWCNT** keywords, i.e. by means of the **MOLECULE** (page 15) and **EXTERNAL** (page 16) keywords.

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 ≤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	1.97
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	2.07	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
Fr	Ra	Ac															
2.61	2.21	2.16															

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
2.07	2.04	2.03	2.01	1.99	1.98	1.98	1.96	1.94	1.92	1.92	1.89	1.90	1.87	1.87
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr
2.16	2.06	2.00	1.97	1.91	1.87	1.80	1.68	1.55	1.50	1.50				

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.
Data for Rn-Es have been inserted in CRYSTAL22, based on Ref. [54].

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 page 37) 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 C, page 437. 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 59).

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 67) option. This information can be obtained by a test run, inserting in the geometry input block the keyword **TESTGEOM** (page 73). 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 36), **ATOMDISP** (page 36), 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 66).

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 D, page 445.

SCELCONF

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

This keyword is specific for configuration counting calculations for disordered systems or solid solutions. Always use **SCELCONF** instead of **SUPERCELL/SUPERCON** in calculations with **CONFCOUNT** and **CONFRAND**.

A supercell is constructed, as in the case of keyword **SUPERCELL** (see page 70), as linear combination of the primitive cell unit vectors.

The number of symmetry operators of the resulting supercell is higher compared to the **SUPERCELL** case. The product group is constructed between the space group of the original system and the group of translation operators associated with the chosen supercell.

SCELPHONO

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

This keyword is specific for phonon dispersion calculations.

A supercell is constructed, as in the case of keyword **SUPERCELL** (see page 70), as linear combination of the primitive cell unit vectors.

The atomic ordering of the resulting supercell is different with respect to the **SUPERCELL** case. The first atoms in the list are the ones of the primitive cell, as required for phonon dispersion calculations. Example:

```
TEST11 - MGO BULK
CRYSTAL
0 0 0
 225
```

```
4.21
2
12 0.    0.    0.
8 0.5   0.5   0.5
SCELPHONO
2 0 0
0 2 0
0 0 2
TESTGEOM
END
=======================  output - atoms in the small cell first ================
 PRIMITIVE CELL - CENTRING CODE 1/0 VOLUME=   149.236922 - DENSITY  3.559 g/cm^3
         A              B              C           ALPHA      BETA       GAMMA
     5.95383910     5.95383910     5.95383910    60.000000  60.000000  60.000000
 ****************************************************************************
 ATOMS IN THE ASYMMETRIC UNIT    5 - ATOMS IN THE UNIT CELL:   16
     ATOM              X/A                Y/B                Z/C
 ****************************************************************************
   1 T  12 MG    0.000000000000E+00  0.000000000000E+00  0.000000000000E+00
   2 T   8 O     2.500000000000E-01  2.500000000000E-01  2.500000000000E-01
   3 T  12 MG    0.000000000000E+00 -5.000000000000E-01  0.000000000000E+00
   4 F  12 MG    0.000000000000E+00 -5.000000000000E-01 -5.000000000000E-01
   5 F  12 MG   -5.000000000000E-01  0.000000000000E+00  0.000000000000E+00
   6 F  12 MG   -5.000000000000E-01  0.000000000000E+00 -5.000000000000E-01
   7 F  12 MG   -5.000000000000E-01 -5.000000000000E-01  0.000000000000E+00
   8 T  12 MG   -5.000000000000E-01 -5.000000000000E-01 -5.000000000000E-01
   9 F  12 MG    0.000000000000E+00  0.000000000000E+00 -5.000000000000E-01
  10 T   8 O     2.500000000000E-01  2.500000000000E-01 -2.500000000000E-01
  11 F   8 O     2.500000000000E-01 -2.500000000000E-01  2.500000000000E-01
  12 F   8 O     2.500000000000E-01 -2.500000000000E-01 -2.500000000000E-01
  13 F   8 O    -2.500000000000E-01  2.500000000000E-01  2.500000000000E-01
  14 F   8 O    -2.500000000000E-01  2.500000000000E-01 -2.500000000000E-01
  15 F   8 O    -2.500000000000E-01 -2.500000000000E-01  2.500000000000E-01
  16 F   8 O    -2.500000000000E-01 -2.500000000000E-01 -2.500000000000E-01


TEST11 - MGO BULK
CRYSTAL
0 0 0
 225
4.21
2
12 0.    0.    0.
8 0.5   0.5   0.5
SUPERCEL
2 0 0
0 2 0
0 0 2
TESTGEOM

=======================  output - atoms of same type first ================

 PRIMITIVE CELL - CENTRING CODE 1/0 VOLUME=   149.236922 - DENSITY  3.559 g/cm^3
         A              B              C           ALPHA      BETA       GAMMA
     5.95383910     5.95383910     5.95383910    60.000000  60.000000  60.000000
 ****************************************************************************
 ATOMS IN THE ASYMMETRIC UNIT    5 - ATOMS IN THE UNIT CELL:   16
     ATOM              X/A                Y/B                Z/C
 ****************************************************************************
   1 T  12 MG    0.000000000000E+00  0.000000000000E+00  0.000000000000E+00
   2 T  12 MG    0.000000000000E+00  0.000000000000E+00 -5.000000000000E-01
   3 F  12 MG    0.000000000000E+00 -5.000000000000E-01  0.000000000000E+00
```

```
  4 F  12 MG    0.000000000000E+00 -5.000000000000E-01 -5.000000000000E-01
  5 F  12 MG   -5.000000000000E-01  0.000000000000E+00  0.000000000000E+00
  6 F  12 MG   -5.000000000000E-01  0.000000000000E+00 -5.000000000000E-01
  7 F  12 MG   -5.000000000000E-01 -5.000000000000E-01  0.000000000000E+00
  8 T  12 MG   -5.000000000000E-01 -5.000000000000E-01 -5.000000000000E-01
  9 T   8 O    -2.500000000000E-01 -2.500000000000E-01  2.500000000000E-01
 10 T   8 O    -2.500000000000E-01 -2.500000000000E-01 -2.500000000000E-01
 11 F   8 O    -2.500000000000E-01  2.500000000000E-01  2.500000000000E-01
 12 F   8 O    -2.500000000000E-01  2.500000000000E-01 -2.500000000000E-01
 13 F   8 O     2.500000000000E-01 -2.500000000000E-01  2.500000000000E-01
 14 F   8 O     2.500000000000E-01 -2.500000000000E-01 -2.500000000000E-01
 15 F   8 O     2.500000000000E-01  2.500000000000E-01  2.500000000000E-01
 16 F   8 O     2.500000000000E-01  2.500000000000E-01 -2.500000000000E-01
```

It can be used to generate a supercell for a phonon dispersion calculation only (keyword **DISPERSI**, page 227).

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 (See **SUPERCON** for crystallographic 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:

$$\begin{aligned}
\mathbf{b}'_\mathbf{1} &= e_{11} \cdot \mathbf{b}_1 + e_{12} \cdot \mathbf{b}_2 + e_{13} \cdot \mathbf{b}_3 \\
\mathbf{b}'_\mathbf{2} &= e_{21} \cdot \mathbf{b}_1 + e_{22} \cdot \mathbf{b}_2 + e_{23} \cdot \mathbf{b}_3 \\
\mathbf{b}'_\mathbf{3} &= e_{31} \cdot \mathbf{b}_1 + e_{32} \cdot \mathbf{b}_2 + e_{33} \cdot \mathbf{b}_3
\end{aligned}$$

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 18). To avoid this operation, insert **NOSHIFT** before **SUPERCEL**
Atoms that are related by translational symmetry in the unit cell are considered nonequivalent 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 37), **ATOMSUBS** (page 38), **ATOMINSE** (page 36), **ATOMDISP** (page 36), **SLAB** (page 66).
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 94). 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 18).

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

The first step in geometry definition in CRYSTAL is the transformation of the cell from crystallographic to primitive, according to the transformation matrices given in appendix A.5, when the crystallographic cell is non-primitive.
From the point of view of CRYSTAL, the crystallographic cell is a supercell, in that case.
Using the keyword **SUPERCON**, the crystallographic cell is built from the primitive one, before applying the expansion matrix.
See **SUPERCEL**, page 70 for expansion matrix input instructions.

Note - The easiest way to work on crystallographic cell is the following input:

```
SUPERCON
1. 0. 0.
0. 1. 0.
0. 0. 1.
```

The expansion matrix is the identity, leaving the crystallographic cell unmodified.

SWCNT - Building a single-wall nanotube from an hexagonal slab (60° unit cell choice)

rec		variable	meaning
•	∗	n_1, n_2	Only for hexagonal cells. Same as *NANOTUBE* keyword but the components of the rolling vector refer to a 60° hexagonal cell.

vspace0.3cm

```
Example: (6,3) SWCNT from graphene
SLAB
77
2.47
1
6   0.333333   0.666667   0.000000
SWCNT
6  3
```

See *NANOTUBE* (page 54) for further details.

Nanotube rebuild: build a nanotube starting from the structure of another one, with same directions but a different radius.

Two restart keywords, *NANORE* and *SWCNTRE*, allow to build a (n_1, n_2) nanotube by starting from the structure of another one (a previously $(nold_1, nold_2)$ optimised one, read from file fort.35). The "old" nanotube is unrolled and re-rolled according to a "new" $\vec{R}$ vector, with minor modifications to the structure. The rolling direction of the two tubes must be the same.

It is particularly helpful for the geometry optimisation of inorganic nanotubes (thick slabs, large systems, the geometry of the tubes is very different from the one of the slab), as the number of optimisation steps is reduced (see Ref. [61] for documentation of computational gain).

SWCNTRE

It is the same as *NANORE* (page 53), with the same input syntax, but for hexagonal lattices with 60° unit cell reference (see *SWCNT*, page 72).

rec		variable	meaning
•	∗	$nold_1, nold_2$	Nanotube rebuild: indices of the starting nanotube.
	∗	n_1, n_2	New indices of the rolling vector.

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 59), 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 49) δ for the allowed distortion are given in input.

WULFF

rec	variable	meaning
• ∗	IPLN	number of crystal planes
		insert IPLN records II
• ∗	h, k, l	Miller indices of the plane
• ∗	ESRF	surface energy in J/m^2

The **WULFF** option is used to determine the equilibrium shape of real (finite) crystals by ensueing the minimization of the total surface free energy according to the Gibbs-Wulff theorem [294].

The execution stops after the Wulff construction is completed. It generates a three dimensional model, wulff.pms, that may be displayed with the free application Mesh Viewer (http://mview.sourceforge.net).

3.2 Basis set input

Symmetry control			
ATOMSYMM	printing of point symmetry at the atomic positions	38	–
Basis set modification			
CHEMOD	modification of the electronic configuration	75	I
GHOSTS	eliminates nuclei and electrons, leaving BS	77	I
Auxiliary and control keywords			
CHARGED	allows non-neutral cell	75	–
NOPRINT	printing of basis set removed	77	–
PRINTOUT	setting of printing options	64	I
SETINF	setting of inf array options	66	I
SETPRINT	setting of printing options	66	I
STOP	execution stops immediately	68	–
SYMMOPS	printing of point symmetry operators	73	–
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	77	–

Basis set input optional keywords

ATOMSYMM

See input block 1, page 38

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 50).

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 (Gaussian 98 or Gaussian03) [111, 110]. The deck can be prepared without the calculation of the wave function by entering the keyword **TESTPDIM** in input block 3 (page 124). 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 **GAUSS98** 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 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 bipolar 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., the symbol is "XX".
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 37) 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.

PRINTOUT - Setting of printing environment

See input block 1, page 64.

SETINF - Setting of INF values

See input block 1, page 66.

SETPRINT - Setting of printing options

See input block 1, page 66.

STOP

Execution stops immediately. Subsequent input records are not processed.

TEST[BS]

Basis set input is checked, then execution stops. Subsequent input records (block3) are not processed.

SYMMOPS

See input block 1, page 73

Effective core pseudo-potentials - ECP

rec	variable	value	meaning
• A	PSN		pseudo-potential keyword:
		HAYWLC	Hay and Wadt large core ECP. (SR)
		HAYWSC	Hay and Wadt small core ECP. (SR)
		BARTHE	Durand and Barthelat ECP. (SR)
		DURAND	Durand and Barthelat ECP. (SR)
		COLUSC	Columbus small-core ECP. (FR) - CRYSTAL22
		COLULC	Columbus large-core ECP. (FR) - CRYSTAL22
		COLUSH	Columbus ECP for super-heavy elements (atomic number > 99, FR) - CRYSTAL22
		STUTSC	Stuttgart-Cologne small-core ECP. (FR) - CRYSTAL22
		STUTLC	Stuttgart-Cologne large-core ECP. (FR) - CRYSTAL22
		STUTSH	Stuttgart-Cologne ECP for super-heavy elements (atomic number > 99, FR) - CRYSTAL22
		INPUT	free ECP (SR) - input follows.
		INPSOC	free ECP (FR) - CRYSTAL22.
		if PSN = INPUT insert	II
• ∗	ZNUC		effective core charge (ZN in eq. 3.17).
	M		Number of terms in eq. 3.18
	M0		Number of terms in eq. 3.19 for ℓ =0.
	M1		Number of terms in eq. 3.19 for ℓ =1.
	M2		Number of terms in eq. 3.19 for ℓ =2.
	M3		Number of terms in eq. 3.19 for ℓ =3.
	M4		Number of terms in eq. 3.19 for ℓ =4. - CRYSTAL09
		insert M+M0+M1+M2+M3+M4 records	II
• ∗	ALFKL		Exponents of the Gaussians: $\alpha_{k\ell}$.
	CGKL		Coefficient of the Gaussians: $C_{k\ell}$.
	NKL		Exponent of the r factors: $n_{k\ell}$.
		if PSN = INPSOC insert - CRYSTAL22	II
• *A*	PSONAME		Convention for SOREP coefficients. - see eqs. 3.26b-3.28
		INTERNAL	
		STUTTGART	
		TOULOUSE	
		COLUMBUS	
• ∗	SOSCALE		Global scaling factor X for the SOREP in eq. 3.23
• ∗	ZNUC		effective core charge (ZN in eq. 3.17).
	M		Number of terms in eq. 3.18
	M0		Number of terms in eq. 3.19 for ℓ =0.
	M1		Number of terms in eq. 3.19 for ℓ =1.
	M2		Number of terms in eq. 3.19 for ℓ =2.
	M3		Number of terms in eq. 3.19 for ℓ =3.
	M4		Number of terms in eq. 3.19 for ℓ =4. - CRYSTAL09
		insert M+M0+M1+M2+M3+M4 records	II
• ∗	ALFKL		Exponents of the Gaussians: $\alpha_{k\ell}$ in eqs. 3.26a and 3.26b .
	CGKL		Coefficient of the AREP Gaussians: $C_{k\ell}^{AREP}$ in eq. 3.26a.
	CSOGKL		Coefficient of the SOREP Gaussians: $C_{k\ell}^{SOREP}$ in eq. 3.26b.
	NKL		Exponent of the r factors: $n_{k\ell}$ in eqs. 3.26a and 3.26b.

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. SR potentials are scalar-relativistic (do not include spin-orbit operators), which FR potentials are fully-relativistic (include spin-orbit operators).

The form of SR 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):

$$W_{ps} = C + W0 + SL \tag{3.16}$$

where:

$$C = -Z_N/r \tag{3.17}$$

$$W0 = \sum_{k=1}^{M} r^{n_k} C_k e^{-\alpha_k r^2} \tag{3.18}$$

$$SL = \sum_{\ell=0}^{4} [\sum_{k=1}^{M_\ell} r^{n_{k\ell}} C_{k\ell} e^{-\alpha_{k\ell} r^2}] P_\ell \tag{3.19}$$

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.

The form of the FR relativistic effective-core potential (REP), implemented in CRYSTAL22 [63], is:

$$W_{ps}^{\mathrm{REP}} = C + \sum_{l=0}^{\infty} \sum_{j=j_\alpha}^{j_\beta} U_{lj}^{\mathrm{REP}} P_{lj}^{\mathrm{Pauli}} \tag{3.20}$$

with C defined in Eq. (3.17), and the abbreviated notation $j_\alpha = |l + 1/2|$ and $j_\beta = |l - 1/2|$ have been introduced. Eq. (3.20) includes a radial part U_{lj}^{REP} and an angular part P_{lj}^{Pauli}. The radial part is expanded in the basis of solid Gaussian functions:

$$U_{lj}^{\mathrm{REP}} = \sum_{k=1}^{M_\ell} r^{n_{k\ell}} C_{k\ell}^{\mathrm{REP}} e^{-\alpha_{k\ell} r^2} \tag{3.21}$$

whereas the angular part is written in the basis of two-component Pauli spinors:

$$P_{lj}^{\mathrm{Pauli}} = \sum_{m_j=-j}^{j} |l, j, m_j\rangle\langle l, j, m_j| \tag{3.22}$$

It is possible to then write W_{ps}^{REP} in terms of spin-averaged (AREP) and spin-dependent (SOREP) contributions, as first suggested by Ermler, Pitzer and co-workers [98]:

$$W_{ps}^{\mathrm{REP}} = C + W_{ps}^{\mathrm{AREP}} + X \times W_{ps}^{\mathrm{SOREP}} \tag{3.23}$$

where X is a global scaling factor for the SOREP operator (inserted in input with the SOSCALE record).
The purely scalar-relativistic AREP reads:

$$W_{ps}^{\mathrm{AREP}} \approx U_L^{\mathrm{AREP}} + \sum_{l=0}^{L} \left[U_l^{\mathrm{AREP}} - U_L^{\mathrm{AREP}}\right] P_l^{\mathrm{Schröd}} \tag{3.24a}$$

and the spin-orbit SOREP reads:

$$W_{ps}^{\mathrm{SOREP}} \approx U_L^{\mathrm{SOREP}} + \sum_{l=0}^{L} \left[U_l^{\mathrm{SOREP}} - U_L^{\mathrm{SOREP}}\right] P_l^{\mathrm{Schröd}}\ \mathbf{L}\cdot\mathbf{S}\ P_l^{\mathrm{Schröd}} \tag{3.24b}$$

where $\mathbf{L}$ is the electron-nuclear angular momentum operator and $\mathbf{S}$ is the one-electron spin operator. In Eqs. (3.24a) and (3.24b), the approximated $\approx$ sign is used because the sum over l is truncated (up to a value of $L = 5$, with the present implementation). The angular operators $P_l^{\mathrm{Schröd}}$ are written in the basis of one-component Schrödinger spinors:

$$P_l^{\mathrm{Schröd}} = \sum_{m=-l}^{l} |l, m\rangle\langle l, m| \tag{3.25}$$

and the radial operators are defined using those of Eq. (3.21) as:

$$U_l^{\rm AREP} = \frac{1}{2l+1}\left[lU_{lj^\beta}^{\rm REP} + (l+1)U_{lj^\alpha}^{\rm REP}\right] \equiv \sum_{k=1}^{M_\ell} r^{n_{k\ell}} C_{k\ell}^{\rm AREP} e^{-\alpha_{k\ell} r^2} \tag{3.26a}$$

and:

$$U_l^{\rm SOREP} = \frac{2}{2l+1}\left[U_{lj^\alpha}^{\rm REP} - U_{lj^\beta}^{\rm REP}\right] \equiv \frac{2}{2l+1}\sum_{k=1}^{M_\ell} r^{n_{k\ell}} C_{k\ell}^{\rm SOREP} e^{-\alpha_{k\ell} r^2} \tag{3.26b}$$

It is important to note that different authors use different definitions for $U_l^{\rm SOREP}$, which must be taken into account when using the keyword **INPSOC**, to define the SOREP manually in the input. This is achieved by the PSONAME record. Eq. (3.26b) corresponds to the definition obtained by invoking the **STUTTGART** keyword.
The **INTERNAL** definition is:

$$U_l^{\rm SOREP} \equiv \sum_{k=1}^{M_\ell} r^{n_{k\ell}} C_{k\ell}^{\rm SOREP} e^{-\alpha_{k\ell} r^2} \tag{3.27}$$

i.e. the factor $\frac{2}{2l+1}$ is absorbed in the definition of $C_{k\ell}^{\rm SOREP}$.
The **COLUMBUS** and **TOULOUSE** definitions are:

$$U_l^{\rm SOREP} \equiv \frac{2}{l}\sum_{k=1}^{M_\ell} r^{n_{k\ell}} C_{k\ell}^{\rm SOREP} e^{-\alpha_{k\ell} r^2} \tag{3.28}$$

1. Hay and Wadt (HW) ECP ([142, 140]) are of the general form 3.16. In this case, the NKL value given in the tables of ref. [142, 140] must be decreased by 2 (2 → 0, 1 → -1, 0 → -2).
2. Durand and Barthelat (DB) ([13] - [84], [14], [15]), and Stuttgart-Dresden [232] 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).

$$SL = \sum_{\ell=0}^{3} e^{-\alpha_\ell r^2}[\sum_{k=1}^{M_\ell} r^{n_{k\ell}} C_{k\ell}]P_\ell \tag{3.29}$$

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 $\varphi_\mu^0\ \varphi_\nu^g$ (page 395) 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 2.3).

As far as SOREP integrals are concerned, their evaluation follows very closely the strategy used for AREP integrals [186], as explained in Refs. [228, 229]. The routines of Ref. [229], have been modified from the ones kindly provided to us by Russell Pitzer, as also implemented in the EPCISO program of Ref. [277]. The same **TOLPSEUD** tolerance, mentionned in the previous paragraph, is used for screening both AREP and SOREP integrals.

Example Manual Input of the Effective Core Potential

The following example provides an input corresponding to the ECP28MWB potential for Eu, available at *http://www.theochem.uni-stuttgart.de/pseudopotentials/clickpse.en.html*:

```
263 17                                  - The '2' in 263 indicates an ECP, rather than AE treatment
INPSOC                                  - INPSOC (rather than INPUT) says that you want a SOREP operator
INTERNAL 1.0                            - Internal convention for defining SOREP coefficients
35. 0 1 1 1 1 1                         - zero l = L = 5 terms, one l = 0,..., one l = 4 terms
23.471384 607.659331  0.000000 0        - exponents of Gaussians, C^AREP, C^SOREP, powers of r (everywhere zero)
16.772479 264.385476 19.869243 0
13.981343 115.381375  1.523881 0        - only three C^SOREP coefficients (for l = 2,3,4) are non-zero
23.962888 -49.400794  0.399191 0
21.232458 -26.748273  0.000000 0
                                        - insert here 17 shells for the valence basis set
```

The following second example provides a manual input, corresponding to the **COLUSC** potential for Cr, but with upwards scaling of the SOREP operator by a factor of 2:

```
224 20                                   - atomic number > 200, so element 24 (Cr) is treated with an ECP
INPSOC                                   - input given manually with both AREP and SOREP. You must be an expert!
COLUMBUS 2.0                             - Scaling of the SOC operator by a factor X = 2.0
14. 4 6 6 0 0 0                          - Z_N = 14, so 24 - 14 = 10 electrons are effectively treated by the ECP
10.0191000       -3.359513  0.019023  0     - 4 exponents, C^AREP and C^SOREP, powers of r for l = L = 2 term
29.9265000      -24.589570  0.108092  0
97.3904000      -61.441873  0.499853  0
329.051200       -8.379354  0.676580 -1
  3.083600      -28.406078  0.000000  0     - 6 exponents C^AREP and C^SOREP, powers of r for l = 0 term
  3.577700       91.176165  0.000000  0
  4.652700     -171.883828  0.000000  0
  6.151000      181.425115  0.000000  0
 23.502400       10.198945  0.000000 -1
 20.169100        3.059802  0.000000 -2
  2.440000      -18.258059 -0.429570  0     - 6 exponents C^AREP and C^SOREP, powers of r for l = 1 term
  2.892200       59.418142  1.547118  0
  3.883000     -113.841801 -3.486250  0
  5.255400      117.491030  4.201010  0
  8.935300       -4.185364 -0.788318 -1
 18.839800        5.408047  0.049254 -2
                                         - 20 shells for valence basis set are inserted here
```

This third example provides the same **COLUSC** potential for Cr, but without the SOREP part (only the scalar-relativistic AREP is provided):

```
224 20
INPUT                                    - INPUT requests manual input of AREP operator
14. 4 6 6 0 0 0
10.0191000       -3.359513  0            - only three rows of data (instead of four, as above)
29.9265000      -24.589570  0
97.3904000      -61.441873  0
329.051200       -8.379354 -1
  3.083600      -28.406078  0            - exponents of the Gaussians, C^AREP coefficients and powers of r
  3.577700       91.176165  0
  4.652700     -171.883828  0
  6.151000      181.425115  0
 23.502400       10.198945 -1
 20.169100        3.059802 -2
  2.440000      -18.258059  0
  2.892200       59.418142  0
  3.883000     -113.841801  0
  5.255400      117.491030  0
  8.935300       -4.185364 -1
 18.839800        5.408047 -2
                                        - 20 shells for valence basis set are inserted here
```

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			

COLUMBUS SMALL CORE ECP (WITH NUMBER OF ELECTRONS TREATED BY THE ECP)

Li	Be											B	C	N	O	F	Ne
2	2											2	2	2	2	2	2
Na	Mg											Al	Si	P	S	Cl	Ar
2	2											10	10	10	10	10	10
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
10	10	10	10	10	10	10	10	10	10	10	10	18	18	18	18	18	18
Rb	Sr	Y	Zr	Ni	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
28	28	28	28	28	28	28	28	28	28	28	28	36	36	36	36	36	36
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg						
46	46	46	60	60	60	60	60	60	60	60	60						

COLUMBUS LARGE CORE ECP (WITH NUMBER OF ELECTRONS TREATED BY THE ECP)

	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
	18	18	18	18	18	18	18	18	18	18	18	28	28	28	28	28
	Y	Zr	Ni	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
	36	36	36	36	36	36	36	36	36	36	46	46	46	46	46	46
	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
	54	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68

Fr	Ra	Ac
78	78	78

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
54	54	54	54	54	54	54	54	54	54	54	54	54	54	54

Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es
78	78	78	78	78	78	78	78	78	78	78

COLUMBUS ECP FOR SUPER-HEAVY ELEMENTS (ATOMIC NUMBER > 99)

Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Nh	Fl	Mc	Lv	Ts	Og
78	78	78	78	78	78	78	78	78	78	78	78	78	78	78

Fm	Md	No	Lr
78	78	78	78

STUTTGART-COLOGNE SMALL CORE ECP (WITH NUMBER OF ELECTRONS TREATED BY THE ECP)

K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Rb	Sr	Y	Zr	Ni	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28
Cs	Ba		Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
46	46		60	60	60	60	60	60	60	60	60	60	60	60	60	60	60

Fr	Ra	Ac
78	78	60

Ac	Th	Pa	U
60	60	60	60

STUTTGART-COLOGNE LARGE CORE ECP (WITH NUMBER OF ELECTRONS TREATED BY THE ECP)

Ge	As	Se	Br
28	28	28	28
Sn	Sb	Te	I
46	46	46	46
Pb	Bi	Po	At
78	78	78	78

STUTTGART-COLOGNE ECP FOR SUPER-HEAVY ELEMENTS (ATOMIC NUMBER > 99)

Rg	Cn	Nh	Fl	Mc	Lv	Ts	Og
92	92	92	92	92	92	92	92

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 [142, 141]. Agreement with published atomic energies data is satisfactory (checked from Na to Ba) for Hay and Wadt small core and large core pseudopotentials, 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.

Only the FR **COLUSC**, **COLULC**, **COLUSH**, **STUTSC**, **STUTLC** and **STUTSH** libraries include spin-orbit SOREP operators in their definition.
The Columbus potentials **COLUSC**, **COLULC** and **COLUSH** [102, 150, 163, 242, 243, 99, 193] are shape-consistent FR potentials obtained from adjustment to Dirac-Fock calculations, employing the Dirac-Coulomb Hamiltonian.
The Stuttgart-Cologne potentials **STUTSC**, **STUTLC** and **STUTSH** [188, 189, 263, 213, 107, 170, 171, 214, 212, 69, 288, 136, 135, 137] are energy-consistent FR potentials obtained from adjustment to multiconfigurational Dirac-Fock calculations, employing the low-frequency or frequency-dependent Dirac-Coulomb-Breit Hamiltonian.
If both the Columbus and Stuttgart-Cologne potentials are available for a calculation, then it is usually recommended to use the Stuttgart-Cologne one. The Stuttgart-Cologne potentials usually have a smaller core than the Columbus ones and are usually obtained by adjustment to models including more accurate approximations to the relativistic electron-electron interaction. We note that the many-body core-polarization operators of the **STUTLC** potentials have not been implemented.

Valence Basis set and pseudopotentials

Hay and Wadt ([142, 141]) 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 93), or any atomic package with effective core pseudopotentials. See Chapter 17 for general comments on atomic basis function optimization. Bouteiller *et al* [31] have published a series of basis sets optimized for Durand and Barthelat ECPs.

As for the Columbus and Stuttgart-Cologne ECP libraries, molecular basis sets are available online in clickable periodic tables at:
`http://people.clarkson.edu/~pchristi/reps.html`
`http://www.tc.uni-koeln.de/PP/clickpse.en.html`
These molecular basis sets can often be modified for calculations in periodic system by decontracting them and (possibly) removing the most diffuse functions.

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 in:
http://www.theochem.uni-stuttgart.de/pseudopotentials/index.en.html

The clickable periodic table supplies, in CRYSTAL format, ECP to be used in CRYSTAL via the INPUT keyword (basis set input, block2, page 79).
http://www.theochem.uni-stuttgart.de/pseudopotentials/clickpse.en.html

The **STUTSC**, **STUTLC** (and **STUTSH**) internal libraries correspond to the so-called "ECPXXMDF" ("ECPXXMDFQ"), as avalable at the address *http://www.theochem.uni-stuttgart.de/pseudopotentials/index.en.html*, where XX denotes the effective core charge.

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. [261, 262]. 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 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.

3.3 Computational parameters, Hamiltonian, SCF control

Single particle Hamiltonian

RHF	Restricted Hartree Fock		118	–
UHF	Unrestricted Hartree-Fock		126	–
ROHF	Restricted Open Shell Hartree-Fock		126	–
DFT	DFT Hamiltonian		126	–
	SPIN	spin-polarized solution	135	–

Choice of the exchange-correlation functionals

	EXCHANGE	exchange functional	129	I
	LDA functionals			
		LDA Dirac-Slater [68] (LDA)		
		VBH von Barth-Hedin [281] (LDA)		
	GGA functionals			
		BECKE Becke 1988[18] (GGA)		
		PBE Perdew-Becke-Ernzerhof 1996 [204] (GGA)		
		PBESOL GGA. PBE functional revised for solids [206]		
		mPW91 modified Perdew-Wang 91 (GGA)		
		PWGGA Perdew-Wang 91 (GGA)		
		SOGGA second order GGA. [302]		
		WCGGA GGA - Wu-Cohen [293]		
	CORRELAT	correlation functional	129	I
	LDA functionals			
		PZ Perdew-Zunger [210] (LDA)		
		VBH von Barth-Hedin [281] (LDA)		
		VWN Vosko,-Wilk-Nusair [282] (LDA)		
	GGA functionals			
		LYP Lee-Yang-Parr [167] (GGA)		
		P86 Perdew 86 [202] (GGA)		
		PBE Perdew-Becke-Ernzerhof [204] (GGA)		
		PBESOL GGA. PBE functional revised for solids [206]		
		PWGGA Perdew-Wang 91 (GGA)		
		PWLSD Perdew-Wang 92 [208, 209, 207] (GGA)		
		WL GGA - Wilson-Levy [291]		
	meta-GGA functionals			
		B95 mGGA - Becke 95 [22]		
	Standalone keywords: exchange+correlation			
	SVWN	see [68, 282]	129	
	BLYP	see [18, 167]	129	
	PBEXC	see [204]	129	
	PBESOLXC	see [206]	129	
	SOGGAXC	see [302]	129	
	SOGGA11	see [216]	129	

Global Hybrid functionals			
Standalone keywords			
B3PW	B3PW parameterization	130	–
B3LYP	B3LYP parameterization	130	–
PBE0	Adamo and Barone [5]	130	
PBESOL0	Derived from PBE0	130	
B1WC	see [25]	130	
WC1LYP	see [60]	130	
B97H	see [4, 101]	130	
PBE0-13	see [42]	130	
SOGGA11X	see [215]	130	
mPW1PW91	see [6]	130	
mPW1K	see [175]	130	
User defined global hybrids			
HYBRID	hybrid mixing	130	I
NONLOCAL	local term parameterization	131	I
Range-Separated Hybrid functionals			
Short-range Corrected RSH functionals			
HSE06	Screened-Coulonb PBE XC functional [2, 204]	132	–
HSESOL	Screened-Coulomb PBESOL XC functional [162, 206]	132	–
SC-BLYP	SC-RSH BLYP functional based on ITYH scheme [152, 271]	132	–
Middle-range Corrected RSH functionals			
HISS	MC based on PBE XC functional [267, 268]	132	–
Long-range Corrected RSH functionals			
RSHXLDA	LC LDA XC functional [1, 151]	132	–
wB97	Chai/Head-Gordon LC functional [153, 4]	132	–
wB97X	Chai/Head-Gordon SC/LC functional [153, 4]	132	–
LC-wPBE	LC hybrid based on PBE XC functional [85]	132	–
LC-wPBESOL	LC hybrid based on PBESOL XC functional [85]	132	–
LC-wBLYP	LC hybrid based on BLYP XC functional [85]	132	–
LC-BLYP	LC-RSH BLYP functional based on ITYH scheme [152, 271]	132	–
CAM-B3LYP	Coulomb-Attenuating method)[270] based on the BLYP XC functional	132	–
LC-PBE	LC-RSH PBE functional based on ITYH scheme [152, 164]	132	–
User defined range separated hybrids			
SR-OMEGA	setting of ω_{SR} for SC hybrids	133	–
MR-OMEGA	setting of ω_{SR} and ω_{LR} for MC hybrids	133	–
LR-OMEGA	setting of ω_{LR} for LC hybrids	133	–
SR-HYB_WB97X	Option to change the amount of SR-HF exchange in the ωB97-X functional	133	–
LSRSH-PBE	User-controllable RSH x-functional based on the PBE functional	133	–

meta-GGA functionals

Pure mGGA functionals

M06L	pure mGGA M06-type functional [296]	134	–
revM06L	revised version of pure mGGA M06L functional [284]	134	–
MN15L	58 parameters pure mGGA by Yu, He, and Truhlar [301]	134	–
SCAN	non-empirical mGGA functional based on SCAN XC [264]	134	–
r^{2}SCAN	modified version of regularized SCAN [114]	134	–

Global hybrid mGGA functionals

B1B95	Becke88 exchange combined with Becke95 correlation and 28% of HF exchange [23]	134	–
mPW1B95	mPW91 exchange combined with Becke95 correlation and 31% of HF exchange [303]	134	–
mPW1B1K	mPW91 exchange combined with Becke95 correlation and 44% of HF exchange [303]	134	–
PW6B95	6 parameters mPW91 exchange combined with re-optimized Becke95 correlation and 28% of HF exchange [304]	134	–
PWB6K	6 parameters mPW91 GGA combined with re-optimized Becke95 correlation and 46% of HF exchange [304]	134	–
M05	Minnesota 2005 functional [300]	134	–
M052x	M05-2X functional [299]	134	–
M06	Minnesota 2006 functional [298]	134	–
M062X	M06-2X functional [298]	134	–
M06HF	M06-type functional with 100% HF [296]	134	–
MN15	59 parameters global hybrid mGGA with 44% of HF exchange [138]	134	–
revM06	revised version of global hybrid mGGA M06 with 40.41% of HF exchange [285]	134	–
SCAN0	global hybrid mGGA based on SCAN XC with 25% of HF exchange [149]	134	–
r^{2}SCANh	global hybrid mGGA based on r^{2}SCAN XC with 10% of HF exchange [40]	134	–
r^{2}SCAN0	global hybrid mGGA based on r^{2}SCAN XC with 25% of HF exchange [40]	134	–
r^{2}SCAN50	global hybrid mGGA based on r^{2}SCAN XC with 50% of HF exchange [40]	134	–

Numerical accuracy control

ANGULAR	definition of angular grid	137	I
RADIAL	definition of radial grid	137	I
[BECKE]	selection of Becke weights (default)	137	–
SAVIN	selection of Savin weights	137	–
OLDGRID	"old" default grid	140	
LGRID	"large" predefined grid	140	
[XLGRID]	"extra large" predefined grid (new default)		
XXLGRID	"extra extra large" predefined grid	140	
XXXLGRID	"extra extra extra large" predefined grid (default grid for mGGA based functionals)	141	
HUGEGRID	"ultra extra large grid" predefined grid (default for SCAN based DFAs)	141	
RADSAFE	safety radius for grid point screening		I
TOLLDENS	density contribution screening [6]	141	I
TOLLGRID	grid points screening [14]	141	I
[BATCHPNT]	grid point grouping for integration	141	I
CHUNKS	max n. points in a batch for numerical int.	142	I
DISTGRID	distribution of DFT grid across nodes	142	
LIMBEK	size of local arrays for integration weights [400]	142	I

Atomic parameters control

RADIUS	customized atomic radius	143	I
FCHARGE	customized formal atomic charge	143	I

Auxiliary

PRINTEXC Enable printing of the Exchange and Correlation energy at each SCF cycle. By default in CRYSTAL22 the Exchange and Correlation energy is printed only when the SCF convergence is reached.

END close DFT input block

Numerical accuracy and computational parameters control

BIPOLAR	Bipolar expansion of bielectronic integrals	95	I
BIPOSIZE	size of coulomb bipolar expansion buffer	95	I
EXCHSIZE	size of exchange bipolar expansion buffer	95	I
EXCHPERM	use permutation of centers in exchange integrals	101	–
ILASIZE	Maximum size of array ILA for 2-electron integral calculation [6000]	112	I
INTGPACK	Choice of integrals package [0]	112	I
MADELIND	reciprocal lattice vector indices for Madelung sums [50]	114	I
NOBIPCOU	Coulomb bielectronic integrals computed exactly	116	–
NOBIPEXCH	Exchange bielectronic integrals computed exactly	116	–
NOBIPOLA	All bielectronic integrals computed exactly	116	–
POLEORDR	Maximum order of multipolar expansion [4]	117	I
TOLINTEG	Truncation criteria for bielectronic integrals [6 6 6 6 12]	125	I
TOLPSEUD	Pseudopotential tolerance [6]	125	I

Type of run

ATOMHF	Atomic wave functions	93	I
SCFDIR	SCF direct (mono+biel int computed)	118	–
EIGS	S(**k**) eigenvalues - basis set linear dependence check	98	–
FIXINDEX	Reference geometry to classify integrals	104	–

Basis set - AO occupancy

FDAOSYM	f and d degeneracies analysis	101	I
FDAOCCUP	f and d orbital occupation guess	102	I
GUESDUAL	Density matrix guess - different Basis set	108	I

Integral file distribution

BIESPLIT	writing of bielectronic integrals in n files [$n = 1$],max=10	95	I
MONSPLIT	writing of mono-electronic integrals in n file [$n = 1$], max=10	115	I

Numerical accuracy control and convergence tools

ANDERSON	Fock matrix mixing	93	I
BROYDEN	Fock matrix mixing	95	I
DIIS	Fock matrix mixing/convergence accelerator	96	I
FMIXING	Fock/KS matrix (cycle i and i-1) mixing [0]	105	I
LEVSHIFT	level shifter [no]	113	I
MAXCYCLE	maximum number of cycles [50]	114	I
SMEAR	Finite temperature smearing of the Fermi surface [no]	121	I
TOLDEE	convergence on total energy [6]	125	I

Initial guess

EIGSHIFT	alteration of orbital occupation before SCF [no]	99	I
EIGSHROT	rotation of the reference frame [no]	99	I
GUESSP	density matrix from a previous run	109	–
GUESSPAT	superposition of atomic densities	110	–

Spin-polarized system

ATOMSPIN	setting of atomic spin to compute atomic densities	94	I
BETALOCK	beta electrons locking	94	I
SPINLOCK	spin difference locking	124	I
SPINEDIT	editing of the spin density matrix used as SCF guess	123	I

Auxiliary and control keywords

END	terminate processing of block3 input		–
FULLTIME	detailed report on running time	106	–
KSYMMPRT	printing of Bloch functions symmetry analysis	112	–
LOWMEM	inhibits allocation of large arrays	114	–
NOLOWMEM	allows allocation of large arrays	114	–
MAXNEIGHB	maximum number of equidistant neighbours from an atom	49	I
NEIGHBOR	number of neighbours to analyse in **PPAN**	59	I
MEMOPRT	Synthetic report about dynamic memory usage	114	–
MEMOPRT2	Detailed report about dynamic memory usage	114	–
PRINTOUT	setting of printing options	64	I
QVRSGDIM	maximum size of mutipole moment gradient array [90000000]	118	I
NOSYMADA	No Symmetry Adapted Bloch Functions	117	–
SYMADAPT	Symmetry Adapted Bloch Functions (default)	124	–
SETINF	setting of inf array options	66	I
SETPRINT	setting of printing options	66	I
STOP	execution stops immediately	68	–
TESTPDIM	stop after symmetry analysis	124	–
TEST[RUN]	stop after integrals classification and disk storage estimate	125	–

Restricted to *MPPcrystal*

CMPLXFAC	Overloading in handling matrices at "complex" k points with respect to "real" k points [2.3]	96	I
REPLDATA	to run *MPPcrystal* as *Pcrystal*	118	–
STDIAG	Enable standard diagonalization method (D&C method disabled)	124	–

Output of data on external units			
NOFMWF	wave function formatted output not written in file fort.98.	116	–
SAVEWF	wave function data written every two SCF cycles	119	–

Post SCF calculations			
POSTSCF	post-scf calculations when convergence criteria not satisfied	117	–
EXCHGENE	exchange energy evaluation (spin polarized only)	101	–
GRADCAL	analytical gradient of the energy	106	–
PPAN	population analysis at the end of the SCF [no]	117	

Computational parameters, Hamiltonian, SCF control optional keywords

ANDERSON

Anderson's method [7], as proposed by Hamann [134], is applied. No input data are required. See test49_dft, a metallic Lithium 5 layers slab, PWGGA Hamiltonian.

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 [240]. 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 343), 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 2.2.

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 **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.

Note: if INF98 is < 0, then the lock duration is controlled by energy difference between successive cycles (disappears when $< 10^{-INF98}$) instead of by number of cycles. This is useful when the locking must be relaxed before the convergence of the SCF -otherwise there is a risk to arrive to a spurious solution- but at the same time one does not know exactly in which SCF cycle the density matrix can be considered to be close enough to the right electronic structure so as to avoid a divergent behavior after the locking finishes.
Situations like this may occur in geometry optimizations of ferromagnetic systems, for instance: the SCF of the first optimization point converges in CYC 20 with the locking finishing in CYC 12; in the remaining points, as the density matrix of the previous one is used as initial guess, the SCF converges in CYC 10 under locking and, so, the true convergence is not ensured. Using FINALRUN=4 this situation could lead to an extremely large or even non convergent optimization process.

On the other hand, to fix locking since the energy difference between cycles is less than a given threshold is a quite handy criterion that could be in several cases preferable than fixing the duration in terms of number of cycles.

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 18
	∗	ITEXCH	overlap threshold for exchange 14

The bipolar approximation is applied in the evaluation of the Coulomb and exchange integrals (page 399). ITCOUL and ITEXCH can be assigned any intermediate value between the default values (18 and 14) (see page 399) and the values switching off the bipolar expansion (20000 and 20000).
Note that default values have become tighter, 18 and 14, from CRYSTAL14, **whereas before they were 14 and 10, respectively. Results are now expected to be more accurate but the program might be slower.**

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 is 4000000, that is 32 Mb, per core). 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 [154]
	∗	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 [37] scheme, following the method proposed by Johnson [154], 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
```

See test50_dft, a metallic Lithium 5 layers slab, PWGGA Hamiltonian.

CMPLXFAC - Weight for diagonalization time

This directive is supported by *MPPcrystal* only.

	rec	variable	meaning
•	∗	WEIGHT	estimated ratio of the computational time required to diagonalize complex and real matrices. Default value: $2.\overline{3}$.

Fock matrix elements at a general k point in reciprocal space are complex numbers. At special k points (such as Γ, for example) those elements are real. Computational times required to diagonalize real (t_r) and complex (t_c) matrices are different. WEIGHT is an estimate of the ratio: WEIGHT $= t_c/t_r$.
An appropriate assessement of WEIGHT improves load balancing. The default value is: $2.\overline{3}$.

DIIS – Direct Inversion of the Iterative Subspace convergence accelerator

The Direct Inversion of Invariant Subspace (DIIS) is a powerful technique to accelerate the SCF convergence that gained success in recent years. Initially proposed by Pulay in 1984 for SCF convergence,[236, 234] it gained considerable reputation during the last 30 years. In CRYSTAL22 this feature is activated by default in sequential and parallel, while it is deactiated by default in MPPcrystal runs.

*S*elect DIIS scheme

keyword	meaning
DIIS	Activates the Γ−point only DIIS procedure [default].
DIIS1K	Equivalent to DIIS
PRTDIIS	Activates detailed printing of the Fock mixing coefficients in the DIIS procedure.
DIISALLK	DIIS is performed over all k-points of the Brillouin zone
SLOSHING	DIIS is performed over all k-points of the Brillouin zone but weighted according to Kerker factors to avoid charge sloshing.
NODIIS	Deactivates DIIS and restores FMIXING 30% + LEVSHIFT

optional additional keywords

	rec	variable	meaning
	A	**HISTDIIS**	Limits DIIS mixing to the most recent $NCYC$ cycles
•	∗	NCYC	maximum number of previous cycles to be kept
	A	**THREDIIS**	Postpones the activation of DIIS by performing Fock mixing until the difference in energy is below DEDIIS.
•	∗	DEDIIS	threshold for ΔE to activate DIIS
	A	**THRKDIIS**	All the Fock matrices having a weight lower than KDIISTHR are discarded from DIIS history.
•	∗	DIISTHR	threshold on DIIS weight for discarding earlier Fock matrices
	A	**SLOSHFAC**	Sets Kerker factor if SLOSHING is active
•	∗	FKERK	Set the parameter used for Kerker factors (k-points weights) [Default: 1.2]

Note: **DIISALLK** *and* **SLOSHING** *are not available in MPPcrystal*

Let us consider an SCF iterative procedure. Cycle n starts with the definition of a density matrix $\mathbf{D}_n(\mathbf{k})$, either obtained as an initial guess (at cycle 0) or from eigenvectors of the previous cycle. A Fock matrix $\mathbf{F}_n(\mathbf{k})$ is then obtained from this density matrix, and then diagonalized in each $\mathbf{k}$ point to obtain eigenvectors $\mathbf{C}_n(\mathbf{k})$. These will form the $\mathbf{D}_{n+1}(\mathbf{k})$ density matrices and so on. In the DIIS procedure, at each iterative cycle n, instead of the Fock matrix $\mathbf{F}_n(\mathbf{k})$, an averaged effective Hamiltonian is generated as a linear combination of the Fock matrices of previous iterations:

$$\overline{\mathbf{F}}_n = \sum_{i=1}^{n} c_i \mathbf{F}_i \ . \tag{3.30}$$

The c_i coefficients are obtained by minimizing a suitable error functional $\mathbf{e}$, subject to the constraint that $\sum_{i=1}^{n} c_i = 1$. This is obtained by solving the linear equation system:

$$\begin{pmatrix} \mathbf{e} & \mathbf{1}^T \\ \mathbf{1} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{c} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \ , \tag{3.31}$$

where $\mathbf{e}$ is an error matrix having the size of the iterative space up to cycle n and λ is a Lagrange multiplier. The error matrix is defined through scalar products of suitable error vectors in each $\mathbf{k}$ point of the Brillouin zone:

$$e_{nm} = \sum_{\mathbf{k}} f_{\mathbf{k}} < \mathbf{e}_n(\mathbf{k}) | \mathbf{e}_m(\mathbf{k}) > \ . \tag{3.32}$$

where k-dependent factors $f_{\mathbf{k}}$ are introduced.

- By default, only gamma point is used in DIIS (**DIIS1K**), and in this case $f_{\mathbf{k}} = 1$ and no summation is performed over k points.
- Sum can be extended to all k points, with factors set all equal to $1/N_{\mathbf{k}}$, by using the keyword **DIISALLK**.
- The keyword **SLOSHING** is similar to **DIISALLK**, but different weights for different k-points are used, according to Kerker's formula

According to Pulay's commutator-DIIS (CDIIS) formulation,[234] we define the error vector for the SCF procedure as:

$$\mathbf{e}_n(\mathbf{k}) = \mathbf{F}_n(\mathbf{k})\mathbf{D}_n(\mathbf{k})\mathbf{S}(\mathbf{k}) - \mathbf{S}(\mathbf{k})\mathbf{D}_n(\mathbf{k})\mathbf{F}_n(\mathbf{k}) \ , \tag{3.33}$$

where $\mathbf{S}(\mathbf{k})$ is the overlap matrix. The CDIIS formulation is particularly convenient because it allows the number of occupied orbitals in a given $\mathbf{k}$ point to vary during iterations.

DFT

The Kohn-Sham [159, 147] DFT code is controlled by keywords, that must follow the general keyword **DFT**, in any order.
The DFT input block ends with the keyword **END** or **ENDDFT**. Default values are supplied for all computational parameters. For further details see Chapter 4.

DOPING - Fractional charge doping

rec	variable	meaning
•	∗ CHGDOP	amount of electrons to be added to the system (real number)

The keyword allows to add (or remove) electrons from the system independently of the number of electrons given as initial guess in the basis set input or using CHEMOD (see page 75). The additional electronic charge is added after the first diagonalization of the SCF procedure. If the resulting cell is not neutral, a uniform compensating charge background is added automatically (similarly to keyword CHARGED, page 75). At difference with CHEMOD, through this keyword it is possible to add a fractional amount of electrons to a periodic cell, obtaining equivalent results with respect to adding one electron to an appropriate supercell of the original cell.

Warning: when adding a fractional number of electrons to a cell, particular care must be taken in verifying that the sampling of reciprocal space is well converged. The appropriate value of the shrinking factor for such a system is in some cases several times larger than for the neutral cell, and is dependent on the amount of charge doping.

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 [160], 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 2.2, page 22).
	SHIF1	α (or total, if Restricted) Fock/KS matrix shift
	[SHIF2	β Fock/KS 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/KS 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 22 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 **ATOMSYMM** (input block 1, page 38) 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 3.3, page 99)

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 (the vector joining atom 3 to atom 1, in general, is not orthogonal to the vectors joining atom 2 to atom 1; it is then orthogonalized.)
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. A rotation of the eigenvectors can be obtained in ***properties*** by entering the keyword **ROTREF**, allowing AO projected Density of States or Population Analysis orienting the cartesian frame along the principal axes of the octahedron.
Example:

```
CoF2
ROTREF
ATOMS
5
0 0 0
2
0 0 0
3
0 0 0
PPAN
END
END
```

END

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

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 18.19, page 398.
No input data are required. See tests 29, 30, 31, 38.

EXCHPERM - Use permutation of center in exchange integrals

In HF and hybrids calculations exact exchange integrals (see equation 18.19, page 398) are normally calculated in the same routine that calculates Coulomb integrals, exploiting the symmetry of the system in order to reduce the number of computed integrals. In case of systems without symmetry (having only the identity as symmetry operator), this option separates the calculation of exchange from that of Coulomb integrals, and exploits invariance of bielectronic integrals under permutation of centers instead of symmetry.
Given a bielectronic integral $(\mu\rho|\nu\sigma)$, where $\mu\ \nu\ \rho\ \sigma$ label the four centers (see chapter 18), the integrals obtained with the permutations:

$$\mu \leftrightarrow \rho\ ,\ \nu \leftrightarrow \sigma \text{ and } (\mu\rho) \leftrightarrow (\nu\sigma) \tag{3.34}$$

are related to the original one by hermiticity. Of the $2^3 = 8$ equivalent integrals only 4 are used to obtain the irreducible part of the Fock matrix, since hermiticity of the Fock matrix is already exploited. With this option the time required for the calculation of the exchange contribution to the Fock matrix is reduced to 40-50% of the normal value. The option is usable only in a SCFDIR run (see page 118) and it is automatically switched off in presence of symmetry.
No input data are required.

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 is 4000000, that is 32 Mb, per core).
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
```

FDAOSYM - f and d degeneracies analysis

	rec	variable	meaning
•	∗	NA	number of atoms for which the AOs mixing check is performed
•	∗	ISCAT(J),J=1,NA	atomic label (output order) of atoms to be checked

This keyword performs the symmetry analysis that permits to know if AOs belonging to the same shell are mixed or not by the symmetry operators that don't move the atom, to which they belong to. This analysis is implemented for d and f shell types only. The FDAOSYM keyword must be inserted in the third block of the CRYSTAL input.
In the output, the subgroup of operators that do not move the atom is performed first (ATOMSYMM keyword). Then the AOs are listed with the indication of mixing with other AOs (if

any).
This keyword is useful for partially occupied shells (d or f). AOs that mix will form a symmetry constrained degeneracy subset. If n AOs of the shell mix generating these subsets, and $m<n$ (open shell case) electrons are supposed to populate the shell, then Jahn-Teller symmetry breaking should be taken into account. If a guess AOs occupation is defined with FDOCCUP keyword or forced with EIGSHIFT keyword, information obtained by FDAOSYM can indicate which AOs will have the same occupation.
Example. Ce_2O_3 bulk. Structure $4d$ shell is completely filled and $4f$ shell contains one electron. Suppose Ce atom is labeled 1. Information produced by FDAOSYM indicates which AOs will have the same occupation. If combined with TESTRUN such information is obtained at zero cost and then a new input with FDOCCUP or EIGSHIFT keywords can be run. The input is

```
FDAOSYM
1
1
TESTRUN
```

Output obtained contains also d shell information, here just f part is reported.

```
 ANALYSIS FOR SHELL TYPE F
  COMPONENT     MIXES WITH     SYMMETRY OPERATORS
      1
      2             3            2  3 10 12
      3             2            2  3 10 12
      4             5            2  3 10 12
      5             4            2  3 10 12
      6
      7
```

That is: AOs 1, 6 and 7 do not mix with any other AO, 2 and 3 mix as well as 4 and 5. Label and characteristic polynomial for quantum angular symmetry d and f are reported in the FDOCCUP keyword description.

FDOCCUP - f and d orbital occupation guess

rec	variable	meaning
• ∗	INSELOC	number of records describing the occupation of d or f shells
		insert INSELOC records II
	IATNUM	atom label, output order
	ISHATO	shell label, atomic list
	SHTYP	angular quantum label of ISHATO(I)
• ∗	IAOCALPH(J)	α occupation of the 5 d (J=1,5) or 7 f (J=1,7) AOs of the shell
• ∗	IAOC BETA(J)	β occupation of the 5 d (J=1,5) or 7 f (J=1,7) AOs of the shell

Label and characteristic polynomial for quantum angular symmetry d and f are reported in the following. Spherical harmonics are used, it follows that 5 and 7 functions are reported for d and f symmetry respectively.

Symmetry d, label 3		Symmetry f, label 4	
Label	Polynomial	label	Polynomial
1	$(2z^2 - x^2 - y^2)$	1	$(2z^2 - 3x^2 - 3y^2)$
2	xz	2	$x(4z^2 - x^2 - y^2))$
3	yz	3	$y(4z^2 - x^2 - y^2)$
4	$(x^2 - y^2)$	4	$z(x^2 - y^2)$
5	xy	5	(xyz)
		6	$x(x^2 - 3y^2)$
		7	$y(3x^2 - y^2)$

The option FDOCCUP permits to define the occupation of specific f or d orbitals in a given shell in the initial guess calculation. This option can be used only for open shell cases, where electrons belonging to partially filled shells can be assigned to selected AOs.

INSELOC specifies the number of shell occupation descriptions; for each shell, the atom IATNUM (output order) and the shell label ISHATO (basis set order) is specified. As a cross check, also the shell label type SHTYP (3 and 4 for d and f respectively) is required. Finally the α and β occupation for the 5 (d) or 7 (f) AOs in the shell is indicated. The occupation numbers in IAOCALPH(J) and IAOCBETA(J) are normalized *a posteriori* to the number of d or f electrons resulting from the input charges.

Note that the input information must be inserted just one time per atomic number (say 26, Fe) and IATNUM can be any of the output order atom labels for the selected atom. See EXAMPLE 1 for test case.

To attribute different AOs guess occupation to atoms with the same atomic number, it is just needed to insert as many input as the number of different occupation one wants to set with IATNUM indicating atom labels of atoms with same atomic number but not symmetry related. If different AOs occupation for atoms symmetry related is required, the program stops and an indication of symmetry relations between atoms is reported (as the one obtained by the use of ATOMSYMM keyword).
FDOCCUP can be used in conjunction with EIGSHIFT, ATOMSPIN and SPINLOCK. Note that FDOCCUP and ATOMSPIN act at CYCLE 0, EIGSHIFT at CYCLE 1 and SPINLOCK works from CYCLE 1 for a defined number of cycles.
On the contrary FDOCCUP, as ATOMSPIN, is incompatible with SPINEDIT, where the initial guess calculation is bypassed because the initial guess is obtained by a previous calculation.

Example. Ce_2O_3. The Ce^{3+} ion is required to have one f electron. Suppose Ce is the first atom in the input list and that the seventh shell in the Basis Set list is the f shell with charge one. Suppose to distribute the f electron in two AOs with similar occupation: $f_{(2z^2-3x^2-3y^2)z}$ (first f AO) and $f_{(x^2-3y^2)y}$ (sixth f AO). Use of FDOCCUP in this case is reported in the following:

```
FDOCCUP
1
1 7 4
1 0 0 0 0 1 0
0 0 0 0 0 0 0
```

Output is printed in the ATOMIC WAVEFUNCTION(S) part and is reported in the following. Note that the nuclear charge is 30 instead of 58 because pseudopotential is used.

```
 NUCLEAR CHARGE 30.0  SYMMETRY SPECIES            S    P    D
 N. ELECTRONS   27.0  NUMBER OF PRIMITIVE GTOS   11   11    5
                      NUMBER OF CONTRACTED GTOS   4    4    1
                      NUMBER OF CLOSED SHELLS     2    2    1
                      OPEN SHELL OCCUPATION       0    0    0

  ZNUC SCFIT  TOTAL HF ENERGY   KINETIC ENERGY   VIRIAL THEOREM ACCURACY
  30.0   7   -4.706164398E+02  1.769968771E+02 -3.658896854E+00  3.4E-06

 FDOCCUP ACTIVE - ATOM   1 SHELL  6 F  SYMMETRY WITH CHARGE  1.0
                 DIAGONAL ELEMENTS OF DENSITY MATRIX FOR SHELL   6
                 ALPHA+BETA   0.500 0.000 0.000 0.000 0.000 0.500 0.000
                 ALPHA-BETA   0.500 0.000 0.000 0.000 0.000 0.500 0.000
```

Note: the printed information refers to $\alpha + \beta$ and $\alpha - \beta$, while in input α and β are inserted separately.

FIXINDEX

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 18) can lead to the selection of different numbers of bi-electronic 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 2.1; basis set, Section 2.2; general information, Section 2.3; SCF, Section 2.3), 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 14
	or
BASE	restart with new basis set
	insert basis set input, page 20
	or
GEBA	restart with new basis set and new geometrical parameters
	insert geometry input, page 14
	insert basis set input, page 20

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 2.1

1b *geometry editing keywords (optional; Section 3.1)*

END

2 standard basis set input (reference basis set). Section 2.2

2b *basis set related keywords (optional; Section 3.2)*

END

3 **FIXINDEX**

3 *block3 keywords*

END

GEOM	**BASE**	**GEBA**
⇓	⇓	⇓
geometry input(p 14)	basis set input(p 20)	geometry input (p 14)
(block 1, *1b*)	(block 2, *2b*)	(block 1, *1b*)
END	**END**	**END**
		basis set input(p 20)
		(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.

Note In geometry optimization (**OPTGEOM**, page 182) FIXINDEX is automatically set, with input geometry as reference geometry. See keyword **FINALRUN**, page 189 to redefine the reference geometry.

See tests 5 and 20.

FMIXING - Fock/KS matrix mixing

rec		variable	meaning	
•	∗	IPMIX	percent of Fock/KS matrices mixing	30

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.
Note that from CRYSTAL14, **a Fock/KS mixing of 30 % is activated by default. Set** IPMIX = 0 **to switch this option off.**

FULLTIME - Detailed timing report

A more detailed report of the timing data is generated:

```
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT SHELXG  TELAPSE   19.68 TCPU   18.42
WWWWWWWWWWWWWW SHELXG      MX    1.07 MN    1.07  MD   1.07
QQQQQQQQQQQQQQ 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.

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 191; **NUMGRATO**, page 191; **NUMGRCEL**, page 191;), analytical gradient is not computed.

GRIMME - Grimme dispersion correction for energy and gradient

The keyword **GRIMME**, inserted in third input block, calculates a London-type pairwise empirical correction to the energy as proposed by Grimme [125], to include long-range dispersion contributions to the computed ab initio total energy and gradients.

Therefore geometry optimization and vibrational frequency calculation can be carried out by including the empirical correction.

	rec	variable	meaning
•	*A*	**GRIMME**	keyword
•	*	s_6	scaling factor
	*	d	steepness
	*	R_{cut}	cutoff distance to truncate direct lattice summation
•	*	NATS	number of atomic species
II			insert NATS records II
•	*	NAT	"conventional" atomic number (page 21).
	*	C_6	dispersion coefficient for atomic number NAT (Jnm^6mol^{-1})
	*	R_{vdw}	van der Waals radius for atomic number NAT (Å)

The keyword GRIMME is followed by a set of computational parameters [i.e. scaling factor, steepness and cutoff radius for **g** (direct lattice) summation], the number of atomic species and for each atomic species the "*conventional atomic number*" (page 21) and the value of C_6 and R_{vdw} (in Jnm^6mol^{-1} and Å, respectively).

Note that the atomic number is defined as the "conventional atomic number". In this way, if one wants to use different atomic parameters for the same atomic species, a different conventional atomic number can be used. For example, 1 and 101 for hydrogen.

The current implementation has been mainly tested and used in combination with the B3LYP method [50, 51, 276], even if it can be applied to whatever level of theory.

The original proposal by Grimme was to augment DFT methods with an empirical London-type correction. To that purpose, Grimme defined a general set of atomic empirical parameters (see

Table 1 of ref. [125]) and used an optimized scaling factor to adjust the dispersion correction for each DFT method.

The total energy is given by

$$E_{DFT-D} = E_{KS-DFT} + E_{disp}$$

where E_{disp}, the empirical dispersion correction, is given as

$$E_{disp} = -s_6 \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} \sum_{\mathbf{g}} \frac{C_6^{ij}}{R_{ij,\mathbf{g}}} f_{dump}(R_{ij,\mathbf{g}})$$

The summation is over all atom pairs and $\mathbf{g}$ lattice vectors with the exclusion of the i=j contribution (i.e. self interaction) for $\mathbf{g} = 0$, C_6^{ij} is the dispersion coefficient for the atom pair ij, s_6 is a scaling factor that depends only on the adopted DFT method (e.g. s_6 is 1.05 for B3LYP) and $R_{ij,\mathbf{g}}$ is the interatomic distance between atoms i in the reference cell and j in the neighbouring cells at distance $|\mathbf{g}|$. For a set of molecular crystals, a cutoff distance of 25.0 Å was used to truncate the summation over lattice vectors which corresponds to an estimated error of less than 0.02 kJ/mol on computed cohesive energies with respect to larger cutoffs [50, 51].

A damping function is used to avoid near-singularities for small interatomic distances:

$$f_{dump}(R_{ij,\mathbf{g}}) = \frac{1}{1+exp^{-d(R_{ij,\mathbf{g}}/R_{vdw}-)}}$$

where R_{vdw} is the sum of atomic van der Waals radii (i.e. $R_{vdw} = R_{vdw}^i + R_{vdw}^j$) and d determines the steepness of the damping function ($d = 20$). Van der Waals radii and atomic C_6 coefficients were taken from Table 1 of ref. [125]. From the latter, the C_6^{ij} dispersion coefficients are computed by using a geometric mean:

$$C_6^{ij} = \sqrt{C_6^i C_6^j}$$

The input requires to specify all parameters that enter equations above and it looks like (urea molecular crystal):

```
Urea
CRYSTAL
0 0 0
113
5.565  4.684
5
6 0.0000 0.5000 0.3260
8 0.0000 0.5000 0.5953
7 0.1459 0.6459 0.1766
1 0.2575 0.7575 0.2827
1 0.1441 0.6441-0.0380
Optional keywords
END (ENDG)
Basis set input
END
DFT
. . . . . .
```

```
END
. . . . .
GRIMME
1.05 20. 25.
1.05 20. 25.      s6 (scaling factor)  d (steepness)  Rcut (cutoff radius)
4
1  0.14  1.001    Hydrogen Conventional Atomic number , C6 , Rvdw
6  1.75  1.452    Carbon   Conventional Atomic number , C6 , Rvdw
7  1.23  1.397    Nitrogen Conventional Atomic number , C6 , Rvdw
8  0.70  1.342    Oxygen   Conventional Atomic number , C6 ,‘Rvdw
SHRINK
. . . . . .
END
```

GUESDUAL - SCF guess (density matrix from a previous run with a different basis set)

rec		variable	value	meaning
•	∗	NFR		number of modification (NFR ≥ 1) in the atomic basis set given in input
		IC	1	complete SCF calculation
			0	stop before the first hamiltonian matrix diagonalization (to be used in CRYSCOR - see *http://www.cryscor.unito.it* when the dual basis set option is activated)
				insert NFR records - for each shell
•	∗	NAT		formal atomic number of the atom whose basis set is modified
	∗	NSH		sequence number of the reference shell in the atomic basis set, starting from which shell(s) is(are) inserted/deleted
	∗	NU		number of shells inserted/deleted after the reference shell NSH

The keyword is to be inserted in the third (Hamiltonian-SCF) input block. It defines how the basis set given in input differs from the one used to compute the density matrix used as SCF guess. The two basis sets must be marginally different. It can be used to add polarization functions, or diffuse functions (see *http://www.cryscor.unito.it*).

The present calculation and the one used to compute the density matrix SCF guess have same symmetry and number of atoms, but different number of shells in the BS of some atoms. Different geometrical parameters and computational conditions are allowed.

The density matrix P^0 (direct lattice) of the previous run, SCF guess, is read from external unit (copy file fort.9 of the previous run in file fort.20). Density matrix is reorganized: elements corresponding to basis functions removed are removed, elements corresponding to basis functions added are set to 0.

Warning. The efficiency of this guess can be dramatically compromised if not congruent BS modifications, i.e. insertion or elimination of significantly populated shells, are made. Check the normalization factor.

Example. LiH bulk. The BS of Li in the reference calculation is referred to as BS1 and reported on the left, the modified one is referred to as BS2 and reported on the right. The GUESDUAL option following BS2 illustrates how BS2 differs with respect to BS1.

```
 BS1                        |      BS2

3   3                      |     3   4
```

```
0   0   6   2.0   1.0      |     0   0   6   2.0   1.0
 700.    .001421            |      700.    .001421
 220.    .003973            |      220.    .003973
  70.    .01639             |       70.    .01639
  20.    .089954            |       20.    .089954
   5.    .31565             |        5.    .31565
   1.5   .4946              |        1.5   .4946
0   0   1   0.0   1.0      |     0   0   1   0.0   1.0
   0.5   1.                 |        0.5   1.
0   2   1   0.0   1.0      |     0   2   1   0.0   1.0
   0.6   1.                 |        0.6   1.
1 4                         |     0 3 1 0. 1.
0 0 5   2.    1.            |        0.4 1.
 120.     .000267           |     1 5
  40.     .002249           |     0 0 5   2.    1.
  12.     .006389           |      120.     .000267
   4.     .032906           |       40.     .002249
   1.2    .095512           |       12.     .006389
0 0 1   0.    1.            |        4.     .032906
   0.45  1.                 |        1.2    .095512
0 0 1   0.    1.            |     0 0 1   0.    1.
   0.13  1.                 |        0.45  1.
0 2 1   0.    1.            |     0 0 1   0.    1.
   0.3   1.                 |        0.13  1.
99   0                      |     0 2 1   0.    1.
END                         |        0.6   1.
                            |     0 2 1   0.    1.
                            |        0.3   1.
                            |     99   0
                            |     END
                                  GUESDUAL
                                  2 0
                                  3 3 1
                                  1 3 1
                                  END
```

In this case two modifications (NFR=2) has been introduced in BS2 with respect to BS1. Complete SCF calculation (IC=1) is required.
First modification in Li (formal atomic number NAT=3) basis set, after the third shell (NSH=3), one shell was added (NU=1).
Second modification in H (atomic number IA=1) basis set, after the third shell (NSH=3), one shell (NU=1) was inserted.

GUESSP - SCF guess - Density matrix from a previous run

The density matrix from a previous run, P^0 (direct lattice), is read from disk, and used as SCF guess. No input data are required.

The density matrix can be edited to modify the spin state. See **SPINEDIT**, page 123.

The density matrix used as SCF guess was written with wave function information in file fort.9 at the end of a previous SCF run, and read as file fort.20.
The two cases, the present one and that used for the restart, must have same symmetry, and 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 density matrix

Program	inp. block	section	comments
crystal	0	2	Title
	1	2.1	geometry input
	2	2.2	basis set input
	3	2.3	computational parameters and SCF input
save wf in file fort.9 (binary) or file.98 (formatted)			

2. Second run - the density 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	2	Title
	1	2.1	geometry input
	2	2.2	basis set input
	3	2.3	computational parameters and SCF input (**GUESSP**)

Warning The modification of the geometry may result in a different order in the storage of the 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 104).

Warning When wave function information are stored formatted in file fort.98, the data must be converted to binary by the keyword **RDFMWF**, page 324 of the ***properties*** program).
Warning SCF guess from a density matrix computed with a different basis set is allowed by the keyword **GUESDUAL**, page 108.

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 20). 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 20).

GUESSYMP - Restart from a previous density matrix with higher or equal symmetry

This keyword allows for the use of a density matrix obtained from a previous run as initial guess for the SCF of a second calculation in which the symmetry of the system has been lowered. The symmetry-group of the second calculation must be a sub-group of the former one.
Both the fort.9 and fort.13 units (symmetry-reducible density matrix; see keyword

SAVEPRED to generate it, page 119) from the first calculation are needed.
The two calculations must have the same number of atoms, basis functions and shells.
Slightly different geometrical parameters, computational conditions or exponents of the Gaussian primitive functions are allowed.
This option can be very useful when the inversion of one or more atomic spins (see the **SPINEDIT** keyword at page 123) leads to a symmetry reduction, as happens in the study of ferro-antiferromagnetic coupling constants or in defective structures where dangling bonds are present and different spin states are thus available, for instance.

HIRSHCHG/HIRSHBLK - Iterative Hirshfeld Population Analysis

The Hirshfeld-I analysis of atomic charge/electron population analysis is performed at the end of the SCF calculation. There are two possible strategies for the input structure. A black-block form, ruled by the **HIRSHCHG** keyword, after which no input data are required and the default conditions concerning spatial integration and convergence are considered for computation of the atomic populations. And a second form in which a block, introduced by the keyword **HIRSHBLK** and completed by the keyword **END**, does contain several particular settings for the calculation of the Hirshfeld-I populations. In particular, the charge integration grid and the parameters of the iterative procedure can be redefined by the keywords described below.

Grids for charge integration

Integration in Hirshfeld-I exploits approaches similar to those considered in the DFT integration scheme (see page 137). The predefined options, that can be eventually used, are the followings:

SGRID: Small grid (default)

This grid correspond to a relatively cheap one that is in general enough for the evaluation of Hirshfeld-I electron population. It is equivalent to **OLDGRID** (see page 140). Other more accurate integration options are:

LGRID: See page 140

XLGRID: See page 140

XXLGRID: See page 140

General keywords to control the iterative procedure

ITCONV

	rec	variable	value	meaning
•	∗	TOLHIRSH	NUMERR	Convergence tolerance

The iterative process stops when the RMS of the differences in atomic population between consecutive cycles is less than a given tolerance. By default, such a value (NUMERR) is actually the numerical error in the estimation of the number of electrons employing the chosen integration grid. As setting **ITCONV** a given real number, the iterative procedure stops when the most restrictive of both criteria is satisfied.

MXNUMITER

	rec	variable	value	meaning
•	∗	NHIRSHIT	200	Maximum number of allowed iterations

If none of the convergence criteria is satisfied, the iterative procedure stops at a NHIRSHIT number of steps.

DETLDINFO

Print a step-by-step detailed information of the iterative procedure. Recommended just for developers.

UNPRUNE

Change the integration grid from pruned to unpruned. This option sometimes provides a numerical integration better than the default but require substantially larger computational time and memory resources. Recommended just for developers.

ILASIZE - set the new size for array ILA used in the calculation of Coulomb 2-electron integrals

	rec	variable	meaning
•	∗	ISIZE	size of array ILA

ILA is an array containing a list of contributions to be computed in the evaluation of the Coulomb series. Use of this option is recommended upon receiving the following error message: "ILA DIMENSION EXCEEDED - INCREASE ILASIZE". Default value of ISIZE is: 6000.

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 [143], if *s* and *sp* shells are involved, and by routines derived from ATMOL [190] 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 [305, 306] (page 124)are used as basis for the Fock/KS 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/KS 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 **SETPRINT**, page 66) 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 [132, 52], 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 105). 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.

Warning - Hamiltonian eigenvalues are modified by the level shifter. Take into account when looking at top and bottom bands eigenvalues printed.

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

LOWMEM - Reduce memory storage requirements

When this option is active, none of the largest arrays is allocated and quantities are recalculated when needed. This option is very efficient in decreasing random access memory requirements for calculations with large unit cells. Usage of **LOWMEM** may affect running time. It is default in *MPPcrystal* calculations.

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".

MADELIND - parameter for Madelung series

	rec	variable	meaning
•	∗	IIND	mql inf(103)

Maximum number of indices for reciprocal lattice vectors to be considered in reciprocal lattice Madelung sums. Default value of IIND: 50.

MEMOPRT - Synthetic report about dynamic memory usage

Memory usage is printed at every step of a CRYSTAL calculation, that is, every time running-time information is also printed. Printed information includes: the total memory allocated by core 0 at that stage of the calculation and the maximum amount of memory required within that step, as well as the total amount of memory allocated by all cores.

MEMOPRT2 - Detailed report about dynamic memory usage

Memory status is printed at every allocation or deallocation of an array. Printed information includes: size of the allocated/deallocated array/matrix and the total amount of memory used by core 0 up to that stage of the calculation.

MK1/MK2/MK2ITER - perturbative dual basis set treatment

Activates the perturbative dual basis set approach. It can be either at First-order (MK1), Second-order (MK2), or iterative second-order (MK2ITER) In all cases the syntax is identical, the user must specify the name label of the Large basis set by using the same syntax as in "Basis set input from external file" 24. The Large basis set must be specified in the external file BASISSETS.DAT.

	rec	variable	meaning
•	*A*	NAME	name tag of the large basis set

Although not mandatory, it is strongly suggested that the small basis set is exactly contained in the large one. That is, the large basis set must be built by adding functions to the small one.

The perturbatively-corrected wave function will be also computed and written on the $fort.9$ unit in place of the small-basis one. For more details on the theory, the user can refer to Ref. [185]

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.

MP2 - Electron correlation energy at second order Møller-Plesset

	rec	variable	meaning
•	A	MP2	Begin MP2 input block
•	A	KNET	reconstruction of the HF Density Matrix
•	I	8	Shrinking factor
•	A	MEMORY	Memory required
•	I	3000	Value in Mbytes
•	A	DFITTING	Density Fitting input block
•	A	PG-VTZ	Density Fitting auxiliarly basis set
•	A	ENDDF	End Density Fitting input block
•	A	DOMMOL	Definition of the excitation domains
•	A	ENDMP2	End MP2 input block

The public CRYSCOR code, a **post-HF local**-correlation program for periodic crystals,[223, 226] has been fully incorporated into the CRYSTAL14 package. This means that it is now possible to perform, on top of and HF-SCF run, a calculation of the electron correlation energy contribution for systems periodic in 1 to 3 dimensions. A simple input example is here reported: for more details about how to run a CRYSCOR calculation please refer to *Cryscor User's Manual*[92] and to CRYSCOR web page: www.cryscor.unito.it.

The post-HF method currently implemented is a perturbative method, namely Møller-Plesset at the second-order (**MP2**). Well localized Wannier Functions (WF) [309] are adopted instead of delocalized Bloch functions for the description of the occupied manifold; this permits the exploitation of the **short-range** nature ($E \propto r^{-6}$) of electron correlation following the general Pulay scheme[235]. An MP2 correction to the HF Density Matrix is feasible so that many of the PROPERTIES can be currently evaluated taking into account the effects of the dynamic correlation.

The code has been widely used to study, among others, the structure and stability of different polymorph,[44, 94] the energetic balance in van der Waals and hydrogen-bonded crystals,[184] interactions between molecules and surfaces,[179] and the upshot of electron correlation on some fundamental properties such as vibrational spectra,[87] electron density and Compton profiles.[96]

Note that the localization procedure and the MP2 technique implemented in CRYSCOR are not suitable for open-shell systems, for conductors, or for semiconductors with very small gap.

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 399). 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 399) 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 59

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 2000000.
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 2000000.
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 2000000. 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.

NOLOWMEM - Disable reduction of memory storage requirements

LOWMEM option (see page 114) is disabled. CRYSTAL calculations are speeded up but the amount of requested memory increases. It is default in Pcrystal calculations.

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 Hamiltonian matrix diagonalization. No input data are required. This choice increases the diagonalization CPU time when the system has symmetry operators.

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 18.3, page 397.

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.

PROJDUAL - dual basis set correction on energy

Activates the dual basis set approach for estimating the energy correction due to a basis set enlargement.
The approach is based on Martin Head-Gordon and coworkers' periodic extension of the method. [169]

	rec	variable	meaning
•	A	NAME	name tag of the large basis set

Differently from the MKn approach (see page 114) this approach only delivers a correction to the total energy.

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 59, 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 D, page 445.

PRINTOUT - Setting of printing environment

See input block 1, page 64.

QVRSGDIM - limiting 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.

NFILE limits the maximum amount of data to be stored to memory. This is a way to reduce memory storage requirements with some possible reduction of performance.

REPLDATA - Replicated-data mode

This option is supported by *MPPcrystal* only. It allows the User to run *MPPcrystal* as *Pcrystal*. All computational parameters and settings are changed accordingly.

RHF [default]

A restricted closed-shell hamiltonian calculation is performed ([241, 224], Chapter 8 of ref. [220]). Default choice.

ROHF

	rec	variable	meaning
•	∗	NSPIN	n_α-n_β electrons (Section 3.3, page 124)

In the Restricted Open Shell[241] Hamiltonian, the variational space is split into three subspaces: doubly and singly occupied, and unoccupied. In order to correctly describe the two occupied subspaces, the difference between the amount of α and β electrons (NSPIN) must be locked at all **k** points; this difference is kept until the end of the SCF. There are three main differences with respect to UHF: (1) only alpha electrons are allowed to be assigned to singly occupied states, (2) the imposed occupancy constraint allows the ROHF approach to provide solutions that are eigenfunctions of the spin operator, $\widehat{S^2}$, and (3) solutions are in general less stable than those obtained with the UHF method. It is important to underline that ROHF solutions never exhibit locally negative spin densities. In CRYSTAL the symmetry adaption of the crystal (molecular) orbitals is disabled for the ROHF method.

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. **This option is activated by default from** CRYSTAL14. It can be deactivated by inserting the **NODIRECT** keyword.

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.
To restart SCF cycles using the density matrix written in file fort.79, it has to be copied in file fort.20
No input data required.

SAVEPRED

The symmetry-reducible density matrix is written in unit fort.13 at the end of the SCF.
No input data required.

SETINF - Setting of INF values

See input block 1, page 66

SETPRINT - Setting of printing options

See input block 1, page 66.

SHRINK - Pack-Monkhorst/Gilat shrinking factors

rec	variable	value	meaning
			if the system is periodic insert II
•	∗ IS		Shrinking factor in reciprocal space (Section 18.7, page 401)
	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:

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

$$H_k A_k = S k A_k E_k$$

A second shrinking factor, ISP, defines the sampling of k points, "Gilat net" [120, 119], 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

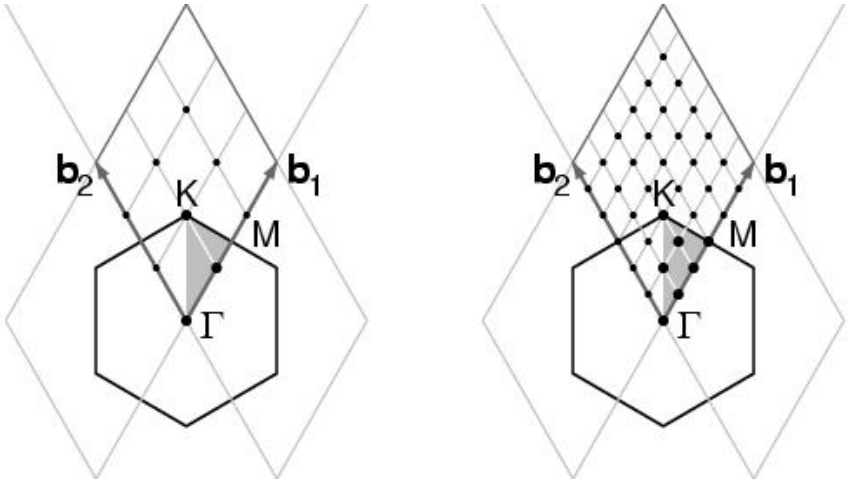
where b1, b2, b3 are the reciprocal lattice vectors, and is1, is2, is3 are integers "shrinking factors".
In 2D crystals, IS3 is set equal to 1; in 1D crystals both IS2 and IS3 are set equal to 1. Only points k_i of the Pack-Monkhorst net belonging to the irreducible part of the Brillouin Zone

(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 18.7, page 401. 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 [224], 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 117.

Some tools for accelerating convergence are given through the keywords **LEVSHIFT** (page 113 and tests 29, 30, 31, 32, 38), **FMIXING** (page 105), **SMEAR** (page 121), **BROYDEN** (page 95) and **ANDERSON** (page 93).
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 125;
TOLDEP (tolerance on SQM in density matrix elements) page **??**;
MAXCYCLE (maximum number of cycles) page 114.

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 18.9, page 396) to the Fermi function;

$$f_j = (1 + e^{\frac{(\epsilon_j - \epsilon_F)}{k_b T}})^{-1} \tag{3.35}$$

where ϵ_F is the Fermi energy and k_bT 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 [187]. In this case the "free energy" of the system may be computed as:

$$\begin{aligned} F &= E(T) - TS(T) \\ &= E + k_b T \sum_{i}^{N_{states}} f_i \ln f_i + (1 - f_i) \ln(1 - f_i) \end{aligned} \tag{3.36}$$

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:

$$\begin{aligned} E(T) &= E(0) + \alpha T^2 \\ S(T) &= 2\alpha T \end{aligned} \tag{3.37}$$

and so the quantity

$$E0 = \frac{F(T) + E(T)}{2} = E(0) + O(T^3) \tag{3.38}$$

may be used as an estimate of E(0).

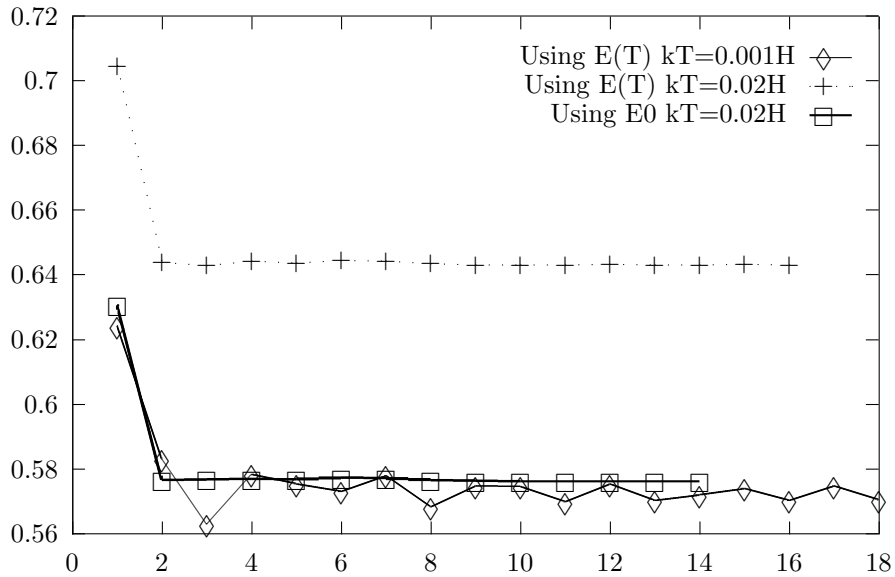


Figure 3.5: The surface energy (J/M^2) of Li(100) for various numbers of layers in a slab model showing the effects of WIDTH (0.02H and 0.001H) and the use of E(T) or E0

Figure 3.5 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.

Note that for conducting systems analytic first derivatives are not fully implemented when the keyword **SMEAR** is used. In that case, *numerical* first derivatives should be computed (see

page 191). For very small value of smearing (around 0.001 hartree) analytical gradients should be reliable in geometry optimization.

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 126) or **DFT/SPIN** (input block 3, page 135) 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 nonequivalent. 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	↑ ↑ ↑ ↑

Warning To lock the difference between α and β electrons α and β eigenvalues are forced to be split. Their printed value is meaningless, until locking is active.

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

Note: if NCYC is < 0, then the lock duration is controlled by energy difference between successive cycles (disappears when $< 10^{-NCYC}$) instead of by number of cycles. This is useful when the locking must be relaxed before the convergence of the SCF -otherwise there is a risk to arrive to a spurious solution- but at the same time one does not know exactly in which SCF cycle the density matrix can be considered to be close enough to the right electronic structure so as to avoid a divergent behavior after the locking finishes.
Situations like this may occur in geometry optimizations of ferromagnetic systems, for instance: the SCF of the first optimization point converges in CYC 20 with the locking finishing in CYC 12; in the remaining points, as the density matrix of the previous one is used as initial guess, the SCF converges in CYC 10 under locking and, so, the true convergence is not ensured.

Using FINALRUN=4 this situation could lead to an extremely large or even non convergent optimization process.
On the other hand, to fix locking since the energy difference between cycles is less than a given threshold is a quite handy criterion that could be in several cases preferable than fixing the duration in terms of number of cycles.

SPINLOC2 - Spin-polarized solutions

rec	variable	meaning
• ∗	SPIN	n_α-n_β electrons
∗	NCYC	number of cycles the difference is maintained or energy thereshold

The keyword functions similarly to SPINLOCK, but instead of shifting the eigenvalues in order to lock the system in a specific spin state, fills separately the α and β bands with the appropriate number of electrons. This allows the system to pass through or converge to a conducting state while the locking is active. Additionally, it is possible to set the value of SPIN to a real number.
Using this keyword the energies for the top of virtual and bottom of valence band , as well as the band gap, are printed for both α and β electrons. These values are not meaningless as in the case of SPINLOCK but should be interpreted with care, since if the spin state is unstable there may be α occupied states at higher energy with respect to β virtual states or vice versa.

STDIAG - Force standard diagonalization method

This option is supported by *MPPcrystal* only. Matrix diagonalization method in reciprocal space is switched from "Divide and Conquer method" (default, more efficient) to standard.

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 [305, 306].
The Symmetry Adapted Functions are used in the Hamiltonian 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 59).

TEST[RUN] - Integrals classification and selection

Full input (geometry, basis set, general information, SCF) is processed.
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 reduces the required space to about 1/3 of the printed value.

No input data 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

Different default values are set for different type of calculation:

SCF single point		6
Geometry optimization	OPTGEOM	7
Frequency calculation	FREQCALC	10
Elastic constants	ELASTCON	8
Equation of state	EOS	8

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

	rec	variable	meaning
•	∗	ITOL1	overlap threshold for Coulomb integrals- page 397 7
		ITOL2	penetration threshold for Coulomb integrals-page 398 7
		ITOL3	overlap threshold for HF exchange integrals-page 398 7
		ITOL4	pseudo-overlap (HF exchange series-page 398) 7
		ITOL5	pseudo-overlap (HF exchange series-page 398) 14

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 18.

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 $\varphi^0_\mu\ \varphi^g_\nu$ (page 395) 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).

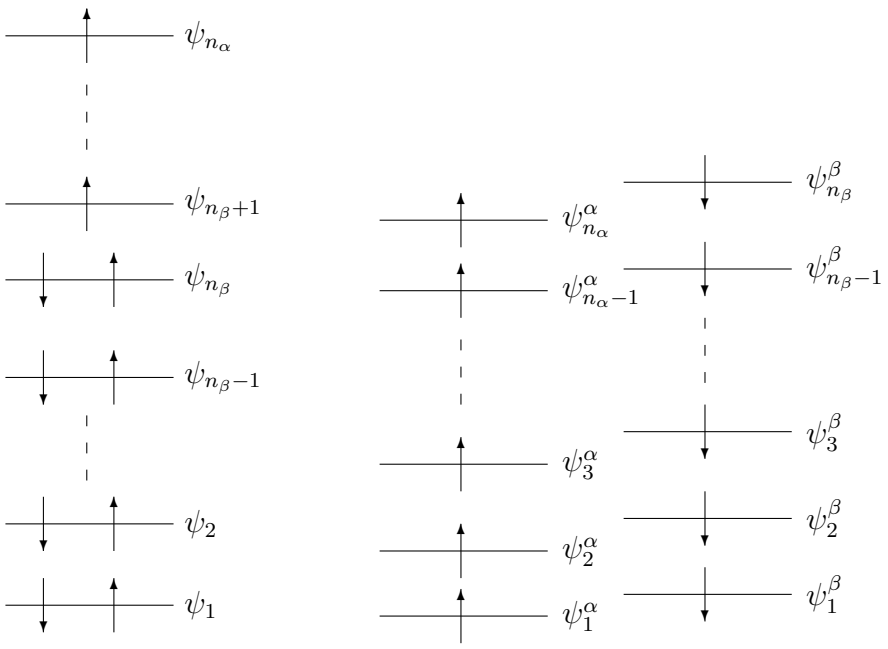


Figure 3.6: Molecular Orbitals diagram for the Restricted Open Shell method (ROHF, left) and for the Unrestricted Open Shell method (UHF, right)

UHF - 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) ([224, 8], Chapter 6 of ref. [220]).

In the Restricted Open Shell [241] 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.

Another mono-determinantal approach for the study of open-shell systems is the UHF method [230]. 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. 3.6.

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.

Chapter 4

Density Functional Methods

The Kohn-Sham [159, 147] 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 functional

1 Dispersion correction to DFT (DFT-D)

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

3 *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.

4.1 Choice of the exchange-correlation functional

Many different approximate exchange-correlation functionals are available in CRYSTAL: from semilocal (i.e. LDA, GGA and mGGA) to global and range-separated hybrid 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.
A list of the exchange and correlation functionals that can be select in input is reported below.

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

Insert one of the following keywords II

LDA functionals	
LDA	LSD. Dirac-Slater [68]
VBH	LSD. von Barth-Hedin [281]
GGA functionals	
BECKE	GGA. Becke 1988 functional [18]
mPW91	GGA. modified Perdew-Wang 1991 functional [**?**]
PBE	GGA. Perdew-Becke-Ernzerhof [204]
PBESOL	GGA. PBE functional revised for solids [206]
PWGGA	GGA. Perdew-Wang [205]
SOGGA	GGA. Second Order corrected GGA functional. It must be used in combination with the PBE correlation functional[302]
WCGGA	GGA - Wu-Cohen 2006 functional [293]

CORRELAT Correlation Potential (default: no correlation).

Insert one of the following keywords II

LDA functionals	
PWLSD	LSD. Perdew-Wang parameterization of the Ceperley-Alder free electron gas correlation results [207]
PZ	LSD. Perdew-Zunger parameterization of the Ceperley-Alder free electron gas correlation results [210]
VBH	LSD. von Barth-Hedin [281]
VWN	LSD. Vosko-Wilk-Nusair parameterization of the Ceperley-Alder free electron gas correlation results. Also known as VWN5. [282]
GGA functionals	
LYP	GGA. Lee-Yang-Parr [167]
P86	GGA. Perdew 86 [202]
PBE	GGA. Perdew-Burke-Ernzerhof [204]
PBESOL	GGA. PBE functional revised for solids [206]
PWGGA	GGA. Perdew-Wang [205]
WL	GGA - Wilson-Levy [291]
meta-GGA functionals	
B95	two parameters meta-GGA correlation functional - Becke [22]

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

exchange	correlation	
—	PWGGA	Hartree-Fock exchange, GGA Perdew-Wang correlation.
LDA	VWN	probably the most popular LDA formulation, also known as S-VWN
VBH	VBH	was the most popular LDA scheme in the early LDA solid state applications (1975-1985).
PBE	PBE	the well-known PBE XC functional
BECKE	LYP	B-LYP

Standalone keywords for common exchange-correlation functionals are also available.

SVWN	Combination of Slater for exchange and VWN for correlation [68, 282]
BLYP	Combination of B88 for exchange and LYP for correlation [18, 167]
PBEXC	Full PBE XC GGA functional [204]
PBESOLXC	XC GGA revised PBE functional for solids [206]
SOGGAXC	Combination of SOGGA for exchange and PBE for correlation [302]
SOGGA11	20 parameters GGA functional based on SOGGA by Peverati, Zhao and Truhlar [216]

Global Hybrid functionals

Global hybrid (GH) functionals have the general formula:

$$E_{xc}^{GH} = (1 - A) * E_x^{DFA} + A * E_x^{HF} + E_c^{DFA}$$

They include a constant fraction of HF exchange (A).

Standalone keywords are available to define global hybrid functionals completely.

B3PW	Becke's 3 parameter functional [19] combined with the non-local correlation PWGGA [203, 208, 209, 207]
B3LYP	Becke's 3 parameter functional [19] combined with the non-local correlation LYP. 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 [282]) it is reported as functional V, which is used to extract the 'local' part of the LYP correlation potential.
PBE0	Hybrid version of the PBE XC functional with 25% (1/4) of HF exchange. Adamo and Barone [5]
PBESOL0	Same as PBE0 but with the PBEsol XC functional instead of PBE
B1WC	One-parameter hybrid functional which combines WC exchange functional with 16% of HF exchange and the PWGGA correlation functional (see [25])
WC1LYP	As for the B1WC functional, but with the LYP correlation functional instead of the PWGGA (see [60])
B97H	Hybrid functional based on the B97 XC functional [4, 101]
PBE0-13	Revised version of the PBE0 functional with 33.3333% (1/3) of HF exchange [42]
SOGGA11X	21 parameters Hybrid version of the SOGGA11 XC functional with 40.15% of HF exchange by Peverati and Truhlar [215]
mPW1PW91	Modified version of PW91 exchange functional combined with original PW GGA correlation with 25% of HF exchange by Adamo and Barone [6]
mPW1K	Modified version of PW91 exchange functional combined with original PW GGA correlation with 42.80% of HF exchange [175]

User-defined global hybrid functionals

It is possible to define other global hybrid functionals by specifying an exchange functional (see the **EXCHANGE** keyword) and a correlation one (see the **CORRELAT** keyword) and then include a given amount of exact Hartree-Fock exchange into the exchange functional through the keyword **HYBRID**.

HYBRID	Hybrid method - 1 record follows:
• ∗A	Fock exchange percentage (default 100.)

Any mixing (0-100) of exact Hartree-Fock and DFT exchange can be used.

The keyword **HYBRID** can also be used to modifiy the fraction of HF exchange in existing GH functionals.

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

NONLOCAL allows users to modify the relative weight of the local and non-local part both in the exchange and the correlation potential with respect to standard definition of Becke's 3 parameter functional types:

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

A, B, and C are the input data of **HYBRYD** and **NONLOCAL**. Becke's 3 parameter functionals currently implemented in CRYSTAL are B3PW and B3LYP.

For example, the following sequences correspond to the stand-alone keywords for some of the available GH functionals.

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

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

PBE0
corresponds to the sequence:
EXCHANGE
PBE
CORRELAT
PBE
HYBRID
25

PBESOL0
corresponds to the sequence:
EXCHANGE
PBESOL
CORRELAT
PBESOL
HYBRID
25

B1WC
corresponds to the sequence:
EXCHANGE
WCGGA
CORRELAT
PWGGA
HYBRID
16

WC1LYP
corresponds to the sequence:
EXCHANGE
WCGGA
CORRELAT
LYP
HYBRID
16

Range-Separated Hybrid functionals

CRYSTAL offers a wide variety of range-separated hybrid (RSH) functionals in which the amount of HF exchange included depends on the distance between electrons
They are obtained from the separation of the Coulomb operator in different ranges (three

ranges in the current implementation) by means of the *error function* as:

$$\frac{1}{r_{12}} = \underbrace{\frac{\mathrm{erfc}(\omega_{SR} r_{12})}{r_{12}}}_{SR} + \underbrace{\frac{1 - \mathrm{erfc}(\omega_{SR} r_{12}) - \mathrm{erf}(\omega_{LR} r_{12})}{r_{12}}}_{MR} + \underbrace{\frac{\mathrm{erf}(\omega_{LR} r_{12})}{r_{12}}}_{LR}$$

where ω is the length scale of separation. Then, the general form of a range-separated hybrid is:

$$E_{xc}^{RSH} = E_{xc}^{DFA} + c_{SR}(E_{x,SR}^{HF} - E_{x,SR}^{DFA}) + c_{MR}(E_{x,MR}^{HF} - E_{x,MR}^{DFA}) + c_{LR}(E_{x,LR}^{HF} - E_{x,LR}^{DFA})$$

According to the values of c_{SR}, c_{MR}, c_{LR}, ω_{SR} and ω_{LR}, short-, middle- and long-range corrected RSH functionals can be defined.

The following RSH functionals are available:

Short-range Corrected (SC) functionals	
HSE06	Screened-Coulomb PBE functional [2, 204] combined with PBE correlation
HSEsol	Screened-Coulomb PBEsol functional [162, 206] combined with PBEsol correlation
SC-BLYP	Short-range corrected RSH functional based on Iirikura-Tsuneda-Yanai-Hirao scheme[152, 271] of the BLYP XC functional
Middle-range Corrected (MC) functionals	
HISS	Middle-range corrected functional based on the PBE exchange hole [267, 268]) combined with PBE correlation (labeled B in ref. [267])
Long-range Corrected (LC) functionals	
RSHXLDA	Long-range corrected LDA exchange functional [1, 151] combined with VWN for correlation
LC-wPBE	Long-range corrected RSH functional[85] based on PBE XC functional
LC-wPBEsol	Long-range corrected RSH functional[85] based on PBEsol XC functional
LC-wBLYP	Long-range corrected RSH functional[85] based on BLYP XC functional
wB97	Chai-Head-Gordon long-range corrected RSH functional[153, 4]
wB97X	Chai-Head-Gordon long-range corrected RSH functional[153, 4] with a small contribution of HF exchange at short-range
LC-BLYP	Long-range corrected RSH functional based on Iirikura-Tsuneda-Yanai-Hirao scheme [152, 271] of the BLYP XC functional
CAM-B3LYP	Long-range corrected RSH functional (Coulomb-Attenuating method)[270] based on the BLYP XC functional
LC-PBE	Long-range corrected RSH functional based on Iirikura-Tsuneda-Yanai-Hirao scheme [152, 164] of the PBE XC functional

Details of the amount of RS-HF exchange and the length scale separation ω for the RSH functionals available in CRYSTAL are given in the table below:

Method	Exchange	c_{SR}	c_{MR}	c_{LR}	$\omega_{SR}(a_0^{-1})$	$\omega_{LR}(a_0^{-1})$	Correlation	Ref.
HSE06	PBE	0.25	0.00	0.00	0.11	0.11	PBE	[2, 204]
HSEsol	PBEsol	0.25	0.00	0.00	0.11	0.11	PBEsol	[162, 206]
SC-BLYP	B88	0.20	0.00	0.00	0.11	0.11	LYP	[152, 271, 18, 167]
HISS	PBE	0.00	0.60	0.00	0.84	0.20	PBE	[267, 268, 204]
LC-ωPBE	PBE	0.00	0.00	1.00	0.40	0.40	PBE	[85, 204]
LC-ωPBEsol	PBEsol	0.00	0.00	1.00	0.60	0.60	PBEsol	[85, 206]
LC-ωBLYP	B88	0.00	0.00	1.00	0.60	0.60	LYP	[85, 18, 167]
RSHXLDA	S	0.00	0.00	1.00	0.40	0.40	VWN	[1, 151, 282]
ωB97	B97	0.00	0.00	1.00	0.40	0.40	B97	[153, 4]
ωB97-X	B97	0.157706	0.00	1.00	0.30	0.30	B97	[153, 4]
LC-BLYP	B88	0.00	0.00	1.00	0.47	0.47	LYP	[152, 271, 18, 167]
CAM-B3LYP	B88	0.157706	0.00	1.00	0.33	0.33	LYP	[270, 18, 167]
LC-PBE	PBE	0.00	0.00	1.00	0.47	0.47	PBE	[152, 164, 204]

Notes for RSH functionals:

(i) the bipolar expansion is not active in the calculation of the exchange integrals (on average, the cost can increase by a factor 2 to 3);

(ii) one- and two-electron repulsion integrals are computed in direct mode;

(iii) implementation of HSE06, HISS, LC-wPBE and related RSH methods is based on the Henderson-Janesko-Scuseria model of the PBE exchange hole[269].

For available RSH functionals, the value(s) of the length scale separation ω can be modified in input by using the following keywords.

SR-OMEGA For screened-Coulomb RSH methods - 1 record follows:
- $*\,\omega_{SR}$ value of the length scale separation ω at short-range (default 0. a_0^{-1})

MR-OMEGA For middle-range corrected RSH methods - 1 record follows:
- $*\,\omega_{SR}$ value of the length scale separation ω at short-range (default 0. a_0^{-1})

  ω_{LR} value of the length scale separation ω at long-range (default 0. a_0^{-1})

LR-OMEGA For long-range corrected RSH methods - 1 record follows:
- $*\,\omega_{LR}$ value of the length scale separation ω at long-range (default 0. a_0^{-1})

SR-HYB_WB97X Option to change the amount of SR-HF exchange in the ωB97-X functional - 1 record follows:
- $*\,c_{SR}$ value of the coefficient for the SR-HF exchange term

The amount of HF exchange (i.e. c_{SR}, c_{MR}, c_{LR}) can be modified by specifying the **HYBRYD** keyword (see above for details).

User-controllable RSH x-functional based on the PBE functional

LSRSH-PBE User-controllable RSH x-functional based on the PBE functional - 1 record follows:
- $*\,\omega$ value of the length scale separation ω (default 0. a_0^{-1})

  c_{SR} value of the coefficient for the SR-HF exchange term

  c_{LR} value of the coefficient for the LR-HF exchange term

The keyword LSRSH-PBE allows users to specify the value of omega, c_{SR} and c_{LR} for a generic RSH functional based on the PBE functional. So, the input looks like:

```
...
DFT
```

```
LSRSH-PBE
0.11 0.25 0.00001
END
...
```

For instance, this corresponds to the HSE06 functional.
According to the general form of the RSH functional, one can also obtain the LC-ωPBE functional as: ω=0.4 a_0^{-1}; c_{SR}=0.00001; c_{LR}=1.00.
Note that c_{SR} and c_{LR} can be small but not zero (threshold: 1E-06).

meta-GGA functionals

Both pures and hybrids meta-GGA (mGGA) functionals are available. Some of them rely on Becke95 mGGA coorelation, while others belong to the Minnesota family of functionals as proposed by Truhlar and co-workers. SCAN by Perdew and and co-workers, r^{2}SCAN by Furness, Sun and co-workers and their hybrid variants are also implemented.

Pure mGGA functionals

M06L	pure mGGA version of the M06 hybrid functional[296]
revM06L	revised version of the M06L pure mGGA functional[284]
MN15L	58 parameters pure mGGA by Yu, He, and Truhlar [301]
SCAN	non-empirical mGGA functional based on SCAN XC [264]
r^{2}SCAN	modified version of regularized SCAN [114]

Global hybrid mGGA functionals

B1B95	Becke88 GGA exchange combined with Becke95 mGGA correlation and 28% of HF exchange[23]
mPW1B95	mPW91 GGA exchange combined with Becke95 mGGA correlation and 31% of HF exchange[303]
mPW1B1K	mPW91 GGA exchange combined with Becke95 mGGA correlation and 44% of HF exchange[303]
PW6B95	6 parameters mPW91 GGA exchange combined with re-optimized Becke95 mGGA correlation and 28% of HF exchange[304]
PWB6K	6 parameters mPW91 GGA exchange combined with re-optimized Becke95 mGGA correlation and 46% of HF exchange[304]
M05	Minnesota 2005 global hybrid mGGA XC functional hybridized with 28% of HF exchange[300]
M052X	as M05 but with a doubled amount of HF exchange[299]
M06	Minnesota 2006 global hybrid mGGA XC functional hybridized with 27% of HF exchange[298]
M062X	as M06 but with twice amount of HF exchange[298]
M06HF	global hybrid mGGA with 100% HF exchange[297]
MN15	59 parameters global hybrid mGGA with 44% of HF exchange[138]
revM06	revised version of global hybrid mGGA M06 with 40.41% of HF exchange[285]
SCAN0	global hybrid mGGA based on SCAN XC with 25% of HF exchange[149]
r^{2}SCANh	global hybrid mGGA based on r^{2}SCAN XC with 10% of HF exchange[40]
r^{2}SCAN0	global hybrid mGGA based on r^{2}SCAN XC with 25% of HF exchange[40]
r^{2}SCAN50	global hybrid mGGA based on r^{2}SCAN XC with 50% of HF exchange[40]

Availability of XC functionals

Not all functionals are available for all types of calculations and computed properties. A summary is given in the following table (Y=Yes, N=Not)

DFA	Energy	Gradients	CPKS
LDA	Y	Y	Y
GGA	Y	Y	Y
mGGA	Y	Y	N
Global hybrids (LDA,GGA)	Y	Y	Y
Global hybrids (mGGA)	Y	Y	N
Range-separated hybrids	Y	Y	Y

Note that only a subset of the pure and hybrid LDA/GGA XC functionals can be used in the CPKS scheme, namely:

LDA/GGA	Global hybrids	Range-separated hybrids
LDA	B3LYP	RSHXLDA
VWN	PBE0	ωB97
BECKE	PBEsol0	ωB97-X
PBE (XC)	B1WC	LC-BLYP
PBEsol (XC)	WC1LYP	SC-BLYP
SOGGA	B97H	CAM-B3LYP
WCGGA		HSE06
LYP		HSEsol
		HISS
		LC-ωPBE
		LC-ωPBEsol
		LC-ωBLYP
		LC-PBE

The calculation of properties that require the solution of the CPKS equations, such as linear and non-linear electric susceptibilities (e.g. dielectric constant), Raman intensities and photoelasticity, is then limited to that subset of XC functionals.

Spin-polarized systems

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)

Self-consistent Hybrids

For a long time, hybrid functionals (either global or screened-exchange) have been characterized by a fixed, system-independent Fock exchange fraction α (0.2 in B3LYP and B3PW91, 0.25 in PBE0 and HSE06, 0.16 in B1WC, for instance). Many properties of solids turn out to be significantly affected by the α parameter, which makes the identification of its optimal value both conceptually and practically important. For instance, it is known that the electronic band structure of small or large band gap solids is better reproduced by use of smaller or larger values of α, respectively.
In recent years, the use of a system-specific optimal Fock fraction, as linked to the static electronic screening of the system, has been proposed, whose practical prescription implies defining α as inversely proportional to the static electronic dielectric constant ϵ_∞ of the material. The dielectric constants entering the definition of these functionals have been either taken from the experiment or computed with different techniques. In this respect, Skone *et al.* [Phys. Rev.

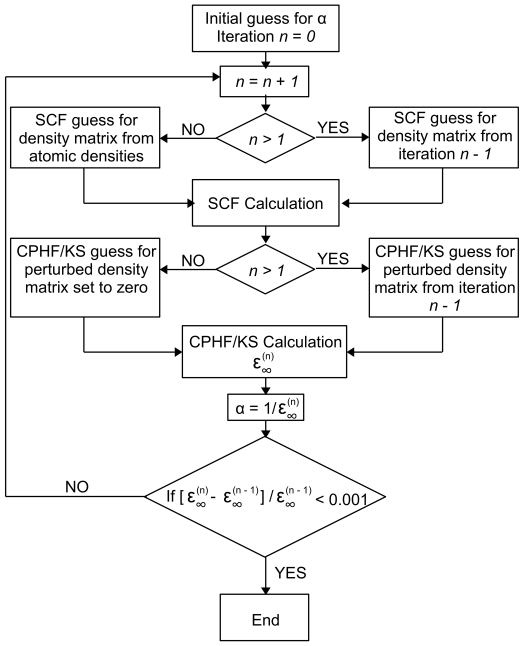


Figure 4.1: Flow chart of the automated algorithm for the system-specific definition of self-consistent hybrid functionals, as implemented in CRYSTAL17.

B **89**, 195112 (2014)] have recently proposed a promising iterative scheme for the calculation of the optimal Fock exchange fraction of global hybrids, where the static electronic dielectric response of the material and α are self-consistently determined. Self-consistent hybrids are becoming popular in solid state applications mainly because of their reliable description of the electronic band structure and some optical properties.
The system-specific character of self-consistent hybrid functionals on the one hand has the advantage of bringing into the exchange-correlation functional the electronic screening proper of the material but, on the other hand, it makes it necessary to determine α for each system as a pre-step to any calculation, which implies a higher computational cost and a reduced user-friendliness compared to standard hybrids, unless computationally-efficient, automated strategies are devised for their definition.
In CRYSTAL17, a fully-automated (i.e. requiring a single calculation), computationally-efficient implementation of self-consistent hybrid functionals has been implemented. The static electronic dielectric tensor of the system is computed by adopting a Coupled-Perturbed-Hartree-Fock/Kohn-Sham (CPHF/KS) approach and full advantage is taken of internal guesses of perturbed and unperturbed density matrices from previous iterations to achieve high computational efficiency.
A schematic representation of the fully-automated algorithm for the determination of the optimal Fock exchange of self-consistent hybrid functionals is reported in Figure 4.1, as implemented into CRYSTAL17. The procedure starts with a guess for the exact exchange fraction α (any value in the range from 0 to 1) and with the choice of the adopted exchange-correlation DFT functional. At each iteration n of the procedure, the static electronic dielectric tensor of the system is computed, from which an average dielectric constant $\epsilon_\infty^{(n)}$ is evaluated. At the end of each iteration n, the Fock exchange fraction is updated according to $\alpha = 1/\epsilon_\infty^{(n)}$ and convergence of the whole process is checked on the average dielectric constant (i.e. convergence is reached when ϵ_∞ changes by less than 0.1% between two subsequent iterations).

The optimal fraction of exact exchange is self-consistently determined by performing a preliminary calculation activated by the keyword **SCHYBRID**, to be inserted at the end of the geometry input block. This keyword opens a block that must be terminated by a keyword **END**.
A possible input reads as follows:

```
...geometry...
SCHYBRID   Keyword to compute the optimal exchange fraction of self-consistent hybrid
END   End of Self-consistent hybrid input
END   End of geometry input block
...basis set...
END   End of basis set input block
DFT    Keyword to set a DFT calculation
PBE0  The PBE0 global hybrid functional is selected, for instance
END   End of DFT input block
END   End of computational parameter input block
```

At the end of the output file, the following information is found:

```
 SCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCH

                    SELF-CONSISTENT HYBRID CALCULATION

                   Implemented by A. Erba on April 2015

          based on the paper by J.H. Skone, M. Govoni and G. Galli
                       Phys. Rev. B, 89, 195112 (2014)

 SCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCHSCH

 CONVERGENCE ON DIELECTRIC CONSTANT REACHED AFTER  4 CYCLES


 AVERAGE DIELECTRIC CONSTANT:  2.7355
   OPTIMAL EXCHANGE FRACTION: 36.5473
```

Users of this module are kindly requested to cite the following reference:
A. Erba, *J. Phys.: Condens. Matter*, **29**, 314001 (2017) *Self-consistent Hybrid Functionals for Solids: A fully-automated Implementation*

The optimal fraction of Fock exchange is given in the last row. At this point, all one needs to do is to use this Fock exchange fraction in the calculation. To do that, the **HYBRID** keyword of the **DFT** block can be used:

```
...geometry...
END   End of geometry input block
...basis set...
END   End of basis set input block
DFT    Keyword to set a DFT calculation
PBE0  The PBE0 global hybrid functional is selected, for instance
HYBRID   Keyword to set the Fock exchange fraction
36.5473
END   End of DFT input block
END   End of computational parameter input block
```

4.2 Dispersion correction to DFT (DFT-D)

The proper description of noncovalent interactions requires the inclusion of long-range electron correlation effects that are missing in both HF and DFT methods. In particular, the treatment of the weak London forces is crucial because of their attractive and ubiquitous nature.

In CRYSTAL, empirical and semi-classical corrections D2/D3 as proposed by S. Grimme[124, 126] are available to take into account dispersive interactions.
For a range of common density functionals, the D3 correction can be invoked by simply specifying the XC functional with the suffix '**-D3**' in the DFT input block by means of stand-alone keywords. For more details, users are referred to Chapther 5.1.

The older D2 corrections can still be used by means of either the option "version 2" of the DFTD3 keyword (see 5.1) or the keyword GRIMME (see 3.3).[50, 51, 276]

4.3 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 [16] for molecular systems and then extended to periodic systems [273]. 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 [17]. Default choice.

SAVIN Savin weights [251]

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 138) [default 13]

Note: A new default grid has been set. It corresponds to the XLGRID in CRYSTAL09

The *default grid* is a pruned (75,974) grid, having 75 radial points and a maximum number of 974 angular points in regions relevant for chemical bonding. Each atomic grid is split into five shells with different angular grids.

This grid guarantees accurate integration of the XC potential 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 also provides accurate results for atoms up to fourth-row and heavier.

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
`75`	Number of radial points in the i-th interval
`ANGULAR`	Keyword to specify the angular grid
`5`	Number of intervals in the angular part
`0.1667 0.5 0.9 3.5 9999.0`	Angular integration limits of the i-th interval
`4 8 12 16 12`	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=      44707

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

 RADIAL INTEGRATION  - INTERVALS (POINTS,UPPER LIMIT):            1( 75,  4.0*R)

 ANGULAR INTEGRATION - INTERVALS (ACCURACY LEVEL [N. POINTS] UPPER LIMIT):
  1(  4[  86]   0.2)  2(  8[ 194]   0.5)  3( 12[ 350]   0.9)  4( 16[ 974]   3.5)
  5( 12[ 350]9999.0)
```

Three more pre-defined grids are available which can be selected to improve accuracy (or reduce the cost) by inputing the following global keywords:

OLDGRID Old default grid

This is the old default grid in CRYSTAL09. It 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

LGRID Large grid

Global keyword to choose a smaller 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. [121].

[**XLGRID**] Extra large grid (default)

XXLGRID Extra extra large grid

The *extra-extra-large grid* is a pruned (99,1454) grid, consisting of 99 radial points and 1454 angular points in the region of chemical interest.
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
```

XXXLGRID Ultra extra extra large grid

The *ultra extra-extra-large grid* is a pruned (150,1454) grid, consisting of 150 radial points and 1454 angular points in the region of chemical interest. This is very large and accurate grid which can be used for benchmark calculations, expecially using meta-GGA functionals.
It corresponds to:

```
RADIAL
1
4.0
150
ANGULAR
5
0.1667 0.5 0.9 3.5 9999.0
6 10 14 18 14
```

HUGEGRID Ultra extra extra large grid for SCAN functional

The *ultra extra-extra-large grid* is a pruned (300,1454) grid, consisting of 300 radial points and 1454 angular points in the region of chemical interest. It should be used in combination with the SCAN functional.
It corresponds to:

```
RADIAL
1
4.0
300
ANGULAR
5
0.1667 0.5 0.9 3.5 9999.0
6 10 14 18 14
```

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 18.12 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]

In CRYSTAL (serial and parallel versions), in the calculation of the exchange-correlation contribution to the Kohn-Sham Hamiltonian matrix elements and energy gradients, a grid of NPOINT points is partitioned into batches of points as suggested by Ahlrichs [274]. The unit cell volume is then divided into a number of equivalent sub-volumes which approximates the following ratio: NPOINT/BATCH. Each grid point is assigned to a sub-volume and the set of points inside a sub-volume forms a batch. Since the distribution of grid points is not uniform in the unit cell (it is an atomic grid) the population of batches may vary from a few points to a few thousands. For this reason BATCH does not correspond to the maximum number of points in a batch. For example, batches of grid points close to the nuclei contain many more points than batches in internuclear regions. This may produce load unbalancing in parallel calculations (see **CHUNKS**).
The number of sub-volumes into which subdivided rhe unit cell is altered by changing BATCH. Reducing BATCH may result in some degree of inefficiency (minimum value: 1). Changing the value of BATCH can also affect results at some extent if TOLDENS is not sufficiently accurate, as the selection of contributions to the total density at a point can be influenced by the definition of the sub-volumes.
Default value of BATCH is 100, as averagely good balance between accuracy and efficiency.

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

This option is available for replicated-data calculations (Pcrystal only). Since different batches can contain a different number of points in a rather wide range, task farming can be unbalanced. **CHUNKS** allows user to set the maximum number of points per batch so improving load balancing. Recommended value of NCHU: 200.
CHUNKS is not implemented under **DISTGRID**.

DISTGRID

This option is available for both replicated-data and massive-parallel calculations. It forces the code to distribute the DFT grid information, which becomes huge for large unit cell cases,

across all available processors. This permits a reduction in memory usage per processor.

LIMBEK

•	∗	LBEK	size of local arrays for integration weights [default 400]

4.4 Atomic parameters

The radius attributed to each atom for the integration is computed from the formal charge as specified in the initial electronic configuration. It is possible to enter for selected atoms a given atomic radius or a different formal charge.

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

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

The atomic radius or formal charge set in this way is used only the construction of the DFT grid. The value is applied to all atoms with the same formal atomic number.

Chapter 5

Semi-classical corrections for HF and DFT and composite methods (3c and sol-3c)

HF and semi-local approximations to DFT have well known limitations in the description on noncovalent interactions, mainly due to the lack of long-range Coulomb correlation. These effects lead to the omnipresent attractive London dispersion interaction (also known as attractive part of the van der Waals force). Though considered as a relatively 'weak' London dispersion interactions play a major role in all organic (organo-metallic) systems and influence the geometry, noncovalent binding energy (lattice energy), frequencies, and even reactivities, i.e. the whole potential energy surface. Thus, for these systems it is crucial to include this interaction. Different correction schemes are described and compared in Ref. [130]. In CRYSTAL, the D3 London dispersion correction, which has been mainly developed in the group of Prof. Grimme, is implemented including three-body dispersion contributions with fast analytical gradients.[126]
Due to restrictions in the computational resources, small atomic orbital basis set expansions are widely used. As the set of orbitals is neither complete nor independent of the geometry, artificial interactions occur, typically referred to as basis set superposition error. A critical analysis of the error sources and its' implications is given in Ref. [265]. Apart from the standard Boys-Bernardi couterpoise correction (see keywords GHOSTS and MOLEBSSE), a semi-empirical geometric counterpoise correction (gCP) has been implemented in CRYSTAL.[161]

3c composite methods
While both dispersion and counterpoise corrections can be applied independently, e.g. with a B3LYP-D3-gCP/6-31G* calculation, the performance is best when the parametrization is done together and basis set incompletenesses errors are simultaneously treated by an additional semi-classical correction or by compensating them with an adjusted exchange-correlation functional. These 'composite' methods are implemented in CRYSTAL and described below, namely: a minimal basis set HF with corrections (HF-3c)[266], a hybrid density functional evaluated in small double-ζ type orbitals with dispersion and counterpoise correction (PBEh-3c)[127], a screened exchange variant of PBEh-3c, dubbed HSE-3c, specifically designed to be numerically robust for molecular solids,[35] and a pure GGA functional based on the B97 XC functional evaluated in medium sized triple-ζ type orbitals with dispersion and short-range correction (B97-3c)[128].
The following schemes have been carefully tested for molecular dimers, larger molecular assemblies, host-guest-complexes, and molecular crystals.
Note that studies on inorganic crystals and metals are less well tested and should be treated carefully. In particular when using a triple-ζ basis set.

Revised composite methods for solid state calculations: sol-3c

In CRYSTAL23, a revised version of the original Grimme's composite methods has been developed with the main aim to extend their applicability to inorganic solids layered systems and metallo-organic materials. These revised methods have been tagged with a label "sol" (as for "solids") to distinguish them from the original ones proposed for molecules and molecular crystals, they share the same semi-classical corrections exploited by the "3c" Grimme's methods (D3, gGP and SRB).
The main guidelines adopted during the revision strategy can be summarized ad follows:

i employ exchange-correlation functionals developed for solids (i.e. PBEsol and HSEsol)

ii reduce the amount of HF exchange in DFT hybrid methods for a better description of electronic properties (e.g. 20-25%)

iii apply a simple recipe to make molecular basis sets suitable for inorganic solids

The resulting methods have been denoted as HFsol-3c, PBEsol0-3c and HSEsol-3c; detailed input examples can be found in the subsections below.

5.1 DFT-D3 - C_6 based London dispersion correction

For a range of common density functionals, the D3 correction can be invoked by simply specifying the XC functional with the suffix '**-D3**' in the DFT input block by means of stand-alone keywords. Here an example follows for the common B3LYP functional.

```
...
DFT
B3LYP-D3
END
...
```

By default, the D3 correction includes a damped (BJ damping) atom-atom pairwise term (no three-body term) to the KS-DFT energies (and atomic and cell gradients) (see below for details).[126]
Note that not all functionals available in CRYSTAL have been parametrized. A list of available D3 dispersion corrected DFT methods follows: BLYP, PBE, B97, B3LYP, PBE0, PW1PW, M06, HSE06, HSEsol, LC-wPBE. For other XC functionals, parameters must be provided in input (see below)

Additionally, the keyword **DFTD3**, inserted in the third input block and closed by **END**, allows users to include the three-body term, modify/specify all free method parameters, change the damping function and other options as detailed in the following.

D3 correction

The D3 augmented KS-DFT energies are obtained as

$$E_{\text{DFT-D3}} = E_{\text{KS-DFT}} + E_{\text{disp}} \tag{5.1}$$

with E_{disp} being the sum of the two- and three-body contributions to the dispersion energy:

$$E_{\text{disp}} = E^{(2)} + E^{(3)} \tag{5.2}$$

$$= -\frac{1}{2}\sum_{n=6,8}\sum_{A,B}^{\text{atoms}} s_n \frac{C_n^{AB}}{R_{AB}^n}\cdot f_n^d(R_{AB}) - \frac{1}{6}\sum_{A,B,C}^{\text{atoms}} \frac{C_9^{ABC}}{\overline{R}_{ABC}^9}\cdot f_9^d(R_{ABC},\theta_{ABC}). \tag{5.3}$$

This expression assumes dispersion coefficients that can be distributed to isotropic atom contributions. While the series should in principle incorporate higher orders, they were shown to be of less importance for many systems.[130]
In order to match the long- and midrange correlation of D3 with the semilocal correlation computed by the functional, an adequate damping function f_n^d must be included. C_n^{AB} denotes the averaged (isotropic) nth-order dispersion coefficient for atom pair AB, and R_{AB} is their internuclear distance. s_n is a functional-dependent scaling factor (see below). In contrast to the D2 method, here the C_6^{AB} coefficients are estimated from first principles by TD-DFT computation of the dynamical polarizability α for model hydrides of all elements and using the Casimir-Polder integration

$$C_6^{AB} = -\frac{3}{\pi}\int \mathrm{d}\omega\, \alpha_A(i\omega)\alpha_B(i\omega) \tag{5.4}$$

These reference C_6 are then mapped to the real system via a geometric coordination number.

Damping functions
In order to avoid near singularities for small distances (R_{AB}), the dispersion contribution needs to be damped at short distances. One possible way is to use rational *Becke-Johnson* damping as proposed by Becke and Johnson [20, 155, 156], which is the recommended option for the two-body contributions as it avoids artificial repulsive forces:

$$f_{n=6,8}^d(R_{AB}) = \frac{R_{AB}^n}{R_{AB}^n + f(R_0^{AB})^n} \tag{5.5}$$

with[156]

$$R_0^{AB} = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}} \tag{5.6}$$

and

$$f(R_0^{AB}) = a_1 R_0^{AB} + a_2. \tag{5.7}$$

In the literature this is typically used as D3(BJ).

Damping the dispersion contribution to zero for short ranges (as in Ref.[126]) is also possible:

$$f_{n=6,8}^d = \frac{1}{1 + 6(R_{AB}/(s_{r,n}R_0^{AB}))^{-\alpha_n}}. \tag{5.8}$$

Note that the R_0^{AB} used with this damping are from Ref.[126], the scheme is typically abbreviated as D3(0). For more information on the supported damping functions, see Ref [129].

The Axildor-Teller-Muto type three-body term is used in the following form:

$$f_9^d = (3\cos\theta_a\cos\theta_b\cos\theta_c + 1) \times \frac{1}{1 + 6(\overline{R}_{ABC}/(s_{r,n}\overline{R}_0^{ABC}))^{-\alpha_n}}. \tag{5.9}$$

where θ_a, θ_b and θ_c are the internal angles of the triangle formed by R_{AB}, R_{BC} and R_{CA}, $\overline{R}_{ABC}$ is the geometric mean of R_{AB}, R_{BC} and R_{AB}. The C_9 coefficient is approximated by:

$$C_9^{ABC} \approx -\sqrt{C_6^{AB}C_6^{AC}C_6^{BC}}. \tag{5.10}$$

The damping function is similar to the zero damping and has not been revised. Inclusion of the three-body term is often denoted by $\mathrm{D3^{atm}}$.

DFTD3 - opens D3 input block

VERSION - chooses between D2, D3 and the different damping functions

	rec	variable	meaning
•	*	NAT	(2) Old D2 correction, (3) D3 with zero-damping, (4) D3 with Becke-Johnson damping (default)

FUNC - chooses the functional for automatic parameter loading

	rec	variable	meaning
•	*	CHAR	List of parametrized functionals available on the homepage `http://www.thch.uni-bonn.de/tc/dftd3`

ABC - switches on the three-body dispersion term

S6 - define s_6 manually

	rec	variable	meaning
•	*	REAL	define s_6 scaling parameter manually (should be always 1.0 for GGA and hybrid functionals)

S8 - define s_8 manually

	rec	variable	meaning
•	*	REAL	define s_8 scaling parameter manually

A1 - define a_1 manually

	rec	variable	meaning
•	*	REAL	define a_1 parameter of rational Becke-Johnson damping manually

A2 - define a_2 manually

	rec	variable	meaning
•	*	REAL	define a_2 parameter of rational Becke-Johnson damping manually

A2 - define a_2 manually

	rec	variable	meaning
•	*	REAL	define a_2 parameter of rational Becke-Johnson damping manually

RS6 - define rs_6 manually

	rec	variable	meaning
•	*	REAL	define rs_6 parameter of zero-damping manually

RS8 - define rs_8 manually

	rec	variable	meaning
•	*	REAL	define rs_8 parameter of zero-damping manually

RADIUS - define two-body real-space cutoff manually

	rec	variable	meaning
•	*	REAL	define distance cutoff for real-space summation of two-body contributions [a.u.]

CNRADIUS - define the real-space cutoff of the of the coordination numbers manually

	rec	variable	meaning
•	*	REAL	define distance cutoff for real-space summation of the coordination number function [a.u.]

ABCRADIUS - define tree-body real-space cutoff manually

	rec	variable	meaning
•	*	REAL	define distance cutoff for real-space summation of three-body contributions [a.u.]

END - closes D3 input block

The older D2 corrections can still be used by means of either the option "version 2" of the DFTD3 keyword or the keyword GRIMME (see 3.3).[50, 51, 276]

Examples

A typical input looks like (urea molecular crystal):

```
Urea
CRYSTAL
0 0 0
113
5.565  4.684
5
6 0.0000  0.5000  0.3260
8 0.0000  0.5000  0.5953
7 0.1459  0.6459  0.1766
1 0.2575  0.7575  0.2827
1 0.1441  0.6441 -0.0380
ENDG
BASISSET
def2-SV(P)
DFT
B3LYP-D3
END
SHRINK
4 4
END
```

This is a simplified version of the explicit input.

```
Urea
CRYSTAL
0 0 0
113
5.565  4.684
5
6 0.0000  0.5000  0.3260
8 0.0000  0.5000  0.5953
7 0.1459  0.6459  0.1766
1 0.2575  0.7575  0.2827
1 0.1441  0.6441 -0.0380
ENDG
BASISSET
def2-SVP
DFT
PBE0
END
DFTD3
VERSION
4
S6
1.0000
S8
```

```
1.9889
A1
0.3981
A2
4.4211
END
SHRINK
4 4
END
```

A change of the damping parameter should be generally avoided as they have been carefully optimized individually for all functionals.
A more useful test is the impact of the Axilrod-Teller-Muto type three body dispersion, e.g. via

```
Urea
CRYSTAL
0 0 0
113
5.565  4.684
5
6 0.0000  0.5000  0.3260
8 0.0000  0.5000  0.5953
7 0.1459  0.6459  0.1766
1 0.2575  0.7575  0.2827
1 0.1441  0.6441 -0.0380
ENDG
BASISSET
def2-SVP
DFT
PBE0
END
DFTD3
ABC
END
SHRINK
4 4
END
```

More input examples can be found at the CRYSTAL tutorials web page.
Note that cohesive energies at the given level will suffer severely from basis set superposition errors.

5.2 GCP - Geometrical counterpoise correction

The keyword **GCP**, inserted in the last section of the CRYSTAL input and closed by **END**, calculates atom-pair wise counterpoise corrections which are added to the KS-DFT energies (and atomic and cell gradients).[161]

Note that the gCP scheme has been carefully tested for molecular crystals. Applications to inorganic crystals and metals are less well tested and should be treated carefully.

gCP correction
The central idea is to add in a semi-empirical fashion an energy correction ΔE_{gCP} to the energies of molecular (or periodic) systems in order to remove artificial overbinding effects from

BSSE.[161, 34] As the focus lies on the contribution of individual atoms a natural outcome is its ability to yield also intramolecular BSSE corrections. The parametrization is constructed such that it approximates the Boys and Bernadi[32] counterpoise (CP) correction ΔE_{CP} in the intermolecular case

$$\Delta E_{\mathrm{CP}} \approx \Delta E_{\mathrm{gCP}} \ , \tag{5.11}$$

where e.g. for a complexation reaction $A + B \rightarrow C$ our correction is given by

$$\Delta E_{\mathrm{gCP}} = E_{\mathrm{gCP}}(C) - E_{\mathrm{gCP}}(A) - E_{\mathrm{gCP}}(B) \ . \tag{5.12}$$

In practice, E_{gCP} can simply be added to the HF/DFT energy

$$E_{\mathrm{total}} = E_{\mathrm{HF/DFT}} + E_{\mathrm{gCP}} \ . \tag{5.13}$$

The central equation over all atoms N reads:

$$E_{\mathrm{gCP}} = \sigma \cdot \sum_{A}^{N} \sum_{\mathbf{g}}{}' \sum_{B}^{N} e_A^{\mathrm{miss}} \cdot f_{\mathrm{dec}}(R_{AB} + \mathbf{g}) \ , \tag{5.14}$$

where the energy e_A^{miss} is a measure for the incompleteness for the chosen target basis set (that is typically small), and $f_{\mathrm{dec}}(R_{AB})$ is a decay function that depends on the inter-atomic distance R_{AB}.

$$f_{\mathrm{dec}}(R_{AB}) = \frac{\exp\left(-\alpha {R_{AB}}^{\beta}\right)}{\sqrt{A_{AB} N_{\mathrm{virt}}^{B}}} \tag{5.15}$$

The potential is normalized by the Slater-type overlap S_{AB}, the number of virtual orbitals N_{virt}^{B}, and the empirical parameters α and β. In molecular systems, the sum over the translation invariant vectors $\mathbf{g}$ is omitted. For periodic systems, a sum over all atompairs inside a supercell is utilized with default distance-cutoff of 60 Bohr analogue to the non-periodic version. The prime ($'$) indicates that for $\mathbf{g} = 0$, $A \neq B$. The scaling factor σ is one out of 4 parameters needed for every `<method>/<basis set>` combination.

Small basis sets show not only a large BSSE, but general shortcomings. These effects are not always clearly distinguishable. If computationally affordable, large basis sets (triple-ζ and higher) are always preferable for a given system.
The gCP has been developed for a range of standard basis sets, including MINIS, def2-SV(P), def2-SVP, def-TZVP, def2-TZVP, 6-31G*, pob-TZVP. The parametrization distinguishes between generalized gradient approximated functionals 'GGA', hybrid functionals 'DFT', and pure Hartree-Fock 'HF'. The DFT parametrization will often also work GGA functionals, but the couterpoise correction is typically underestimated. The following combinations are parametrized:

	GGA	DFT	HF
MINIS	yes	no	yes
SV	yes	yes	yes
SV(P)	no	yes	yes
SVP	yes	yes	yes
def-TZVP	no	yes	yes
def2-TZVP	no	yes	yes
6-31G*	no	yes	yes
pob-TZVP	no	yes	no

A more extensive list can be found at the web page:

`http://www.thch.uni-bonn.de/tc/?section=downloads`.

The gCP correction can be invoked by a single line in the third CRYSTAL input block

GCPAUTO - It includes the gCP correction with automatic parameter setup

This is however only possible if the basis set has been defined with a general **BASISSET** keyword. Then, the routine automatically selects the most suitable available parametrization. Otherwise, the gCP correction has to be defined by an individual input block:

GCP - opens gCP input block

METHOD - choose gCP method combination

	rec	variable	meaning
•	*	CHAR	Choose between different gCP parametrization. The method distinguishes between different types of Hamiltonians (HF, GGA, DFT) and the used basis set name, separated by a slash. E.g., dft/svp

SIGMA - define σ manually

	rec	variable	meaning
•	*	REAL	define σ scaling parameter manually

ALPHA - define α manually

	rec	variable	meaning
•	*	REAL	define α parameter manually

BETA - define β manually

	rec	variable	meaning
•	*	REAL	define β parameter manually

ETA - define η manually

	rec	variable	meaning
•	*	REAL	define η parameter manually

RADIUS - define real-space cutoff manually

	rec	variable	meaning
•	*	REAL	define distance cutoff for real-space summation [a.u.]

PRINTEMISS - extensive output with e_{miss}, N_{virt}, and individual E_{gcp} contributions

END - closes gCP input block

Examples

A typical input looks like (urea molecular crystal):

```
Urea
CRYSTAL
0 0 0
113
5.565  4.684
5
```

```
6 0.0000  0.5000  0.3260
8 0.0000  0.5000  0.5953
7 0.1459  0.6459  0.1766
1 0.2575  0.7575  0.2827
1 0.1441  0.6441 -0.0380
BASISSET
def2-SVP
DFT
PBE0
END
DFTD3
ABC
END
GCPAUTO
SHRINK
4 4
END
```

This is a simplified version of the following explicit input.

```
Urea
CRYSTAL
0 0 0
113
5.565  4.684
5
6 0.0000  0.5000  0.3260
8 0.0000  0.5000  0.5953
7 0.1459  0.6459  0.1766
1 0.2575  0.7575  0.2827
1 0.1441  0.6441 -0.0380
BASISSET
def2-SVP
DFT
PBE0
END
DFTD3
VERSION
4
S6
1.0000
S8
1.9889
A1
0.3981
A2
4.4211
END
GCP
METHOD
DFT/SVP
SIGMA
0.2424
ETA
1.2371
ALPHA
0.6076
BETA
```

```
1.4078
END
SHRINK
4 4
END
```

5.3 *3c* - composite methods

5.3.1 HF-3C - Minimal basis HF composite method

The keyword **HF3C**, inserted in the last section of the CRYSTAL input and closed by the keyword **END**, calculates three semi-classical corrections (D3, gCP, SRB), which are added to the HF energies (and atomic and cell gradients).[266]

The HF-3c method
This composite method is based on a minimal basis set HF calculation as:

$$E_{\text{HF-3c}} = E_{\text{HF/MINIX}} + E_{\text{D3}} + E_{\text{gCP}} + E_{\text{SRB}} \tag{5.16}$$

It is designed to compete with semi-empirical (tight–binding or NDO) methods without neglecting any of the many-center integrals. This keyword should only be used in combination with a pure HF calculation in the MINIX orbital basis set. Its original design targeted organic complexes[266], but it has been used for periodic systems[36], and specifically tested as implemented in CRYSTAL.[55].

Note that the HF-3c method has been carefully tested for molecular crystals. Applications to inorganic crystals and metals are less well tested and should be treated carefully.

The keyword **HF3C** automatically switches on all needed correction schemes with proper parameter setup. While the D3 and gCP schemes have been described before (see sections 5.1), a third short-range basis (SRB) correction is introduced. This correction mainly improves the bond distances of electronegative elements that have a systematic error in the restricted basis set expansion.

HF3C - opens HF-3c input block

RESCALES8 - scale C_8 contribution in D3 correction

	rec	variable	meaning
•	*	REAL	scales the C_8 contribution in the D3 dispersion correction. Can also be adjusted individually in the **DFTD3** input block.

SCALEGCP - scale contribution from gCP correction

	rec	variable	meaning
•	*	REAL	scales the contribution from the gCP correction. Can also be adjusted individually in the **GCP** input block.

END - closes HF-3c input block

Examples
The urea crystal can be treated like

```
Urea
CRYSTAL
```

```
0 0 0
113
5.565  4.684
5
6 0.0000  0.5000  0.3260
8 0.0000  0.5000  0.5953
7 0.1459  0.6459  0.1766
1 0.2575  0.7575  0.2827
1 0.1441  0.6441 -0.0380
BASISSET
MINIX
HF3C
END
SHRINK
4 4
END
```

The scaled s-HF-3c with scaled down C_8 dispersion contribution as introduced in Ref. [55] can be defined by

```
Urea
CRYSTAL
0 0 0
113
5.565  4.684
5
6 0.0000  0.5000  0.3260
8 0.0000  0.5000  0.5953
7 0.1459  0.6459  0.1766
1 0.2575  0.7575  0.2827
1 0.1441  0.6441 -0.0380
BASISSET
MINIX
HF3C
RESCALES8
0.70
END
SHRINK
4 4
END
```

5.3.2 PBEH-3C - Small basis DFT composite method

The keyword **PBEH3C**, inserted in the DFT Hamiltonian input block as a standalone keyword (see section 4), corresponds to a modified global hybrid functional PBEh with semi-classical corrections (D3, gCP).[127]

$$E_{\text{PBEh-3c}} = E_{\text{PBEh/def2-mSVP}} + E_{\text{D3}} + E_{\text{gCP}} \tag{5.17}$$

This composite method is based on a modified double-ζ orbital set dubbed def2-mSVP, which should be always used when performing PBEh-3c calculations. The basis set development and explicit definition is given in Ref [127]. The PBE type correlation and exchange enhancement factors have been optimized to absorb certain incompleteness errors of the restricted basis set. The percentage of HF exchange (42%) is adjusted to get covalent bond length that are on average correct.

Note that the PBEh-3c method has been carefully tested for molecular crystals. Applications to inorganic crystals and metals are less well tested and should be

treated carefully..

The keyword **PBEH3C** automatically switches on all needed correction schemes with proper parameter setup. The D3 and gCP schemes have been described before (see sections 5.1), in order to avoid contributions of the gCP for covalent bond length and thermochemistry, a short-range damping has been introduced. The D3 correction is used in its rational Becke-Johnson damping scheme and includes per default the three-body Axilrod-Teller-Muto term.

Note that vibrational frequencies computed with PBEh-3c will be scaled by 0.95 as recommended in its original publication. See section (8.1) for changing the frequency scaling factor manually.

Example
The urea crystal can be treated like

```
Urea
CRYSTAL
0 0 0
113
5.565  4.684
5
6 0.0000  0.5000  0.3260
8 0.0000  0.5000  0.5953
7 0.1459  0.6459  0.1766
1 0.2575  0.7575  0.2827
1 0.1441  0.6441 -0.0380
BASISSET
def2-mSVP
DFT
PBEH3C
END
SHRINK
4 4
END
```

5.3.3 HSE-3C - Small basis DFT composite method

The keyword **HSE3C**, inserted in the DFT Hamiltonian block as standalone keyword (see section 4), corresponds to a modified screened exchange hybrid functional HSE with semi-classical corrections (D3, gCP).[35]

$$
\begin{aligned}
E_{HSE-3c} =& E_{\text{HSE/def2-mSVP}} + E_{D3} + E_{gCP} \\
E_{\text{xc}}^{\text{HSE-3c}} =& a\, E_{\text{x}}^{\text{HF,SR}}(\omega) + (1-a)\, E_{\text{x}}^{\text{PBEh-3c,SR}}(\omega) + E_{\text{x}}^{\text{PBEh-3c,LR}}(\omega) + E_{\text{c}}^{\text{PBEh-3c}}\,. \qquad (5.18)
\end{aligned}
$$

In close correspondence to PBEh-3c, this composite method is based on the modified double-ζ basis set def2-mSVP, which should be always used when performing HSE-3c calculations.[127]. The Henderson-Janesko-Scuseria (HJS) exchange hole model[269, 85] has been parametrized to closely reproduce the PBEh-3c exchange enhancement factor. The correlation enhancement is kept fixed, the short-range HF exchange is $a = 0.42$, and the standard error function separation (controlled via the parameter $\omega = 0.11$)[2] switches the HSE-3c potential to a pure generalized gradient approximation (GGA) at long-range.

Note that the HSE-3c method has been carefully tested for molecular crystals. Applications to inorganic crystals and metals are less well tested and should be treated carefully.

The keyword **HSE-3c** automatically switches on all needed correction schemes with proper parameter setup. The D3 and (SR-damped) gCP schemes have been described before (see sections 5.1), The D3 correction is used in its rational Becke-Johnson damping scheme and includes per default the three-body Axilrod-Teller-Muto term. The accuracy of HSE-3c has been shown to be identical to PBEh-3c, but we see a significantly computational speed-up. This holds especially for organic solids with smaller band gaps (e.g. oligoacenes).[35] Note that frequencies computed with HSE-3c will be scaled by 0.95 as recommended in its original publication. See section (8.1) for changing the frequency scaling faction manually.

The urea crystal can be treated like

```
Urea
CRYSTAL
0 0 0
113
5.565  4.684
5
6 0.0000  0.5000  0.3260
8 0.0000  0.5000  0.5953
7 0.1459  0.6459  0.1766
1 0.2575  0.7575  0.2827
1 0.1441  0.6441 -0.0380
BASISSET
def2-mSVP
DFT
HSE3C
END
SHRINK
4 4
END
```

5.3.4 B97-3C - Medium basis DFT composite method

The keyword **B973C**, inserted in the DFT Hamiltonian block as standalone keyword (see section 4), corresponds to a modified GG functional of the B97 form with semi-classical corrections (D3, SRB).[128]

$$E_{\text{B97-3c}} = E_{\text{B97/def2-mTVP}} + E_{\text{D3}} + E_{\text{SRB}} \tag{5.19}$$

This composite method is based on a modified triple-ζ orbital set dubbed mTZVP, which should be always used when performing B97-3c calculations. The basis set development and explicit definition is given in Ref [128]. The B97 type correlation and exchange enhancement factors have been optimized to absorb certain incompleteness errors of the still not converged basis set. Due to the rather large basis set and the flexible exchange-correlation form most basis set errors can be absorbed into the functional parametrization. Thus, the semi-empirical gCP scheme is not needed. However, all bond-length are systematically too long, which actually holds for all GGA functionals. To correct this, the SRB atom-pair correction (see section 5.3.1) has been adopted.

Note that the B97-3c method has been carefully tested for molecular crystals. Applications to inorganic crystals and metals are less well tested and should be treated carefully..

The keyword **B973C** automatically switches on all needed correction schemes with proper parameter setup. The D3 and SRB schemes have been described before (see sections 5.1), The D3 correction is used in its rational Becke-Johnson damping scheme and includes per default the three-body Axilrod-Teller-Muto term.

The urea crystal can be treated like

```
Urea
CRYSTAL
0 0 0
113
5.565  4.684
5
6 0.0000  0.5000  0.3260
8 0.0000  0.5000  0.5953
7 0.1459  0.6459  0.1766
1 0.2575  0.7575  0.2827
1 0.1441  0.6441 -0.0380
BASISSET
mTZVP
DFT
B973C
END
SHRINK
4 4
END
```

5.4 Revised composite methods for solid state calculations: *SOL-3c*

5.4.1 HFSOL-3C - Minimal basis HF revised composite method

The keyword **HFSOL3C**, inserted in the last section of the CRYSTAL input and closed by the keyword **END**, calculates three semi-classical corrections (D3, gCP, SRB), which are added to the HF energies (and atomic and cell gradients).[70]

The HFSOL-3C method
This composite method is based on a minimal basis set revised for inorganic solids HF calculation as:

$$E_{\text{HFSOL-3c}} = E_{\text{HF/SOLMINIX}} + E_{\text{D3}} + E_{\text{gCP}} + E_{\text{SRB}} \tag{5.20}$$

This keyword should only be used in combination with a pure HF calculation in the SOLMINIX orbital basis set. The keyword **HFSOL3C** automatically switches on all needed correction schemes with proper parameter setup. While the D3 and gCP schemes have been described before (see sections 5.1), a third short-range basis (SRB) correction is introduced. This correction mainly improves the bond distances of electronegative elements that have a systematic error in the restricted basis set expansion.

Examples
The MgO crystal can be treated like

```
MgO
CRYSTAL
0 0 0
225
4.207
2
12  0.0 0.0 0.0
 8  0.5 0.5 0.5
BASISSET
```

```
SOLMINIX
HFSOL3C
END
SHRINK
8 8
END
```

5.4.2 PBESOL0-3C - Small basis DFT revised composite method

The keyword **PBESOL03C**, inserted in the DFT Hamiltonian input block as a standalone keyword (see section 4), corresponds to the global hybrid functional PBESOL0 with semi-classical corrections (D3, gCP).

$$E_{\text{PBESOL0-3C}} = E_{\text{PBESOL0/sol-def2-mSVP}} + E_{\text{D3}} + E_{\text{gCP}} \tag{5.21}$$

This composite method is based on a modified double-ζ orbital set revised for solid state dubbed sol-def2-mSVP, which should be always used when performing PBESOL0-3C calculations. The basis set revision and explicit definition is given in Ref [70]. The keyword **PBESOL03C** automatically switches on all needed correction schemes with proper parameter setup. The D3 and gCP schemes have been described before (see sections 5.1), in order to avoid contributions of the gCP for covalent bond length and thermochemistry, a short-range damping has been introduced. The D3 correction is used in its rational Becke-Johnson damping scheme and includes per default the three-body Axilrod-Teller-Muto term.

Example
The Diamond crystal can be treated like

```
C (diamond)
CRYSTAL
0 0 1
227
3.53713419
1
6  0.0000  0.0000  0.0000
BASISSET
SOLDEF2MSVP
DFT
PBESOL03C
END
SHRINK
8 8
END
```

5.4.3 HSESOL-3C - Small basis DFT revised composite method

The keyword **HSESOL3C**, inserted in the DFT Hamiltonian block as standalone keyword (see section 4), corresponds to the screened exchange hybrid functional HSESOL with semi-classical corrections (D3, gCP).

$$E_{HSESOL-3C} = E_{\text{HSESOL/sol-def2-mSVP}} + E_{D3} + E_{gCP} \tag{5.22}$$

In close correspondence to PBESOL0-3C, this composite method is based on the modified double-ζ basis set sol-def2-mSVP, which should be always used when performing HSESOL-3c calculations.[70].

The keyword **HSESOL-3C** automatically switches on all needed correction schemes with proper parameter setup. The D3 and (SR-damped) gCP schemes have been described before (see sections 5.1), The D3 correction is used in its rational Becke-Johnson damping scheme and includes per default the three-body Axilrod-Teller-Muto term.
The α-quartz crystal can be treated like

```
QUARTZ ALFA (init. geom.: expt.)
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
BASISSET
SOLDEF2MSVP
DFT
HSESOL3C
END
SHRINK
6 6
END
```

Chapter 6

Spin-Orbit Coupling, Non-Collinear Magnetism, Two-Component Spinors

Relativistic effects in quantum chemistry and materials physics, refer to corrections to the Schrödinger equation from an account that the speed of light is finite and constant. Such corrections become increasingly important moving down the periodic table, to heavier elements, in which the effective velocity of the electrons becomes non-negligeable when compared to the speed of light. They can be classified into two categories, depending on whether or not they are described through scalar operators in the Hamiltonian. The first category comprises of scalar-relativistic (SR) effects, and the second is here loosely referred to as spin-orbit coupling (SOC) effects.

The reformulation of the (one-component) Schrödinger equation that is consistent with the postulates of special relativity is the four-component Dirac equation. In the Dirac equation, a crystalline orbital (CO) is a 4×1 vector function (a four-component "spinor"), rather than a 1×1 scalar function $\psi_{\mathbf{k}}^{\text{Schröd}}(\mathbf{r})$ of space (a one-component, 1c, spinor), as well as the electron quasi-momentum $\mathbf{k}$ (i.e. the sampling point in the first-Brillouin zone, FBZ). More explicitly, a Schrödinger CO ket reads:

$$|\psi_{\mathbf{k}}\rangle = \int d\mathbf{r}\ \psi_{\mathbf{k}}^{\text{Schröd}}(\mathbf{r})\ \ |\mathbf{r}\rangle \tag{6.1}$$

while a Dirac ket is decomposed as:

$$|\psi_{\mathbf{k}}\rangle = \int d\mathbf{r} \sum_{\sigma=\uparrow,\downarrow} \sum_{a=L,S} \psi_{\sigma,a,\mathbf{k}}^{\text{Dirac}}(\mathbf{r})\ \ |\mathbf{r},\sigma,a\rangle \tag{6.2}$$

in which σ is a spin index and a denotes the so-called "large" and "small" components of the Dirac wavefunction, principally related to positive (electronic) and negative (positronic) energy solutions of the Dirac equation, respectively.

In quantum chemistry and materials physics, interest is dominated by the electronic solution of the Dirac equation. It is therefore common practice to write $|\psi_{\mathbf{k}}\rangle$ in a basis of two-component (2c, or Pauli) spinors:

$$|\psi_{\mathbf{k}}\rangle = \int d\mathbf{r} \sum_{\sigma=\uparrow,\downarrow} \psi_{\sigma,\mathbf{k}}^{\text{Pauli}}(\mathbf{r})\ \ |\mathbf{r},\sigma\rangle \tag{6.3}$$

$\psi_{\sigma,\mathbf{k}}^{\text{Pauli}}(\mathbf{r})$ are components of the 2×1 vector CO $\boldsymbol{\psi}_{\mathbf{k}}^{\text{Pauli}}(\mathbf{r})$:

$$\boldsymbol{\psi}_{\mathbf{k}}^{\text{Pauli}}(\mathbf{r}) = \begin{pmatrix} \psi_{\uparrow,\mathbf{k}}^{\text{Pauli}}(\mathbf{r}) \\ \psi_{\downarrow,\mathbf{k}}^{\text{Pauli}}(\mathbf{r}) \end{pmatrix} \tag{6.4}$$

Comparing Eq. (6.1) and Eq. (6.3), one immediate consequence is that the Kohn-Sham equation:

$$\hat{F}|\psi_{\mathbf{k}}\rangle = \epsilon(\mathbf{k})|\psi_{\mathbf{k}}\rangle \tag{6.5}$$

in Pauli spinor representation leads to a 2×2 Fock operator, $\hat{F}$, instead of a 1×1 Fock in the Schrödinger spinor basis. A relativistic theory including spin, is then associated to a two-component self-consistent field (2c-SCF), rather than 1c-SCF, procedure. Such a 2c-SCF strategy has recently been implemented in the CRYSTAL code [62, 63, 65, 64, 66].
In matrix form, Eq. (6.5) reads, in the Schrödinger spinor basis:

$$\mathbf{F_k C_k} = \mathbf{S_k C_k E_k} \tag{6.6}$$

and in the Pauli spinor basis:

$$\begin{pmatrix} \mathbf{F}_\mathbf{k}^{\uparrow\uparrow} & \mathbf{F}_\mathbf{k}^{\uparrow\downarrow} \\ \mathbf{F}_\mathbf{k}^{\downarrow\uparrow} & \mathbf{F}_\mathbf{k}^{\downarrow\downarrow} \end{pmatrix} \begin{pmatrix} \mathbf{C}_\mathbf{k}^{\uparrow} \\ \mathbf{C}_\mathbf{k}^{\downarrow} \end{pmatrix} = \begin{pmatrix} \mathbf{S}_\mathbf{k}^{\uparrow\uparrow} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_\mathbf{k}^{\downarrow\downarrow} \end{pmatrix} \begin{pmatrix} \mathbf{C}_\mathbf{k}^{\uparrow} \\ \mathbf{C}_\mathbf{k}^{\downarrow} \end{pmatrix} \mathbf{E_k} \tag{6.7}$$

where $\mathbf{E_k}$ is the diagonal matrix of band structure energies $\epsilon_i(\mathbf{k})$, having size $N \times N$ for a calculation with N basis functions in the Schrödinger spinor basis, or size $2N \times 2N$ in the Pauli spinor basis. $\mathbf{C}_\mathbf{k}^\sigma$ is the $N \times 2N$ matrix of expansion coefficients $C_{\mu,i}^\sigma(\mathbf{k})$ of the Pauli COs in Bloch functions:

$$\psi_{\sigma,i,\mathbf{k}}^{\text{Pauli}}(\mathbf{r}) = \sum_\mu C_{\mu,i}^\sigma(\mathbf{k})\, \phi_{\mu,\mathbf{k}}(\mathbf{r}) \tag{6.8}$$

while, for Schrödinger spinors, we make use of the $N \times N$ matrix $\mathbf{C_k}$ with elements $C_{\mu,i}(\mathbf{k})$:

$$\psi_{i,\mathbf{k}}^{\text{Schröd}}(\mathbf{r}) = \sum_\mu C_{\mu,i}(\mathbf{k})\, \phi_{\mu,\mathbf{k}}(\mathbf{r}) \tag{6.9}$$

Finally, in Eq. (6.7), $\mathbf{S}_\mathbf{k}^{\sigma\sigma'}$ and $\mathbf{F}_\mathbf{k}^{\sigma\sigma'}$ are the $N \times N$ spin-blocks of the Bloch function overlap and Fock matrices.
For CRYSTAL, in Eqs. (6.8) and (6.9), the Bloch functions are the inverse Fourier transform of *pure-real* atom-centered local functions (termed atomic orbitals, AOs):

$$\phi_{\mu,\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{g}=-\infty}^{\infty} e^{\imath \mathbf{k}\cdot\mathbf{g}}\, \chi_\mu(\mathbf{r} - \mathbf{g} - \mathbf{a}_\mu) \tag{6.10}$$

with $\mathbf{a}_\mu$ being the position of the atom on which χ_μ is centered in the reference cell $\mathbf{0}$, $\imath$ is the imaginary unit, and Ω is the volume of the FBZ.

Users of this module are kindly requested to cite:

1. *Desmarais, J. K., Flament, J. P., Erba, A. (2019). Spin-orbit coupling from a two-component self-consistent approach. I. Generalized Hartree-Fock theory. The Journal of chemical physics, 151(7), 074107.*
2. *Desmarais, J. K., Flament, J. P., Erba, A. (2019). Fundamental role of fock exchange in relativistic density functional theory. The Journal of Physical Chemistry Letters, 10(13), 3580-3585.*
3. *Desmarais, J. K., Flament, J. P., Erba, A. (2020). Adiabatic connection in spin-current density functional theory. Physical Review B, 102(23), 235118.*
4. *Desmarais, J. K., Flament, J. P., Erba, A. (2020). Spin-orbit coupling in periodic systems with broken time-reversal symmetry: Formal and computational aspects. Physical Review B, 101(23), 235142.*
5. *Desmarais, J. K., Komorovsky, S., Flament, J. P., Erba, A. (2021). Spin–orbit coupling from a two-component self-consistent approach. II. Non-collinear density functional theories. The Journal of Chemical Physics, 154(20), 204110.*

The following keywords for the SCF block of the input file are compatible with a 2c-SCF calculation. Note, in particular, that the present 2c-SCF implementation only supports pure LDA and GGA functionals, as well as global-hybrid LDA and GGA functionals. geometry optimizations, frequency calculations, response properties, etc... are not available in the two-component spinor basis. All options available with the 2c-SCF are contained in the present chapter. ***Any feature not explicitly mentionned in this chapter is not supported for calculations in the two-component spinor basis.***

6.1 Computational Parameters, Hamiltonian

Single particle Hamiltonian

[GHF]	Generalized Hartree-Fock (Default)	–	–
[SC]DFT	Spin-Current DFT Hamiltonian	126	–
SDFT	Spin DFT Hamiltonian	177	–

Choice of the exchange-correlation functionals

EXCHANGE	exchange functional	129	I

LDA functionals

LDA	Dirac-Slater [68] (LDA)
VBH	von Barth-Hedin [281] (LDA)

GGA functionals

BECKE	Becke 1988[18] (GGA)
PBE	Perdew-Becke-Ernzerhof 1996 [204] (GGA)
PBESOL	GGA. PBE functional revised for solids [206]
mPW91	modified Perdew-Wang 91 (GGA)
PWGGA	Perdew-Wang 91 (GGA)
SOGGA	second order GGA. [302]
WCGGA	GGA - Wu-Cohen [293]

CORRELAT	correlation functional	129	I

LDA functionals

PZ	Perdew-Zunger [210] (LDA)
VBH	von Barth-Hedin [281] (LDA)
VWN	Vosko,-Wilk-Nusair [282] (LDA)

GGA functionals

LYP	Lee-Yang-Parr [167] (GGA)
P86	Perdew 86 [202] (GGA)
PBE	Perdew-Becke-Ernzerhof [204] (GGA)
PBESOL	GGA. PBE functional revised for solids [206]
PWGGA	Perdew-Wang 91 (GGA)
PWLSD	Perdew-Wang 92 [208, 209, 207] (GGA)
WL	GGA - Wilson-Levy [291]

Standalone keywords: exchange+correlation

SVWN	see [68, 282]	129
BLYP	see [18, 167]	129
PBEXC	see [204]	129
PBESOLXC	see [206]	129
SOGGAXC	see [302]	129

Global Hybrid functionals

Standalone keywords

B3PW	B3PW parameterization	130	–
B3LYP	B3LYP parameterization	130	–
PBE0	Adamo and Barone [5]	130	
PBESOL0	Derived from PBE0	130	
B1WC	see [25]	130	
WC1LYP	see [60]	130	
B97H	see [4, 101]	130	
PBE0-13	see [42]	130	
mPW1PW91	see [6]	130	
mPW1K	see [175]	130	

User defined global hybrids

HYBRID	hybrid mixing	130	I

Spin(-current) DFT Formulation				
[**COLLINEAR**]	Collinear spin(-current) dft (Default)		–	–
NONCOLC	Canonical non-collinear spin(-current) dft		176	–
NONCOLSF	Scalmani-Frisch non-collinear spin(-current) dft		176	–
TOLM	Developers only		177	I
Numerical accuracy control				
ANGULAR	definition of angular grid	137	I	
RADIAL	definition of radial grid	137	I	
[**BECKE**]	selection of Becke weights (default)	137	–	
SAVIN	selection of Savin weights	137	–	
OLDGRID	"old" default grid	140		
LGRID	"large" predefined grid	140		
[**XLGRID**]	"extra large" predefined grid (new default)			
XXLGRID	"extra extra large" predefined grid	140		
RADSAFE	safety radius for grid point screening		I	
TOLLDENS	density contribution screening 6	141	I	
TOLLGRID	grid points screening 14	141	I	
[**BATCHPNT**]	grid point grouping for integration	141	I	
CHUNKS	max n. points in a batch for numerical int.	142	I	
DISTGRID	distribution of DFT grid across nodes	142		
LIMBEK	size of local arrays for integration weights 400	142	I	
Atomic parameters control				
RADIUS	customized atomic radius	143	I	
FCHARGE	customized formal atomic charge	143	I	
Auxiliary				
END	close DFT input block			

The **TWOCOMPON** block : 2c-SCF control keywords

TWOCOMPON	Open keyword block and request 2c-SCF	165	–
2NDVARIAT	second-variational calculation	168	I
GCOREROT	Core Hamiltonian guess with infinitesimal magnetization	165	I
GUESSPATNC	Superposition of atomic densities with magnetization	166	I
GUESSP[NOSO]	Guess from a previous 1c-SCF calculation	168	–
GUESSPSO	Guess from a previous 2c-SCF calculation	168	–
GUESSROTM	Rotate the magnetization from a previous 1c-SCF or 2c-SCF	168	I
PRTENESOC	Detailed printing of energy contributions	170	–
SPINORLOCK	Fix number of occupied spinors for a preset number of cycles	170	–
SOC	Include a SOC operator in the Hamiltonian	170	–
END	close **TWOCOMPON** block	–	–

Numerical accuracy and computational parameters control

BIPOLAR	Bipolar expansion of bielectronic integrals	95	I
BIPOSIZE	size of coulomb bipolar expansion buffer	95	I
EXCHSIZE	size of exchange bipolar expansion buffer	95	I
[**EXCHPERM**]	use permutation of centers in exchange integrals (Default)	101	–
[**GUESSPAT**]	Superposition of atomic densities (Default)	110	–
ILASIZE	Maximum size of array ILA for 2-electron integral calculation 6000	112	I
INTGPACK	Choice of integrals package 0	112	I
MADELIND	reciprocal lattice vector indices for Madelung sums 50	114	I
NOBIPCOU	Coulomb bielectronic integrals computed exactly	116	–
NOBIPEXCH	Exchange bielectronic integrals computed exactly	116	–
NOBIPOLA	All bielectronic integrals computed exactly	116	–
POLEORDR	Maximum order of multipolar expansion 4	117	I
TOLINTEG	Truncation criteria for bielectronic integrals 6 6 6 6 12	125	I
TOLPSEUD	Pseudopotential tolerance 6	125	I
Type of run			
ATOMHF	Atomic wave functions	93	I
SCFDIR	SCF direct (mono+biel int computed)	118	–
EIGS	S(**k**) eigenvalues - basis set linear dependence check	98	–
FIXINDEX	Reference geometry to classify integrals	104	–
Basis set - AO occupancy			
FDAOSYM	f and d degeneracies analysis	101	I
FDAOCCUP	f and d orbital occupation guess	102	I
Numerical accuracy control and convergence tools			
FMIXING	Fock/KS matrix (cycle i and i-1) mixing 0	105	I
LEVSHIFT	level shifter no	113	I
MAXCYCLE	maximum number of cycles 50	114	I
SMEAR	Finite temperature smearing of the Fermi surface no	121	I
TOLDEE	convergence on total energy 6	125	I
Auxiliary and control keywords			
END	terminate processing of block3 input		–
MAXNEIGHB	maximum number of equidistant neighbours from an atom	49	I
NEIGHBOR	number of neighbours to analyse in **PPAN**	59	I
QVRSGDIM	maximum size of mutipole moment gradient array 90000000	118	I
SETINF	setting of inf array options	66	I
STOP	execution stops immediately	68	–
TESTPDIM	stop after symmetry analysis	124	–
TEST[**RUN**]	stop after integrals classification and disk storage estimate	125	–
Post SCF calculations			
POSTSCF	post-scf calculations when convergence criteria not satisfied	117	–
PPAN	population analysis at the end of the SCF no	117	

6.2 The Two-Component Self-Consistent Field Procedure Control

To activate a 2c-SCF, rather than 1c-SCF calculation, all that is necessary is to open a **TWOCOMPON** block in the SCF part of the input file:

```
...
99 0
END          end of basis set input
...
TWOCOMPON    insert TWOCOMPON block to request
END          performing the SCF in a basis of 2c spinors
...
```

Opening the **TWOCOMPON** block deactivates DIIS and symmetry (see keywords **NODIIS** and **SYMMREMO** at pages p. 96 and 73). Time-reversal symmetry is not exploited in the present 2c-SCF implementation, so that both closed-shell and open-shell systems are treated on the same footing. The following additional keywords, specific to the 2c-SCF procedure, may be inserted in the **TWOCOMPON** block:

TWOCOMPON	open 2c-SCF control keywords block
	Available optional keywords II
Options for the guess density matrix	
[**GUESSPAT**]	Guess from a superposition of scalar-relativistic atomic densities (Default)
GCOREROT	Core Hamiltonian guess with infinitesimal magnetization along selected direction
GUESSPATNC	Guess from a superposition of scalar-relativistic atomic densities with non-collinear magnetization
GUESSP[**NOSO**]	Guess from a previous 1c-SCF calculation
GUESSPSO	Guess from a previous 2c-SCF calculation
GUESSROTM	Rotate the magnetization in the starting guess from a previous 1c-SCF or 2c-SCF
Other keywords specific to the 2c-SCF	
2NDVARIAT	second-variational calculation
PRTENESOC	detailed printing of energy contributions
SPINORLOCK	fix number of occupied spinors for a preset number of cycles
SOC	include a spin-orbit coupling operator in the Hamiltonian
END	close TWOCOMPON block

GCOREROT - Core Hamiltonian guess with infinitesimal magnetization along selected direction

	rec	variable	meaning
•	*	RTHETA,RPHI	real-valued polar and azimuthal angles to define orientation of magnetization

Use the core Hamiltonian as a starting guess (guess density matrix $\mathbf{P} = \mathbf{0}$), with an infinitesimal magnetization along a selected direction. The direction of the (vanishingly small) guess magnetization is determined by the polar and azimuthal angles RTHETA and RPHI. Refer to the keyword **GUESSPATNC** at p. 166 for the explicit definition of RTHETA and RPHI angles.

The following provides an example input sequence with infinitesimal magnetization along the y-axis:

```
...
TWOCOMPON
GCOREROT
90.0 0.0
END
...
```

GUESSPATNC - Guess from a superposition of scalar-relativistic atomic densities with non-collinear magnetization

	rec	variable	meaning
•	*	NB	no. of distinct electronic configurations to be modified
			insert NB records II
•	*	LA,RTHETA,RPHI,RSCALE	*atom labels*, polar angle, azimuthal angle, and scale factor for the magnetization

As in the 1c-SCF, the standard initial starting guess for the 2c-SCF is a superposition of non- or scalar-relativistic atomic densities (see keyword **GUESSPAT** at p. 110). Such non-interacting atomic densities are calculated using a multiconfigurational Hartree-Fock approach described in Ref. [53]. The electronic configuration, including the total spin (or magnetization) on each atomic center is fully determined by the shell occupation numbers entered in the basis set input (see p. 20). By default, the magnetization is oriented along the z-axis.
A distinct atomic guess is calculated for each distinct basis set (i.e. typically each distinct chemical element) provided in the basis set input section. *A different atomic guess on atoms with the same atomic number requires using the keyword* **CHEMOD** (see p. 75).
The present keyword (**GUESSPATNC**, i.e. Guess "P" atomic non-collinear) allows to modify the above described atomic guess by allowing for a magnetization that is oriented along a selected direction (instead of only the z-axis) on different atomic centers.
The record NB specifies the number of distinct atoms for which the magnetization is modified.
LA contains the label of an atom whose electronic configuration is to be modified.
The polar RTHETA θ^a and azimuthal RPHI ϕ^a angles which orient the magnetization for atom a in the Cartesian frame are defined as follows.
Let $m^a(\mathbf{r})$ denote the magnitude of the scalar-relativistic magnetization calculated for atom a at point $\mathbf{r}$ in space. Then:

$$\mathbf{m}^a(\mathbf{r}) = m^a(\mathbf{r}) \left[\sin(\theta^a)\sin(\phi^a), \sin(\theta^a)\cos(\phi^a), \cos(\theta^a)\right]^T \; S^a \tag{6.11}$$

where S^a (inserted through the RSCALE record) is a scaling factor for the magnetization of atom a.

Some examples follow.

Example 1 Reorient the magnetization of all iodine atoms along the x-axis:

```
MOLECULE
1
2
253  1.335 0. 0.
253 -1.335 0. 0.
END
...
```

```
TWOCOMPON
GUESSPATNC
1
1 90.0 90.0 1.0
ENDTWO
...
```

Example 2 Reorient the magnetization of all iodine atoms along the y-axis and scale their magnetization in the starting guess by a factor of $1/2$:

```
MOLECULE
1
2
253  1.335 0. 0.
253 -1.335 0. 0.
END
...
TWOCOMPON
GUESSPATNC
1
1 90.0 0.0 0.5
ENDTWO
...
```

Example 3 Reorient the magnetization of all iodine atoms along the xyz diagonal:

```
MOLECULE
1
2
253  1.335 0. 0.
253 -1.335 0. 0.
END
...
TWOCOMPON
GUESSPATNC
1
1 54.7356 45.0 1.0
ENDTWO
...
```

In this example, let us clarify that the proper angles are found by writing:

$$\sin(\theta^a)\sin(\phi^a) = \sin(\theta^a)\cos(\phi^a) = \cos(\theta^a) \tag{6.12}$$

The value of ϕ^a is obtained from the first equality:

$$\begin{aligned} &\sin(\phi^a) = \cos(\phi^a) \\ &\phi^a = 45^\circ \end{aligned} \tag{6.13}$$

The value of θ^a can then be found from the second equality:

$$\begin{aligned} &\cos(\theta^a) = \sin(\theta^a)\cos(45^\circ) \\ &\theta^a = \operatorname{arccot}[\cos(45^\circ)] \sim 54.7356^\circ \end{aligned} \tag{6.14}$$

Example 4 Orient the magnetization of the first Cr atom along the $+x$ direction and the other Cr atom along the $-x$ direction:

```
MOLECULE
1
2
224 0. 0. 0.       Use different atomic numbers (224 and 324)
324 0. 0. 1.7      to permit a different atomic guess. Can
END                alternatively be done with keyword CHEMOD
...
TWOCOMPON
GUESSPATNC
2
1  90.0 90.0 1.0
2 270.0 90.0 1.0
ENDTWO
...
```

GUESSP[NOSO] - Guess from a previous 1c-SCF density matrix

Use a density matrix from a previous one-component calculation as a starting guess for the 2c-SCF. ***The following requirements must be met by the reference 1c-SCF calculation:***

- The 1c-SCF must be unrestricted (see keyword **UHF** at p. 126 or **SPIN** at p. 135)
- The 1c-SCF must have been performed without exploitation of space-group (or point-group) symmetry (see keyword **SYMMREMO** at p. 73)

GUESSPSO - Guess from a previous 2c-SCF density matrix

Use a density matrix from a previous two-component calculation as a starting guess for the 2c-SCF.

GUESSROTM - Rotate the guess magnetization from a previous 1c-SCF or 2c-SCF

	rec	variable	meaning
•	*	IROT	rotation of magnetization
			If IROT $\neq 0$ insert II
•	*	RTHETA,RPHI	real-valued polar and azimuthal angles to define orientation of magnetization

2NDVARIAT - Second-Variational Calculation

	rec	variable	meaning
•	*	IROT	rotation of magnetization
			If IROT $\neq$ 0 insert II
•	*	RTHETA,RPHI	real-valued polar and azimuthal angles to define orientation of magnetization

Perform a second-variational calculation for SOC, as initially proposed in Ref. [158] (provided that the keyword **SOC** is also given, see p. 170). That is to say, request to perform one cycle of the 2c-SCF, using a previously calculated scalar-relativistic density matrix (without SOC) as a starting guess.

The sequence:

```
...
TWOCOMPON
SOC
2NDVARIAT
0
END
...
```

is equivalent to the combination:

```
...
TWOCOMPON
SOC
GUESSPNOSO
PRTENESOC
END
FMIXING
0
MAXCYCLE
1
...
```

For performing the second-variational calculation, it is important to remember to not overwrite the requested values for FMIXING and MAXCYCLE (see pages 105 and 114).
A non-zero input value for IROT requests to globally rotate the magnetization from the z-axis to a desired orientation. This option can be used to estimate, for instance, the magnetic anisotropy energy. Refer to p. 166 for an explicit discussion on the definition of the polar RTHETA and azimuthal RPHI angles used in the definition of the global orientation of the magnetization vector.

The following provides an example input sequence making use of IROT to orient the guess magnetization along the x-axis:

```
...
TWOCOMPON
SOC
2NDVARIAT
1
90.0 90.0
END
...
```

and is equivalent to the following combination of keywords:

```
...
TWOCOMPON
SOC
GUESSPNOSO
GUESSROTM
90.0 90.0
PRTENESOC
END
FMIXING
0
MAXCYCLE
1
...
```

PRTENESOC - Detailed printing of energy contributions at each cycle of the 2c-SCF

This option provides a detailed printing of contributions to the total energy at each cycle of the self-consistent field procedure. Strictly speaking, such individual energy contributions have no physical meaning, but may be useful for numerical analysis.

An example output follows:

```
 ::: SPIN-ORBIT COUPLING PART OF ENERGY           -1.2238967134367E-02
 ::: SCALAR-RELATIVISTIC PART OF ENERGY           -2.2905922090883E+01
 ::: BIELET ZONE COULOMB & EXACT-EXCHANGE          3.0545080411043E+01
 ::: EXCHANGE ENERGY                              -4.2718032122717E+00
 ::: ELECTROSTATIC (MONO+COULOMB) ENERGY          -1.8169042370996E+01
 ::: CORRELATION ENERGY                           -4.6507650761531E-01
 ::: NUCLEAR-NUCLEAR REPULSION ENERGY              9.7114918377154E+00
 ::: 2-COMPONENT RELATIVISTIC TOTAL ENERGY        -2.2918161058017E+01
```

Note that the first line DOES NOT provide the true contribution of SOC to the total energy. This is because the inclusion of a SOC operator in the 2c-SCF changes the density matrix, which, in turn, also changes all other individual energy contributions.
Please note that the *only way* to obtain the SOC contribution to the total energy is to perform two calculations, one with the keyword **SOC** and one without the keyword **SOC**.

SOC - Include spin-orbit coupling in the Hamiltonian.

For a calculation with SOC, a SOREP operator must be provided as part of an effective core potential (ECP), either through internal libraries (e.g. the **STUTSC**, **STUTLC**, **STUTSH**, **COLUSC**, **COLULC**, **COLUSH** internal libraries of ECPs), or through manual input of the ECP, using the keyword **INPSOC** (see section 3.2)

SPINORLOCK - Fix number of occupied spinors for a preset number of cycles

	rec	variable	meaning
•	*	NSPIN	number of occupied spinors
•	*	NCYC	number of cycles the locking is maintained

Lock the number of occupied spinors to NSPIN for a preset number of cycles NCYC. This option can be used, for instance, to lock the number of occupied bands as being equal to the number of electrons in the unit cell, so as to avoid passing through a metallic state for the first NCYC cycles of the 2c-SCF procedure.
This keyword is similar to **SPINLOCK** used in the 1c-SCF (see p. 124). However, unlike in the 1c-SCF, it is not possible to define a number of occupied spin-up or spin-down bands, because spin is no longer a good quantum number to label the states in the 2c-SCF. Thus, rather than locking the number of up- or down- levels (as in **SPINLOCK**), this keyword simply locks the number of total levels.

6.3 Non-Collinear Density Functional Theory

For studying open-shell molecules and materials, it is nowadays common practice to employ the spin-DFT [281]. In this theory, the exchange-correlation energy functional is allowed to depend on the electron density ρ, as well as the spin or magnetization vector $\mathbf{m} = [m_x, m_y, m_z]$. Usually the magnetization is quantized along the z direction (i.e. the *"collinear"* approach) and the functional then reads $E_{xc} = E_{xc}[\rho, m_z]$. We also note that using the spin densities $\rho^{\uparrow}$ and $\rho^{\downarrow}$, as well as the relations $\rho = \rho^{\uparrow} + \rho^{\downarrow}$ and $m_z = \rho^{\uparrow} - \rho^{\downarrow}$, allows to write this same functional as $E_{xc} = E_{xc}\left[\rho^{\uparrow}, \rho^{\downarrow}\right]$.
The limitation for the functional to depend on only one Cartesian component of the magnetization is insufficient if the electronic state can be described by a Hamiltonian including an electron-electron *ee* interaction that depends explicitly on spin. This is the case of the so-called geometrically frustrated states of matter, occuring for instance, in face-centered cubic Fe, as well as certain Jarosite and Garnet minerals that exhibit the "Kagomé" lattice, with nearest neighbours in a triangular arrangement. But perhaps a more important example is for an open-shell system with strong spin-orbit coupling (SOC). In such cases, (*ee* interaction depending on spin), the total energy obtained from a $E_{xc} = E_{xc}[\rho, m_z]$ collinear calculation becomes dependent on the arbitrary orientation in space of the system w.r.t. the Cartesian frame. It is then clearly necessary to use instead an exchange-correlation energy formula like $E_{xc} = E_{xc}[\rho, m_x, m_y, m_z]$, to restore rotational invariance of the total energy. This is the idea of the non-collinear (NC) DFT approach. In CRYSTAL, the NC-DFT approach is useable for a calculation in the two-component spinor basis (see section 6.2). It may be activated by using one of two keywords **NONCOLC** or **NONCOLSF**, as in the following two examples:

```
...
99 0
END
...
TWOCOMPON
END
...
DFT
NONCOLC
PBE0
END
...
```

or:

```
...
99 0
END
...
TWOCOMPON
END
...
DFT
NONCOLSF
PBE0
END
...
```

The implementation is fully outlined in the following article:[66]

"Desmarais, J. K., Komorovsky, S., Flament, J. P., and Erba, A. (2021). Spin–orbit coupling from a two-component self-consistent approach. II. Non-collinear density functional theories. The Journal of Chemical Physics, 154(20), 204110."

In spin-DFT, the exchange-correlation potential has the general form:[281]

$$\hat{V} = \frac{\delta E_{\mathrm{DFA}}}{\delta \rho}\hat{\sigma}_0 + \sum_{a=x,y,z} \frac{\delta E_{\mathrm{DFA}}}{\delta m_a}\hat{\sigma}_a \tag{6.15}$$

where $\hat{\sigma}_0$ is a 2×2 unit matrix and $\hat{\sigma}_a$ are the 2×2 Pauli matrices.
This expression is simplified in the collinear approach:

$$\hat{V} \equiv \hat{V}^{col} = \frac{\delta E_{\mathrm{DFA}}}{\delta \rho}\hat{\sigma}_0 + \frac{\delta E_{\mathrm{DFA}}}{\delta m_z}\hat{\sigma}_z \tag{6.16}$$

where $\hat{V}^{col}$ is block diagonal, that is to say $\hat{V}^{\sigma\sigma'} = \delta_{\sigma\sigma'}\hat{V}^{\sigma\sigma}$. In the LDA:[66]

$$\hat{V}^{\sigma\sigma} = \frac{\partial \varepsilon_{\mathrm{LDA}}}{\partial \rho^{\sigma}} \tag{6.17}$$

in which $\varepsilon_{\mathrm{LDA}}$ is the LDA energy-density, $\rho^{\uparrow} = \frac{1}{2}(\rho + m_z)$ and $\rho^{\downarrow} = \frac{1}{2}(\rho - m_z)$.
In the GGA, it turns out to be convenient to not work directly with $\hat{V}^{\sigma\sigma}$, but rather with its matrix elements $V^{\sigma\sigma}_{\mu\nu}(\mathbf{g})$ in an AO basis, which read:[66]

$$V^{\sigma\sigma}_{\mu\nu}(\mathbf{g}) = \int \frac{\partial \varepsilon_{\mathrm{GGA}}}{\partial \rho^{\sigma}} \varrho^{\mathbf{g}}_{\mu\nu} d\mathbf{r} + \int \left[2\frac{\partial \varepsilon_{\mathrm{GGA}}}{\partial |\boldsymbol{\nabla}\rho^{\sigma}|^2} \boldsymbol{\nabla}\rho^{\sigma} + \frac{\partial \varepsilon_{\mathrm{GGA}}}{\partial \gamma^{\mathrm{col}}_{+-}} \boldsymbol{\nabla}\rho^{\sigma'} \right] \cdot \boldsymbol{\nabla}\varrho^{\mathbf{g}}_{\mu\nu} d\mathbf{r} \tag{6.18}$$

where $\sigma \neq \sigma'$ and the overlap distribution $\varrho^{\mathbf{g}}_{\mu\nu}$ reads:

$$\varrho^{\mathbf{g}}_{\mu\nu}(\mathbf{r}) = \chi_{\mu}(\mathbf{r} - \mathbf{a}_{\mu})\,\chi_{\nu}(\mathbf{r} - \mathbf{g} - \mathbf{a}_{\nu}) \tag{6.19}$$

and χ_{ν} is an AO centered at $\mathbf{a}_{\nu}$ in cell $\mathbf{g}$.
In Eq. (6.18), $\gamma^{\mathrm{col}}_{+-}$ are defined as:

$$\gamma^{\mathrm{col}}_{+-} = \frac{1}{4}\boldsymbol{\nabla}[\rho + m_z] \cdot \boldsymbol{\nabla}[\rho - m_z]$$

For the *non-collinear* approach, both the canonical **NONCOLC**, as well as the more elaborate Scalmani and Frisch **NONCOLSF** approaches are available. These two approaches coincide in the LDA, but differ in the GGA. In the NC LDA case, the exchange-correlation potential reads:

$$\hat{V} = \hat{V}^{\mathrm{ncol}} = \frac{\partial \varepsilon_{\mathrm{LDA}}}{\partial \rho}\hat{\sigma}_0 + \sum_{a=x,y,z} \frac{\partial \varepsilon_{\mathrm{LDA}}}{\partial m_a}\hat{\sigma}_a \tag{6.20}$$

while, in the GGA, it is again convenient to work with matrix elements of $\hat{V}$, which may be expressed as follows, in terms of quantities defined in Eq. (6.15). For the functional derivative w.r.t. ρ:[66]

$$\begin{aligned}\langle\mu^{\mathbf{0}}|\frac{\delta E_{\mathrm{GGA}}}{\delta\rho}|\nu^{\mathbf{g}}\rangle &= \frac{1}{2}\int\left[\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\rho_+}+\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\rho_-}\right]\varrho^{\mathbf{g}}_{\mu\nu}d\mathbf{r}+\frac{1}{2}\int\Big[2\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{can}}_{++}}\boldsymbol{\nabla}\rho_+ + 2\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{can}}_{--}}\boldsymbol{\nabla}\rho_- \\ &+ \frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{can}}_{+-}}\left(\boldsymbol{\nabla}\rho_+-\boldsymbol{\nabla}\rho_-\right)\Big]\cdot\boldsymbol{\nabla}\varrho^{\mathbf{g}}_{\mu\nu}d\mathbf{r} \qquad (6.21a)\end{aligned}$$

and for the functional derivative w.r.t. magnetization components:[66]

$$\begin{aligned}\langle\mu^{\mathbf{0}}|\frac{\delta E_{\mathrm{GGA}}}{\delta m_a}|\nu^{\mathbf{g}}\rangle &\approx \frac{1}{2}\int\frac{m_a}{m}\left[\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\rho_+}-\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\rho_-}\right]\varrho^{\mathbf{g}}_{\mu\nu}d\mathbf{r}+\frac{1}{2}\int\frac{m_a}{m}\Big[2\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{can}}_{++}}\boldsymbol{\nabla}\rho_+ - 2\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{can}}_{--}}\boldsymbol{\nabla}\rho_- \\ &- \frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{can}}_{+-}}\left(\boldsymbol{\nabla}\rho_+-\boldsymbol{\nabla}\rho_-\right)\Big]\cdot\boldsymbol{\nabla}\varrho^{\mathbf{g}}_{\mu\nu}d\mathbf{r} \qquad (6.21b)\end{aligned}$$

where we have introduced the following auxiliary semi-local quantities proper to the canonical non-collinear formulation:

$$\gamma^{\mathrm{can}}_{\pm\pm}=\frac{1}{4}\boldsymbol{\nabla}\left[\rho\pm m\right]\cdot\boldsymbol{\nabla}\left[\rho\pm m\right] \qquad (6.22a)$$

and:

$$\rho_\pm=\frac{1}{2}\left(\rho\pm m\right) \qquad (6.22b)$$

in which $m=\sqrt{m_x^2+m_y^2+m_z^2}$ is the modulus of the magnetization vector. In Eq. (6.21b), an approximated equal sign has been used, because contributions originating from the gradient of m_a/m have been dropped, which corresponds to assuming that the gradient of the magnetization locally follows the direction of the magnetization itself.
In the alternate SF formulation (keyword **NONCOLSF**), the functional derivatives are calculated as:[66]

$$\begin{aligned}\langle\mu^{\mathbf{0}}|\frac{\delta E_{\mathrm{GGA}}}{\delta\rho}|\nu^{\mathbf{g}}\rangle &= \frac{1}{2}\int\left[\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\rho_+}+\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\rho_-}\right]\varrho^{\mathbf{g}}_{\mu\nu}d\mathbf{r}+\frac{1}{2}\int\Bigg[\left(\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{sf}}_{++}}+\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{sf}}_{--}}+\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{sf}}_{+-}}\right)\boldsymbol{\nabla}\rho \\ &+ \left(\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{sf}}_{++}}-\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{sf}}_{--}}\right)f_\nabla\left\{(\boldsymbol{\nabla}\rho\cdot\boldsymbol{\nabla}\mathbf{m})\circ(\boldsymbol{\nabla}\rho\cdot\boldsymbol{\nabla}\mathbf{m})\right\}^{-\frac{1}{2}}(\boldsymbol{\nabla}\rho\cdot\boldsymbol{\nabla}\mathbf{m})\circ\boldsymbol{\nabla}\mathbf{m}\Bigg]\cdot\boldsymbol{\nabla}\varrho^{\mathbf{g}}_{\mu\nu}d\mathbf{r} \qquad (6.23a)\end{aligned}$$

and:

$$\begin{aligned}\langle\mu^{\mathbf{0}}|\frac{\delta E_{\mathrm{GGA}}}{\delta m_c}|\nu^{\mathbf{g}}\rangle &\approx \frac{1}{2}\int\frac{m_c}{m}\left[\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\rho_+}-\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\rho_-}\right]\varrho^{\mathbf{g}}_{\mu\nu}d\mathbf{r}+\frac{1}{2}\int\Bigg[\left(\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{sf}}_{++}}+\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{sf}}_{--}}-\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{sf}}_{+-}}\right)\boldsymbol{\nabla}m_c \\ &+ \left(\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{sf}}_{++}}-\frac{\partial\varepsilon_{\mathrm{GGA}}}{\partial\gamma^{\mathrm{sf}}_{--}}\right)f_\nabla\left\{(\boldsymbol{\nabla}\rho\cdot\boldsymbol{\nabla}\mathbf{m})\circ(\boldsymbol{\nabla}\rho\cdot\boldsymbol{\nabla}\mathbf{m})\right\}^{-\frac{1}{2}}(\boldsymbol{\nabla}\rho\cdot\boldsymbol{\nabla}m_c)\,\boldsymbol{\nabla}\rho\Bigg]\cdot\boldsymbol{\nabla}\varrho^{\mathbf{g}}_{\mu\nu}d\mathbf{r} \qquad (6.23b)\end{aligned}$$

where the product $\cdot$ is understood to act on the components of $\boldsymbol{\nabla}$, while the product $\circ$ acts on the components of $\mathbf{m}$, and we have used the following quantities proper to the SF non-collinear formulation:

$$\gamma^{\mathrm{sf}}_{++}\text{ or }\gamma^{\mathrm{sf}}_{--} = \frac{1}{4}\left[\boldsymbol{\nabla}\rho\cdot\boldsymbol{\nabla}\rho+\boldsymbol{\nabla}\mathbf{m}\cdot\circ\boldsymbol{\nabla}\mathbf{m}\right]\pm\frac{f_\nabla}{2}\left\{(\boldsymbol{\nabla}\rho\cdot\boldsymbol{\nabla}\mathbf{m})\circ(\boldsymbol{\nabla}\rho\cdot\boldsymbol{\nabla}\mathbf{m})\right\}^{\frac{1}{2}} \qquad (6.24a)$$

along with:

$$\gamma^{\mathrm{sf}}_{+-}=\frac{1}{4}\left[\boldsymbol{\nabla}\rho\cdot\boldsymbol{\nabla}\rho-\boldsymbol{\nabla}\mathbf{m}\cdot\circ\boldsymbol{\nabla}\mathbf{m}\right] \qquad (6.24b)$$

in which:

$$f_\nabla=\mathrm{sgn}\left[\boldsymbol{\nabla}\rho\cdot(\boldsymbol{\nabla}\mathbf{m})\circ\mathbf{m}\right] \qquad (6.24c)$$

where sgn is the signum function [66]. In Eq. (6.23b), an approximated equal sign is used, because terms originating from the gradient of f_∇ are dropped.

A screening algorithm has been developped for dealing with those terms in Eqs. (6.21b) and (6.23b), that contain m_a/m, which cannot be numerically evaluated at those points on the integration grid where m is vanishing. This procedure, described in Ref. [[66]] has been documented to provide rotational invariance of the total energy down to an order of at least 1×10^{-09} E_h for both the canonical and SF non-collinear formulations, employing GGA functionals, with SOC.
Briefly, the screening procedure evaluates m_a/m explicitly, only if at least two Cartesian components of **m** exceed (in absolute value) a preset tolerance `TOLM` (the default value being 1×10^{-14} a.u.). If, instead only one Cartesian component exceeds `TOLM` in absolute value, then the procedure reduces to the collinear problem along the corresponding component. Finally, if none of the Cartesian components of **m** exceed `TOLM`, then terms proportional to the magnetization itself are set to zero in Eq. (6.21b) or Eq. (6.23b), whereas terms proportional to the gradient of the magnetization are calculated using a value $m_a/m \to \langle m_a/m \rangle$ which is averaged over the atomic basin in which the relevant point in the DFT grid is situated [66].
We recall the following keywords which are available for the **DFT** block with a two-component spinor basis:

[SC]DFT	Spin-Current DFT Hamiltonian	126	–
SDFT	Spin DFT Hamiltonian	177	–
	Choice of the exchange-correlation functionals		
	EXCHANGE exchange functional	129	I
	LDA functionals		
	LDA Dirac-Slater [68] (LDA)		
	VBH von Barth-Hedin [281] (LDA)		
	GGA functionals		
	BECKE Becke 1988[18] (GGA)		
	PBE Perdew-Becke-Ernzerhof 1996 [204] (GGA)		
	PBESOL GGA. PBE functional revised for solids [206]		
	mPW91 modified Perdew-Wang 91 (GGA)		
	PWGGA Perdew-Wang 91 (GGA)		
	SOGGA second order GGA. [302]		
	WCGGA GGA - Wu-Cohen [293]		
	CORRELAT correlation functional	129	I
	LDA functionals		
	PZ Perdew-Zunger [210] (LDA)		
	VBH von Barth-Hedin [281] (LDA)		
	VWN Vosko,-Wilk-Nusair [282] (LDA)		
	GGA functionals		
	LYP Lee-Yang-Parr [167] (GGA)		
	P86 Perdew 86 [202] (GGA)		
	PBE Perdew-Becke-Ernzerhof [204] (GGA)		
	PBESOL GGA. PBE functional revised for solids [206]		
	PWGGA Perdew-Wang 91 (GGA)		
	PWLSD Perdew-Wang 92 [208, 209, 207] (GGA)		
	WL GGA - Wilson-Levy [291]		
	Standalone keywords: exchange+correlation		
	SVWN see [68, 282]	129	
	BLYP see [18, 167]	129	
	PBEXC see [204]	129	
	PBESOLXC see [206]	129	
	SOGGAXC see [302]	129	
	Global Hybrid functionals		
	Standalone keywords		
	B3PW B3PW parameterization	130	–
	B3LYP B3LYP parameterization	130	–
	PBE0 Adamo and Barone [5]	130	
	PBESOL0 Derived from PBE0	130	
	B1WC see [25]	130	
	WC1LYP see [60]	130	
	B97H see [4, 101]	130	
	PBE0-13 see [42]	130	
	mPW1PW91 see [6]	130	
	mPW1K see [175]	130	
	User defined global hybrids		
	HYBRID hybrid mixing	130	I
	Spin(-current) DFT Formulation		
	[**COLLINEAR**] Collinear spin(-current) dft (Default)	–	–
	NONCOLC Canonical non-collinear spin(-current) dft	176	–
	NONCOLSF Scalmani-Frisch non-collinear spin(-current) dft	176	–
	TOLM Developers only	177	I

Numerical accuracy control			
ANGULAR	definition of angular grid	137	I
RADIAL	definition of radial grid	137	I
[**BECKE**]	selection of Becke weights (default)	137	–
SAVIN	selection of Savin weights	137	–
OLDGRID	"old" default grid	140	
LGRID	"large" predefined grid	140	
[**XLGRID**]	"extra large" predefined grid (new default)		
XXLGRID	"extra extra large" predefined grid	140	
RADSAFE	safety radius for grid point screening		I
TOLLDENS	density contribution screening 6	141	I
TOLLGRID	grid points screening 14	141	I
[**BATCHPNT**]	grid point grouping for integration	141	I
CHUNKS	max n. points in a batch for numerical int.	142	I
DISTGRID	distribution of DFT grid across nodes	142	
LIMBEK	size of local arrays for integration weights 400	142	I
Atomic parameters control			
RADIUS	customized atomic radius	143	I
FCHARGE	customized formal atomic charge	143	I
Auxiliary			
END	close DFT input block		

NONCOLC - Canonical Non-Collinear Spin(-current) DFT Formulation

This keyword activates the canonical non-collinear spin(-current) DFT formulation for LDA or GGA functionals [66]. This option is only available with unpruned (i.e. uniform) integration grids, see section 4.3.
Activation of the keyword also sets the default integration grid as:

```
...
DFT
RADIAL
1
4.0
75
ANGULAR
1
9999.0
16
END
...
```

i.e. an unpruned grid with 75 radial points and a Lebedev accuracy level of 16, corresponding to 974 angular points - see p. 138. The unpruned grid may be modified by changing these two input values.

NONCOLSF - Scalmani-Frisch Non-Collinear Spin(-current) DFT Formulation

This keyword activates the Scalmani-Frisch (SF) non-collinear spin(-current) DFT formulation

for LDA or GGA functionals. It is equivalent to **NONCOLC** (canonical formulation) for LDA functionals, but differs in the case of GGA functionals [66, 252]. In the GGA case, the difference is that the SF formulation allows for a local magnetic torque from the exchange-correlation potential (being instead always vanishing with the canonical formulation for non-hybrid functionals) and is also slightly more numerically stable. The local magnetic torque refers to the ability of the exchange-correlation potential to locally rotate the magnetization along the self-consistent field process.
As with **NONCOLC**, this option is only available with uniform (unpruned) integration grids, see section 4.3. The default grid is the same as for the canonical formulation.

TOLM - Developer's option

	rec	variable	meaning
•	*	8 × TOL	eight integer tolerances for magnetization screening

Tolerances for magnetization screening in the algorithm of Ref. [66].
Default values are:

```
...
DFT
TOLM
14 14 14 14 14 14 14 14
END
...
```

SDFT - Request to Perform a Spin DFT (Rather than Spin-current DFT) Calculation

This keyword opens an input block that must be closed by the keyword **END**.
For density functional theory calculations with hybrid exchange-correlation functionals in a two-component spinor basis (i.e. employing the **TWOCOMPON** block, see p. 165), in CRYSTAL it is possible to perform the calculation with two different formulations for exact-exchange.
The first, and most accurate strategy (activated by default with the keyword **DFT** or **SCDFT**, see p. 128) follows an adiabatic-connection exchange-correlation energy E_{xc}^{SCDFT} formula according to the spin-current DFT formalism [227, 64]:

$$E_{xc}^{\mathrm{SCDFT}}\left[\rho, m_x, m_y, m_z, \mathbf{j}, \mathbf{J}^x, \mathbf{J}^y, \mathbf{J}^z\right] = \alpha\, E_{F}^{\mathrm{SCDFT}}\left[\rho, m_x, m_y, m_z, \mathbf{j}, \mathbf{J}^x, \mathbf{J}^y, \mathbf{J}^z\right] + \int d\mathbf{r}\; \varepsilon_{\text{LDA or GGA}}\left[\rho, m_x, m_y, m_z\right] \qquad (6.25)$$

where α is the fraction of Fock exchange (with energy E_F^{SCDFT}) and $\varepsilon_{\text{LDA or GGA}}$ is the exchange-correlation energy-density in an LDA or GGA treatment. In Eq. (6.25), ρ is the particle-number density (electron density), $\mathbf{m}$ is the magnetization, $\mathbf{j}$ is the particle current and $\mathbf{J}^x, \mathbf{J}^y, \mathbf{J}^z$ are the three spin currents.
The second possible strategy, activated by the present keyword, uses instead an analogous formula from the spin DFT formalism [29]:

$$E_{xc}^{\mathrm{SDFT}}\left[\rho, m_x, m_y, m_z\right] = \alpha\, E_{F}^{\mathrm{SDFT}}\left[\rho, m_x, m_y, m_z\right] + \int d\mathbf{r}\; \varepsilon_{\text{LDA or GGA}}\left[\rho, m_x, m_y, m_z\right] \qquad (6.26)$$

A comparison of predictions from Eqs. (6.25) and (6.26) (using the keywords **DFT** or **SCDFT** vs. **SDFT**) allows to directly discuss the effect of including current densities in the electron-electron interaction on the electronic structure. See Ref. [29] for a more ample discussion and example applications.

An example input follows:

```
...
99 0          end of basis set input
END
...
TWOCOMPON     open block to request a calculation
END           in the 2c-spinor basis
...
SDFT          open this block with SDFT to deactivate contributions from current densities
PBE0          request, for example, to use the PBE0 hybrid functional
NONCOLC       request, for example, to use the canonical non-collinear formulation
END           close the SDFT block
...
```

6.4 Properties from the 2c-SCF Solution

One-electron properties may be computed from the 2c-SCF solution by running the program **properties**, as outlined in chapter 14. In the present 2c-SCF implementation, the following keywords are supported:

Preliminary calculations			
NEWK	Eigenvectors calculation	341	I
Properties computed from the density matrix			
BAND	Band structure	304	I
ECHG	Charge density and charge density gradient - 2D grid	319	I
ECH3	Charge density - 3D grid	317	I
PPAN	Mulliken population analysis	117	
PROPS2COMP	Specify which densities to compute on a grid of points	179	I
Properties computed from eigenvectors (after keyword NEWK)			
ANBD	Printing of principal AO component of selected CO	302	I
BWIDTH	Printing of bandwidth	311	I
DOSS	Density of states	315	I
New properties			
TOPO	Topological analysis of the electron density	361	I
Auxiliary and control keywords			

ADDSPACE	add an extra space delimiter in the fort.25 for densities on a grid	179	–
ANGSTROM	Set input unit of measure to Ångstrom	35	–
BOHR	Set input unit of measure to bohr	38	–
END	Terminate processing of properties input keywords		–
FRACTION	Set input unit of measure to fractional	47	–
MAPNET	Generation of coordinates of grid points on a plane	339	I
MAXNEIGHB	maximum number of equidistant neighbours from an atom	49	I
NEIGHBOR	Number of neighbours to analyse in **PPAN**	59	I
RAYCOV	Modification of atomic covalent radii	64	I
SETINF	Setting of inf array options	66	I
STOP	Execution stops immediately	68	–
SYMMOPS	Printing of point symmetry operators	73	–

Info - Output of data on external units

ATOMIRR	Coordinates of the irreducible atoms in the cell	303	–
ATOMSYMM	Printing of point symmetry at the atomic positions	38	–
COORPRT	Coordinates of all the atoms in the cell	42	–
EXTPRT	Explicit structural/symmetry information	44	–

ADDSPACE - Insert extra space delimiter for data written on the 2D grid in file fort.25

This keyword adds an extra space delimiter between each density datum calculated on the 2D grid in the formatted fort.25 output from an **ECHG** calculation (see p. 319). This keyword may be useful when plotting magnetization $\boldsymbol{m}$, particle-current $\boldsymbol{j}$ and spin-current $\boldsymbol{J}^x$, $\boldsymbol{J}^y$, and $\boldsymbol{J}^z$ densities (see **PROPS2COMP** at p. 179). The extra space may prove useful in subsequent data reading and processing with negative values of the densities.

PROPS2COMP - Specification of density variables to be computed on a 2D or 3D grid

rec	variable	meaning
•	*A* keyword	enter a keyword to choose the type of property
•	**DENSITY**	
•	**MAGNETIZ**	
•	**ORBCURDENS**	
•	**SPICURDENS**	
•	**GRID3D**	
		if **GRID3D** is specified, insert ECH3 input records II
•	**END**	end of **PROPS2COMP** input block

This keyword opens a block that must be closed with the keyword **END**. It specifies which density variables are to be computed from the 2c-SCF solution in a 2D or 3D grid of points, by a subsequent call to **ECHG** or **ECH3**, respectively, see p. 319 and 317. For the case of a 3D grid of points, it is also possible to insert the corresponding input records directly after the keyword **GRID3D**, instead of using a subsequent call to **ECH3**.
The possible density variables that may be plotted are those of the spin-current density-functional theory of Vignale and Rasolt [279] as generalized to SOC by Bencheikh [24].
Namely, they are:

1. The electron or particle-number density, using the keyword **DENSITY**:

$$\rho(\boldsymbol{r}) = \sum_i \int_{\Omega_F} d\boldsymbol{k}\ \boldsymbol{\psi}_{i\boldsymbol{k}}^{\text{Pauli}\dagger}(\boldsymbol{r})\, \boldsymbol{\psi}_{i\boldsymbol{k}}^{\text{Pauli}}(\boldsymbol{r})\ , \tag{6.27}$$

where Ω_F is the subvolume inside the first Brillouin zone associated with band-structure energies $\epsilon_{i\boldsymbol{k}}$ below the Fermi level.

2. The particle-current (the currents of the electrons) or orbital-current density, using the keyword **ORBCURDENS**:

$$\boldsymbol{j}(\boldsymbol{r}) = \frac{1}{2\imath} \sum_i \int_{\Omega_F} d\boldsymbol{k}\ \boldsymbol{\psi}_{i\boldsymbol{k}}^{\text{Pauli}\dagger}(\boldsymbol{r}) \left[\boldsymbol{\nabla} \boldsymbol{\psi}_{i\boldsymbol{k}}^{\text{Pauli}}(\boldsymbol{r})\right] + h.c.\ , \tag{6.28}$$

where $h.c.$ denotes the Hermitian conjugate and $\imath = \sqrt{-1}$.

3. The magnetization or spin-density, using the keyword **MAGNETIZ**:

$$m_a(\boldsymbol{r}) = \sum_i \int_{\Omega_F} d\boldsymbol{k}\ \boldsymbol{\psi}_{i\boldsymbol{k}}^{\text{Pauli}\dagger}(\boldsymbol{r})\, \hat{\sigma}^a \boldsymbol{\psi}_{i\boldsymbol{k}}^{\text{Pauli}}(\boldsymbol{r}) \quad \forall a = x, y, z\ , \tag{6.29}$$

4. And finally, the spin-current densities (the currents of the spin), using the keyword **SPICURDENS**:

$$\boldsymbol{J}^a(\boldsymbol{r}) = \frac{1}{2\imath} \sum_i \int_{\Omega_F} d\boldsymbol{k}\ \boldsymbol{\psi}_{i\boldsymbol{k}}^{\text{Pauli}\dagger}(\boldsymbol{r})\, \hat{\sigma}^a \left[\boldsymbol{\nabla} \boldsymbol{\psi}_{i\boldsymbol{k}}^{\text{Pauli}}(\boldsymbol{r})\right] + h.c. \quad \forall a = x, y, z\ . \tag{6.30}$$

ρ is, of course, non-vanishing for all systems that contain electrons. $\boldsymbol{J}^x, \boldsymbol{J}^y, \boldsymbol{J}^z$ are non-vanishing for systems with SOC. m_x, m_y, m_z, as well as $\boldsymbol{j}$, on the other hand, are only non-vanishing for systems with broken time-reversal symmetry (open-shell systems) [275, 64, 29]. See keywords **ECHG** and **ECH3** at p. 319 and 317 for an explanation of the data format written on output for calculations on 2D and 3D grids, respectively.

Some example input decks follow.

Example 1 Calculate all possible variables for a time-reversal symmetry breaking polymer with SOC on a 3D grid extending 5 Bohr on each side of the origin in the non-periodic y and z directions:

```
PROPS2COMP
DENSITY              request to calculate the electron density
MAGNETIZ             request to calculate the magnetization
ORBCURDENS           request to calculate the orbital current density
SPICURDENS           request to calculate the three spin current densities
GRID3D               request to directly insert the ECH3 input record
25                   25 points along the periodic x direction
RANGE
-5 -5                y,z directions are from -5 to 5 Bohr
 5  5
END                  close props2comp block
END                  end of properties input
```

and is equivalent to the following input deck:

```
PROPS2COMP
DENSITY
MAGNETIZ
ORBCURDENS
SPICURDENS
END
```

```
ECH3
25
RANGE
-5 -5
 5  5
END
```

Example 2 Calculte the electron density ρ and the three spin current densities $\boldsymbol{J}^x$, $\boldsymbol{J}^y$ and $\boldsymbol{J}^z$ in a time-reversal symmetry preserving TaAs crystal with SOC on a 2D grid of points:

```
PROPS2COMP
DENSITY            request to calculate the electron density
SPICURDENS         request to calculate the spin current densities
END                close props2comp block
PPAN               Also Mullkien populations, why not? they're free!
ADDSPACE           Add an extra space as a delimiter in the fort.25 file
ECHG               Call ECHG for specifications on how to plot rho and J^x, J^y, J^z in 2D
0                  0th order derivatives of rho and J^x, J^y, J^z are calculated
500                five hundred points along the B-A segment
COORDINA           the plane of the plot is defined by the three points A, B, C
1 0 0              point A
0 0 0              point B
0 0 1              point C
MARGINS            enlarge the window of the plot by 5 Angstrom along all directions
5. 5. 5. 5.
END                end of "MAPNET" input block
END                end of properties input
```

Chapter 7

Geometry Optimization

Geometry optimization at constant symmetry is invoked by the keyword **OPTGEOM** in input block 1 (geometry). OPTGEOM must be the last keyword in geometry input. OPTGEOM input block admits several options (sub keywords), and terminates with keyword **END** (or END[OPT], END[—]: the first three characters only are processed).

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 either symmetrized fractional coordinates with [default] and without cell parameters, or redundant internal coordinates (optional choice, page 192).

OPTGEOM sub keywords can be classified as follow:

1. *General sub keywords*:
   - A - Optimization type (page 186)
   - B - Initial Hessian (page 186)
   - C - Hessian updating technique (page 184)
   - D - Convergence criteria (page 183)
   - E - Step control (page 188)
   - F - Coordinate system related options (page 189)
   - G - Optimization procedure control (page 189)
   - H - Numerical (first) derivatives (page 191)
   - I - Printing options (page 192)
2. *Geometry optimization in redundant coordinates* (page 192).
3. *Geometry optimization with constraints* (page 196).
   - *A* - Constant volume optimization (page 197)
   - *B* - Fixing lattice deformations (page 197)
   - *C* - Linear constraints between atomic coordinates (page 199)
   - *D* - Partial optimization of atomic positions (page 200)
   - *E* - Fixing internal coordinates (page 200)
4. *Geometry optimization with application of an external stress* (page 202)
5. *Searching a transition state (page 203)*

Default values are supplied for all computational parameters.

By default a full geometry optimization (atomic positions and cell) is performed. This was not the case with previous versions of the program, where atomic positions only were optimized by default.

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).

7.1 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 default choice for the initial Hessian matrix is obtained from a model Hessian, as proposed by Schlegel and updated by using the BFGS algorithm[38, 39, 109, 122, 259].

By default the step considered is the Newton step (direction and length) controlled by the Trust Radius scheme (see **ALLOWTRUSTR** page 188). **NOTRUSTR** to remove trust radius control (CRYSTAL06 default choice).

HF and DFT (pure and hybrid functionals) analytical gradients for atomic positions and cell parameters, 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** (page 121) is used. In that case, *numerical* first derivatives should be computed (see page 191). For very small value of smearing (around 0.001 hartree) analytical gradients can be used.

For atomic positions (**ATOMONLY** option), 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 64) 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.

To optimize the lattice parameters a set of symmetry preserving cell deformations (see Symmetry Allowed Elastic Distortions, **USESAED**, page 73) related to changes of isotropic volume and of axial ratios is defined. 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 conveniently normalized combined set of symmetrized directions and deformations is adopted.

Optional choice (keyword **INTREDUN**, page 192) 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.

7.2 Default setting

7.2.1 Type of optimization

The default geometry optimization type is the relaxation of both the nuclear coordinates and the lattice parameters. Optional choices: see page 186.

7.2.2 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 page. 203), 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

Optimization convergence criteria are set to different values according to the context where geometry optimization is performed.

	RMS on gradient	RMS on displacement
Standard geometry opt	0.0003	0.0012
preopt in frequency calculation	0.00003	0.00012
preopt in EOS	0.00006	0.00012
preopt in elastic constants	0.00006	0.00012

7.2.3 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 Es (element 1 to 99).

Warning - To define bonds the sum of covalent radii (see page 64) is used. For ionic systems it may be necessary to modify the default values (see **RAYCOV**, page 64).

7.2.4 Hessian updating technique

BFGS Broyden-Fletcher-Goldfarb-Shanno scheme [38, 39, 109, 122, 259].

Optional choices:

1. Schlegel's updating scheme [254], (**OLDCG**, page 187), optimization scheme as in CRYSTAL03
2. Powell's updating scheme (**POWELL**, page 187)

7.2.5 SCF convergence and guess

The default value for SCF convergence criterion on total energy is set to 10^{-7} (TOLDEE in input block 3 to modify it: never reduce accuracy).

After the first step, at each SCF cycle, the density matrix is recovered from the previous geometry optimization step (implicit GUESSP, page 109 option).

This choice may be modified by inserting the keyword **NOGUESS**. A superposition of atomic densities is then adopted on each step as SCF initial guess.

If the SCF solution at a given optimization step 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 may 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**).

7.2.6 Output files

The following formatted files 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 to file fort.33 in a simple *xyz* format (see Appendix D, page 444). This file is suitable to be read by molecular graphics programs (e.g. Molden...) to display the animation of a geometry optimization run.

fort.34 If optimization is successful, the last geometry in written in file fort.34 (format described in Appendix D, page 446).

The file can be read to define the basic geometry input. See **EXTERNAL**, page 16

opta(c)xxx 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 D, page 446). The file must be renamed "fort.34" if used to enter geometry input (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, page 191).

HESSOPT.DAT The hessian matrix is written, and can be read to define the initial guess for the Hessian (keyword **HESSOPT**) in geometry optimization of a system with same geometry and symmetry (it may have different BS, Hamiltonian, computational parameters).

SCFOUT.LOG SCF and optimization process printout is routed to file SCFOUT.LOG after the first cycle. Keyword **ONELOG**: full printing on standard output.

7.3 General sub-keywords

A number of optional keywords allow tuning of the optimization procedure.

A - Type of optimization (page 186)

B - Initial Hessian (page 186)

C - Hessian updating technique (page 187)

D - Convergence criteria (page 183)

E - Step control (page 188)

F - Coordinate system related options (page 189)

G - Optimization procedure control (page 189)

H - Numerical first derivatives (page 191)

I - Printing options (page 192)

7.3.1 Type of optimization

Optional choices:

ATOMONLY	Only atomic coordinates are optimized. This was the default before CRYSTAL14
FULLOPTG	full optimization, atom coordinates and cell parameters (default in CRYSTAL14). The cell volume may change (see **CVOLOPT**, page 197, to optimize at constant volume)
CELLONLY	only cell parameters are optimized. Default: the cell volume may change (see **CVOLOPT**, page 197, to optimize at constant volume)
ITATOCEL	full optimization, iterative procedure optimization: atoms-cell-atoms-cell-. . . .
INTREDUN	full optimization of atomic positions and cell parameters in redundant internal coordinates (page 192).

7.3.2 Initial Hessian

By default an estimated model Hessian is adopted. The Hessian matrix is stored in file HESSOPT.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 (HESSNUM), but this process can be very expensive.

HESSFREQ	initial guess for the hessian - input from file HESSFREQ.DAT obtained from frequencies calculations (developers only)
HESSIDEN	initial guess: identity matrix
HESSOPT	external guess (read from file HESSOPT.DAT)
HESSMOD1	initial guess: Lindh's model Hessian [173]

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 [255, 292] [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 Es.

Warning - To define bonds the sum of covalent radii (see page 64) is used. For ionic systems it may be necessary to modify the default values (see **IONRAD** in what follows in order to use tabulated ionic radii in place of the covalent ones provided by default or **RAYCOV**, page 64, for customizing the value of covalent radii, page 64).

HESSNUM	initial guess: numerical estimate

7.3.3 Tabulated atomic radii

The construction of the set of internal coordinates which is used in the definition of the Schlegel model Hessian relies on tabulated values for the atomic radii, according to which bonds, angles and dihedrals are defined. Two different tables can be choosen in input with the following keywords:

COVRAD	The covalent radii table is used in the construction of the internal coordinates. This is the default choice, which is reasonable in most cases but may lead to a ill defined Schlegel model Hessian in some compact ionic structures.
IONRAD	The ionic radii table is used in the construction of the internal coordinates. This leads to a better initial Schlegel model Hessian in ionic compounds.

7.3.4 Hessian updating technique

Different Hessian updating schemes are available for minimization:

BFGS	Hessian update - Broyden-Fletcher-Goldfarb-Shanno scheme [38, 39, 109, 122, 259] - [default]
OLDCG	Hessian updating - old Schlegel updating scheme[254] (CRYSTAL03)
BERNY	Synonym
POWELL	Hessian update - symmetric Powell scheme [231]

7.3.5 Optimization 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)

7.3.6 Step control

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

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

Trust Radius [default]
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 run CRYSTAL06 as CRYSTAL09 the keyword **ALLOWTRUSTR** must be specified along with **BFGS**.

To run CRYSTAL09 as CRYSTAL06, the keyword **NOTRUSTR** must be specified in geometry optimization input

Related keywords are discussed below:

ALLOWTRUSTR activate the trust radius technique to control the step size [0.5 for geometry optimization; 0.1 for transition state search] [default]

The step at each cycle is computed by means of a Newton-like scheme, in case it exceeds trust radius it is re-scaled using the scheme proposed by Simmons and Nichols [Simmons, J., and Nichols, J.: , Int. J. Quantum Chem. Symp. 24, volume 24, 263, (1990)] (see also page 203).

MAXTRADIUS optional
• ∗ TRMAX maximum value allowed for the trust radius

This is useful in transition state optimizations or in minimizations along flat potential surfaces in order to avoid too large displacements from one point to the next one. Default value: geometry optimization: 4.0 ; transition states search: 0.3.

NOTRUSTR not using trust radius to limit displacement

TRUSTRADIUS
• ∗ TRADIUS set the initial value for trust radius

Set the initial value of the trust radius to [TRADIUS]. The trust radius limitates the step length of the displacement at each cycle. The value is updated at each optimization point by analysis of the local quadraticity of the potential energy function.
Default: geometry optimization 0.5; transition state search 0.1.

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, written to file optc(a)xxx, being xxx the optimization cycle number.

7.3.7 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 (page 197) - 3D only.
FRACTION	optimization in fractional coordinates
FRACTIOO	optimization in normalized fractional 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 constraints])
RENOSAED	renormalize symmetry allowed deformations [default when FULLOPTG is requested]

7.3.8 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 202. **Note that in the previous versions of the program the default choice was 0. Now it is 4.**

- ∗ ICODE Action type code:
  - 0 the program stops
  - 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
  - 4 step 3 is iterated until full stable optimization (default)

FIXDELTE

- ∗ IE 10^{-ie} hartree: threshold on the total energy change for redefining the geometry to which integral classification is referred - see **FIXINDEX**, page 104 - [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]

FITDEGR

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

HESEVLIM limits for the allowed region of hessian eigenvalues (hartree)

- ∗ VMIN lower limit [default 0.001]
- VMAX upper limit [default 1000.]

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]

N.B. When optimization is restarted, the first restarted optimization cycle number is the last of the previous run + 1. Set MAXCYCLE value accordingly.

NOGUESS SCF guess at each geometry point: superposition of atomic densities at each SCF calculation (default choice in geometry optimization: **GUESSP**

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.
See page 203

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

7.3.9 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 conducting systems and the **SMEAR** option - page 121) and therefore to optimize the atoms coordinates.

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

NUMGRALL geometry optimization - numerical atomic and cell gradient

---

NUMGRATO geometry optimization - numerical atomic gradients

NUMGRCEL geometry optimization - numerical cell gradients

STEPSIZE modify step for numerical gradient [default 0.001 au] (developers only)
- ∗ STEP new stepsize value

7.3.10 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

2 - Optimization in redundant internal coordinates

INTREDUN	geometry optimization in valence internal redundant coordinates

An optimization in redundant internal coordinates is performed when the keyword **INTREDUN** is inserted in **OPTGEOM** input block.

Optional keywords related to geometry optimization in redundant internal coordinates must follow.

A symmetrized set of internal coordinates (i.e. bonds, angles and torsions) is automatically defined, which contains much 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 64) is used. For ionic systems the default values can be automatically set by using the **IONRAD** directive (see pag. 187), or explicitly setting them with the **RAYCOV** keyword (page 64). In case of systems constituted by unconnected fragments (*ie* some molecular crystals or adsorption complexes), fragments are automatically linked to each other by *pseudo* "bond lengths" between the closest pair of atoms belonging to each fragment.

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.

By default, optimization of internal redundant coordinates involves both atomic positions and cell parameters. To avoid optimizing cell parameters the keyword **FIXCELL** page 196 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.

INTLMIXED geometry optimization in selected valence internal mixed with or embedded in a full set of general symmetrized fractional and cell parameters

---

The optimization is performed in a mixed coordinate system made up of a set of selected valence internal parameters (bond length, angles and dihedrals) and the full set of symmetry adapted fractional displacements and elastic distortions. As in the previous case, the coordinate system is redundant but the number of parameters generated is in general substantially smaller than in the case of **INTREDUN**. Therefore, it is convenient for calculations with a large number of atoms in which the automatic generation of valence internal parameters makes the dimension of the optimization parameters space huge. It is also useful when the lack of connectivity of the system causes *quasi*-linear dependencies reflected in a very high condition number of the Wilson B-Matrix. The latter fact usually makes the optimization either to fail or to exhibit an erratic behavior. It is recommended in transition state optimizations of large systems, in particular when freezing selected valence parameters, **SCANREDU**, or **PATHFOLLOW**/**FITTOPATH** options.
This option does not generate automatically valence internal coordinates unless the user explicitly indicate atomic connections through keywords **DEFLNGS** and **DEFANGLS**. The tree is automatically completed generating the necessary lengths, angles or dihedrals. For the back-transformation to the non-redundant set, the **BKTRNSF2** algorithm is considered. This option is not compatible with **DBANGLIST**, page 194.

Managing with almost linear angles

Linear or almost linear angles (i.e. close to 180°) 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. Double angles are defined as a couple of angles obtained by projection of the vectors onto two suitable perpendicular planes. This avoids the indetermination around 180°. 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°]

The default value is set to 165°. This means that all angles larger than 165° are automatically split into two.
This option can be required, 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°.

A list of angles 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

This keyword provides the list of angles chosen to be converted in double angles (i.e. defined by the angles obtained by projection of the vectors onto two suitable perpendicular planes) in order to avoid the indetermination around 180 degrees). This option is not compatible with **INTLMIXED**, page 193 . The labels used for the angles are those provided by a previous automatic generation of internal coordinates computed in a test run (**TESTREDU** keyword).

Double angles have to be defined at the starting of the optimization. If any single angle approaches 180° the program stops with a message. For this reason, it is strongly recommended foresee **before** the optimization which are the angles that may evolve to close to 180° degree and protect them making them explicitly double. A less recommendable alternative is to set an **ANGTODOUBLE** value very small (< 90°) so as to make double any eventual tricky angle. Such a procedure requires less effort to the user but must be used with caution as the number of angles and dihedrals based on them may explode.

Explicitly defining internal coordinates - bonds and angles

When some relevant internal coordinates are missing (e.g. intermolecular bonds) they can be explicitly defined 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

Choosing method for back-transformation

By default CRYSTAL employs an iterative method (P. Pulay and G. Fogarasi, "Geometry optimization in redundant internal coordinates", *J. Chem. Phys. 96*, 2856 (1992)) so as to compute the atomic positions and cell parameters corresponding to the set of redundant internal coordinates obtained at each optimization point. For those periodic systems in which

the generated internal coordinates have a high degree of redundancy and/or the displacement is relatively large (for instance in the starting points of the optimization when geometry is very far from the target one) such a method may fail for large steps and provide inaccurate displacements making the whole optimization less efficient. As an alternative you might change the strategy for back-transformation using the keyword **BKTRNSF2**

Under such an option, the program performs back-transformations from redundant (internal) to fractional+cell coordinate systems through an alternative procedure based on the conjugate gradient (CG) algorithm. The idea of the method is as follows. Given a reference set of redundant parameters, the best point in terms of atomic positions and lattice parameters is the one that is the closest as possible to the reference one in the redundant parameter space. The distance function is defined as a weighted mean square of the differences betwen redundant parameters of the trial and reference sets. This function is fastly minimized by the CG strategy.

The main drawback is that the new method is slightly more costly in cpu time and memory space than the default iterative one.

In all tests done, this new scheme for the back-transformation works better than the iterative procedure originally proposed by Pulay for molecules and extended in Crystal to periodic systems:

1. It is safer in its convergence; specially when the displacement step in terms of redundant coordinates is very large and the structure displays a large number of connectivity loops (this happens in most natural crystals). In such cases the "old" method fails and the program performs a rough conversion that is often very poor in accuracy.
2. As concerns the optimization process, the use of this more accurate alterative allows in most cases to save a few points leading to an overall gain in computational time, even if the conjugate gradient scheme is a bit more costly than the simple iterative procedure.

An additional advantage of the algorithm is that it allows to set weights to the squared deviations of the redundant parameters used in the deviation function so as to force some of them to be better approched than the others in the back-transformation procedure. This permits a more controlled definition of the steps in terms of internal redundant coordinates. The choice is optative in the **MODINTCOOR** option, and during the optimization it can be set by means of the keyword **WGHTDREDU**.

WGHTDREDU Assign weights for back-transformation

- ∗ NMODI — Number of internal coordinates to be given an specific weight
  DEFWGTH — Default weight for internal coordinates not explicitly defined

- ∗ NRED(I), WEIGHT(I), I=1,NMODI — Label of coordinate (in the list of internal redundants coordinates produced in output with directive **TESTREDU**) and specific weight.

Specific weights may be used to provide priorities when different internal coordinates are in conflict in the back-transformation, as the corresponding redundant set of values does not correspond to any real atomic arrangement. This typically occurs for instance when the connectivity graph exhibits a high degree of nested loops.

Modifying geometry before the optimization through internal coordinates

Geometry modification in terms of internal coordinates. This option allows to modify the value of any internal coordinate. The keyword is set in the **OPTGEOM** block (together with **INTREDUN/INTLMIXED** and the syntaxis is as follows:

MODINTCOOR Modification of internal coordinates
(**ADJUSTGEO)**

- $*$ NMODI — Number of internal coordinates to be modified
  IWGHT > 0 Weight given to the parameters to be modified using **BKTRNSF2** as back-transformation scheme. The min value (1) means the new parameters will be approached trying to move as less as possible the values of the remaining ones in the redundant set. The max value is 1000 and means the new value of the chosen parameters will be as close as possible to the new provided values while the rest of redundant parameters will accomodate to allow this.
  < 0 Use the old back-transformation; no weight is actually assigned.
- $*$ NRED(I), VALNEW(I), I=1,NMODI — Label (in the list of internal redundants coordinates produced in output with directive **TESTREDU**) and new value of the internal coordinates to be modified.

Together with **TESTREDU** it allows to perform geometry modifications in terms of internal coordinates without performing any optimization. In such a case, a file called "optc000" is written in the execution directory that contains the modified geometry in external format. Conveniently renamed, this file can be used with the keyword **EXTERNAL** (page 16) as starting geometry for a new calculation.

Other optional keywords

FIXCELL keep cell parameters fixed in internal coordinates optimization

---

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})

.

---

3 - Geometry 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 (page 197)

B - Fixing lattice deformations (page 197)

C - Linear constraints between atomic coordinates (page 199)

D - Partial optimization of atomic positions (page 200)

E - Fixing internal coordinates (page 200)

Constraining strategies *A*-*D* 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** page 192).
The examples in 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), **FULLOPTG** (atom coordinates and cell parameters optimization) or **INTREDUN** (redundant coordinates 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 E vs V curves point-by-point 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 - 3D only
• ∗	NFIXC	number of constraints relating pairs of cell deformations
		insert NFIXC records
• ∗	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** (p. 189, 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.

For instance, the following symmetry-allowed elastic deformations reported in the output correspond to the sequence: 1 a, 2 γ, 3 b, 4 β, 5 α, 6 c.

```
SYMMETRY ALLOWED ELASTIC DISTORTION 1
 1.0000000  0.0000000  0.0000000
 0.0000000  0.0000000  0.0000000
 0.0000000  0.0000000  0.0000000

SYMMETRY ALLOWED ELASTIC DISTORTION 2
 0.0000000  0.7071068  0.0000000
 0.7071068  0.0000000  0.0000000
 0.0000000  0.0000000  0.0000000

SYMMETRY ALLOWED ELASTIC DISTORTION 3
 0.0000000  0.0000000  0.0000000
 0.0000000  1.0000000  0.0000000
 0.0000000  0.0000000  0.0000000

SYMMETRY ALLOWED ELASTIC DISTORTION 4
 0.0000000  0.0000000  0.7071068
 0.0000000  0.0000000  0.0000000
 0.7071068  0.0000000  0.0000000

SYMMETRY ALLOWED ELASTIC DISTORTION 5
 0.0000000  0.0000000  0.0000000
 0.0000000  0.0000000  0.7071068
 0.0000000  0.7071068  0.0000000

SYMMETRY ALLOWED ELASTIC DISTORTION 6
 0.0000000  0.0000000  0.0000000
 0.0000000  0.0000000  0.0000000
 0.0000000  0.0000000  1.0000000
```

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 (synonym **ATOMFREE**). 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 16.4, page 380.

E - Fixing internal coordinates

Constraints on internal coordinates can be easily imposed during geometry optimization.
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 printing internal coordinates).

Note that for a better control over the selected frozen internal coordinates we suggest using the keywords **LNGSFROZEN** and **ANGSFROZEN**. These options reorder the internal coordinates list so as the frozen parameters to be the first ones in the lengths and angles entries of the output. The frozen coordinates will appear with the label (T) in the FRZ column.

Constrained 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 behavior in the optimization process.

The constraining of internal coordinates is performed with numerical techniques (particularly in the back-transformation from redundant internal to cell/atomic coordinate systems) and the fixed values may be affected by some small changes (in general of the order of 10^{-4} au). The use of **BKTRNSF2** (page 194) may improve the numerical behaviour.

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.

FREEZDIH	freeze a list of dihedral (active with **INTREDUN** only):
• ∗ NDH	number of dihedrals to be frozen
• ∗ IFR(I), I=1,NDH)	list of dihedrals to be frozen

The list of dihedrals, to choose the ones to be frozen, is obtained performing a previous run with the keyword **TESTREDU** into **OPTGEOM** input block (the program stops after the

printing of the internal coordinates, see page 196).

After using **FREEZINT** (page 200) it turns out that the order of the dihedral angles in the output changes: the frozen parameters appear at the beginning of the list. All of them are now labeled as frozen ("FRZ=T") in the printed list of redundant coordinates.

4 - Geometry optimization with application of an external stress

Geometry optimization can be done under an hydrostatic pressure (EXTPRESS).

EXTPRESS	to apply external, hydrostatic pressure
• ∗ pres	pressure in hartree/bohr^3

Input example:

```
EXTPRESS
0.001    hydrostatic pressure of hartree/bohr^3 s applied
```

The directive should be introduced in any place within the **OPTGEOM** block.

7.4 Notes on geometry optimization

7.4.1 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 18.12), depends on the geometry of a crystal. With default 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 four 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.
4. step 3 is iterated until full stable optimization

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=4 in FINALRUN input). This is the default in *crystal*.

7.4.2 Optimization of flat surfaces

Flat regions of surfaces often 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 using 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.

7.4.3 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.
The geometry at each step is written to file optc_number_of_step, and can be read by EXTERNAL (see page 16).
If optimization ended successfully, the optimized geometry is written to file fort.34 (EXTERNAL format).

When restarting an optimization, information on previous optimization steps is read from file OPTINFO.DAT. Optimization then proceeds, saving information in the same file at each step.

The SCF guess, read from file fort.20, is the density matrix written in file fort.9 at the end of SCF of 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

CRYSTAL output is read by the software MOLDRAW:
http://www.moldraw.unito.it to visualize the optimization process.

File fort.33 contains the geometry at each step, in *xyz* format.

7.5 Searching a transition state

TSOPT transition state search requested [default: false]

Transition state optimization is invoked by the keyword **TSOPT** in **OPTGEOM** input block.

The reference to be quoted is:

C. M. Zicovich-Wilson, M. L. San-Roman, A. Ramirez-Solis,
Mechanism of F^- Elimination from Zeolitic D4R Units: A Periodic B3LYP Study on the Octadecasil Zeolite
J. Phys. Chem. C 114 (2010) 2989-2995.

An example of transition state search is presented in http://www.crystal.unito.it/tutorials

By default the eigenvector that corresponds to the lowest eigenvalue is followed uphill according to the scheme proposed by Simmons and Nichols [Simmons, J., and Nichols, J.: , Int. J. Quantum Chem. Symp. 24, volume 24, 263, (1990)].
To adopt other choices for directions to be followed uphill see keywords: **MODEFOLLOW**, **PATHFOLLOW**, **FITTOPATH** and **CHNGTSFOL**.

7.5.1 Transition state control keywords

MODEFOLLOW
- ∗ MODEFOL mode to follow

ABS(MODEFOL) is the label of the eigenvector to be followed uphill initially, namely DIR(0).

If MODEFOL < 0, the initial uphill direction, DIR(0), is the opposite to that of the eigenvector of label ABS(MODEFOL)

In a general optimization step, NSTEP, the current uphill direction DIR(NSTEP) is chosen as the hessian eigenvector of maximum overlap with the direction chosen in the previous step, DIR(NSTEP-1). In this scheme the uphill direction is allowed to smoothly change along the optimization. Some problems might appear when there are quasi-degeneracies between the Hessian eigenvalue of the uphill direction and other that corresponds to a direction to be followed downhill. In such a case the optimization might go in troubles. Using **PATHFOLLOW** is a safer way to define the uphill direction and so the reaction path.

PATHFOLLOW only with redundant internal coords
- ∗ NPATHFOL max coord to choose the mode to follow

Only valid together with **INTREDUN**/**INTLMIXED**.

The uphill direction is the eigenvector that has maximal absolute contribution of the internal valence coordinate labeled ABS(NPATHFOL), which is thus supposed to dominate the reaction path.

If NPATHFOL < 0, the uphill search is such that the value of coordinate ABS(NPATHFOL) decreases along the reaction coordinate, otherwise the opposite direction is chosen.

At variance with the MODEFOLLOW case, where the reference direction changes from step to step, here the same strategy is employed in every step of the optimization. This prevents troubles when near-degeneracies occur (see keyword **MODEFOLLOW**).

FITTOPATH only with redundant internal coords
- ∗ NPATHFOL2 integer
  NPATHWEIGHT integer

Only valid together with **INTREDUN**/**INTLMIXED** and **PATHFOLLOW**

ABS(NPATHFOL2) is the label of a second internal valence coordinate, namely coordinate (II), that together with the one labeled ABS(NPATHFOL), coordinate (I), mostly contributes to the reaction coordinate.

Once the eigenvector with maximum contribution of coordinate (I), namely XMAXCONTR, is

obtained (see **PATHFOLLOW** keyword), the eigenvectors are once more scanned and those having a contribution of this coordinate larger in absolute value than XMAXCONTR*(100-NPATHWEIGHT)/100 selected. If NPATHFOL2 > 0 the previously selected eigenvector with maximum contribution of coordinate (II) with the same sign of XMAXCNTR is chosen as uphill direction. Otherwise, the one with maximum contribution having opposite sign is considered.

CHNGTSFOL only with redundant internal coords

Valid together with **INTREDUN/INTLMIXED**, **PATHFOLLOW (FITTOPATH)**.

The optimization follows uphill the path according to the **PATHFOLLOW (+FITTOPATH)** scheme while the chosen eigenvector is not the first one in the list ordered by increasing eigenvalues, i.e. it has not the lowest eigenvalue. Once such a situation reverts, the scheme changes to **MODEFOLLOW** in the following steps and the uphill direction is chosen according to the criterion of maximum overlap with the previous uphill direction (see **MODEFOLLOW** keyword).

7.5.2 Scan keywords

SCANATOM

• ∗	NATSCAN	(integer): label of the atom to be scanned
	TARGET	(real array dim 3): last position of the atom in the scan
	MSCAN	number of steps in which the previous displacement is carried out

Only for P1 structures.

Perform a series of optimizations in which one atom is kept fixed at different contiguous positions and the remainder of the structure fully or partially relaxed.

Compatible with the optimization of atomic positions (default) and atoms+cell (**FULLOPTG**).

Not compatible with **INTREDUN/INTLMIXED**. *The atom chosen for scan and any other one with the only condition that belongs to that part of the system chosen to remain associated to the center of mass, must be defined as fixed with the* **FRAGMENT** *option.*

This directive may be used associated with **FRAGSCAN**. The keyword has the same syntax as **FRAGMENT** and defines a set of atoms that after each step of the scan have the same displacement of the scanned atom, NATSCAN, so as to set the geometry of the starting point of the next optimization of the sequence.

SCANREDU		To be used with **INTREDUN/INTLMIXED** only
• ∗	IREDSCA	(integer): type of valence coordinate to be scanned (1, bond length; 2 angle; 3 dihedral)
	ENDSCA	(real): last value taken by the chosen coordinate along the scan The initial value is the current one with the geometry defined for the structure.
	MAXSCA	number of points considered in the scan

Perform a series of optimizations (scan) in which one (or two) redundant valence internal coordinate(s) are kept fixed at different values while the remainder are fully relaxed.

To be used only with **INTREDUN/INTLMIXED**.

The directive SCANREDU must be accompanied with freezing the redundant valence internal coordinate(s) one wants to scan (see keywords **FREEZINT** (page 200), **FREEZDIH** (page 201), **ANGSFROZEN** (page 200), **LNGSFROZEN** (page 200).

According to the order of the coordinates given in the output (see **INTREDUN, TESTREDU**) the scan is performed on the last frozen coordinate (indicated in the output by a "T") of type *IREDSCA*. If **SCANREDU** is requested twice with the same *IREDSCA*, the second time it refers to the last but one frozen coordinate of type *IREDSCA*.

SCANREDU may be requested at maximum twice so as to carry out a bidimensional scan.

The use of **BKTRNSF2** (see pag. 194) may improve the accuracy in the displacements between two consecutive optimizations of the scan, particularly when they are rather large.

Chapter 8

Lattice Dynamics - Vibration Frequencies

8.1 Calculation of Harmonic Vibration Frequencies

The calculation of the vibration harmonic frequencies is invoked by the keyword **FREQCALC** in input block 1 (geometry). **FREQCALC** must be the last keyword in the geometry input block. **FREQCALC** admits several options (subkeywords) listed below and terminates with keyword **END** (or END[FREQ], END[—]: only the first three characters are processed). References to be quoted when using this module:

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

Besides harmonic vibration frequency calculation at the Γ-point [default] it allows:

1. Calculation of infrared (IR) intensities (**INTENS**, page 215).
2. Calculation of Raman intensities (**INTRAMAN**, page 217).
3. Scanning of geometry along selected normal modes (**SCANMODE**, page 220).
4. Calculation of the infrared spectra (**IRSPEC**, page 223).
5. Calculation of the Raman spectra (**RAMSPEC**, page 225).
6. Phonon dispersion at the harmonic level (**DISPERSION**, page 227).
7. Calculation of the atomic Anisotropic Displacement Parameters (ADP) at any temperature (**ADP**, page 229).
8. Calculation of total and projected vibrational density-of-states (**PDOS**, page 231) and neutron-weighted PDOS for comparison with Inelastic Neutron Scattering spectra (**INS**, page 231).
9. Calculation of the vibrational contribution to Second-Harmonic Generation (SHG) and Pockels tensors (**BETAVIB**, page 219).
10. Thermodynamic properties at different temperatures and pressures (see page 213).

The second derivatives of the energy are computed numerically by using the analytical first derivatives (gradients). Frequencies are obtained by diagonalizing the mass-weighted Hessian in Cartesian coordinates.

8.1.1 Symmetry exploitation

The point group symmetry of the lattice is used to reduce the number of SCF+gradient calculations to be performed. At each point, the residual symmetry is exploited for the SCF calculation.
Second derivatives calculations are done on the irreducible atoms only. The full Hessian matrix is then generated by applying the point group symmetry to the irreducible part. The mass-weighted Hessian matrix is diagonalized to obtain eigenvalues, which are converted in frequencies (cmISOTROPIC^{-1}), and eigenvectors, i.e. the normal modes.

8.1.2 Equilibrium Geometry

The first step to compute frequencies is the calculation of the wave function at the equilibrium geometry. The geometry of the system **must** correspond to a stationary point on the potential energy surface.
Geometry optimization can be either performed on a previous run (recommended) or controlled by two sub-keywords of **FREQCALC**:

NOOPTGEOM do not perform previous optimization in **FREQCALC** job [default]

PREOPTGEOM perform optimization of atomic positions before starting the vibrational modes calculation. For a full relaxation of both lattice parameters and stomic postions the keywork **FULLOPTG** should be inserted within the **PREOPTGEOM** block

insert **OPTGEOM** keywords (close with **END**)

keyword **END**

An input block is opened that must finish with keyword **END**. The numerical conditions for the optimization are controlled by means of the same keywords as documented in page 182 (keyword **OPTGEOM**).
The conditions adopted by default in geometry optimization before frequency calculation are different than those considered for normal optimizations in order to obtain much more accurate minima numerical second derivatives. This ensures a good accuracy in the computation of the frequencies and modes. The defaults are:

```
TOLDEG          0.00003
TOLDEX          0.00012
FINALRUN        4
MAXTRADIUS      0.25
TRUSTRADIUS     .TRUE.
```

If frequency calculation is restarted (keyword **RESTART**, page 214) the input geometry must be the final optimized geometry found by **PREOPTGEOM**.

8.1.3 Default settings

The SCF wave-function starting guess for the calculations of all the displaced geometries necessary to compute the numerical second derivatives of the total energy is the density matrix obtained at the equilibrium geometry.

The default value for SCF convergence criterion on total energy is set to 10^{-9} (use the **TOLDEE** keyword in input block 3 to modify it).
The default choice for DFT grid, when a DFT Hamiltonian is used, corresponds to **XLGRID** (page 140).
The calculation of longitudinal-optical (LO) frequencies and infrared (IR) intensities is not performed by default. If the **INTENS** (page 215) keyword is used, intensities are evaluated.

The **FREQCALC** input block must be closed by the keyword **END** (or **ENDFREQ**). All additional keywords (see below) to be put in between are optional.

8.1.4 Output files

Here is a list of the files that are written during a frequency calculation, to be saved in order 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 for the displaced configurations 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.

HESSFREQ.DAT Formatted. Contains the Hessian in Cartesian coordinates to be read by **HESSFREQ** in geometry optimization input block.

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

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

fort.28 Binary. Data for restart of IR intensities calculation through Berry phase approach.

fort.80 Binary. Localized Wannier functions, computed only if IR intensities are computed through Wannier functions.

8.1.5 Optional keywords

This is a list of some possible (optional) subkeywords to be inserted after **FREQCALC**. This is just a partial list since the main subkeywords are described into details in the following sections of this chapter:

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

- A **CHI2TENS** to be used if **INTRAMAN** is active (pages 215,217). Reads the second-order anisotropic dielectric tensor for the calculation of the LO Raman intensities.
- $*$ TENS(1:27) second-order dielectric tensor matrix TENS (3x9 elements, input by rows: 27 reals (3D)).
  The tensor elements can be computed using the **CPHF** keyword (see page 265) at **THIRD** or **FOURTH** order (page 10.3). The tensor elements should be inserted in the following order:

```
xxx xxy xxz
xyx xyy xyz
xzx xzy xzz
yxx yxy yxz
yyx yyy yyz
yzx yzy yzz
zxx zxy zxz
zyx zyy zyz
zzx zzy zzz
```

- A **COMBMODE** Evaluation of transverse optical combination modes and overtones at the Γ point. A set of options are available, which are described below. This keyword opens an input block that must be closed by **END.**
- A .. Optional sub-keywords
- A **END**

List of optional sub-keywords

- A **IR** Only infrared (IR) active combination modes and overtones are displayed in the output.
- A **RAMAN** Only Raman active combination modes and overtones are displayed in the output.
- A **IRRAMAN** IR and Raman active combination modes and overtones are displayed in the output. [default]
- A **ALL** All vibrational combination modes and overtones are displayed in the output.
- A **FREQ** Combination modes and overtones are sorted according to the frequency value. [default]
- A **IRREP** Combination modes and overtones are sorted according to the IRREP.
- A **FREQRANGE** Only combination modes and overtones with frequency included within FMIN and FMAX are displayed in the output.
- ∗ FMIN,FMAX Boundaries of the frequency range. Default: [0., 1.3 ν_{max}].

- A **DIELISO** to be used if **INTENS** is active (page 215). Reads the isotropic diagonal dielectric tensor (dielectric constant) for the calculation of the LO/TO splitting. The dielectric constant has to be computed on a previous run with options **CPHF** (page 265) or **SUPERCEL** - **FIELD** and **DIEL** (see page 44) applied for each axis of the system.
- ∗ DIEL dielectric constant

- A **DIELTENS** to be used if **INTENS** is active (page 215). Reads the anisotropic dielectric tensor for the calculation of the LO/TO splitting. In some low-symmetry systems the polarization of LO modes is not determined by symmetry. In this case, LO frequencies are computed only if symmetry analysis of the vibrational modes is inactive (see keyword NOKSYMDISP). If **INTRAMAN is active, the keyword CHI2TENS should also be declared.**
- ∗ TENS(1:9) Dielectric tensor matrix TENS (3x3 elements, input by rows: 9 reals (3D).
  The dielectric tensor elements can be obtained from the literature or computed using the **CPHF** keyword (see page 265) or the **SUPERCEL - FIELD** keywords (see page 44).

- A **ECKART** Eckart conditions imposed to project out of the Hessian matrix purely translational and rotational degrees of freedom (see page 411 for more details). **Note that this option is now active by default. This was not the case in the previous versions**

•	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 a fragment contains symmetry related atoms, 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 **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** in **FREQCALC** input block, and supplying the file FREQINFO.DAT written by the previous run.

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

- A **MULTITASK** This keyword allows the simultaneous execution of independent SCF+G calculations. It is aimed at the exploitation of the large processor counts available on High Performance Computing facilities. Allocation of too many processors results in performance degradation due to the communication overhead between the processors. When many independent tasks are usually performed sequentially, MULTITASK performs N tasks at the same time, so that an N-fold increase in the number of allocated processors can be requested, without losing performance.
- $*$ N Number of tasks. For example if N=8 is requested and the job is running on a total of NCPU=128 cores, 8 tasks with 16 cores each will be created. N should be greater than one and smaller than the total number of cores NCPU ($1<N\leq NCPU$). N should be a divider of NCPU, so that each task will run on the same amount of processors.

- A **NEGLEFRE** Reads the number of lowest vibration frequencies to be neglected in the computation of the thermodynamical properties. By default, the (rotational)+translational degrees of freedom are automatically neglected. This keyword allows to neglect further soft vibrations with low frequencies which may carry numerical issues on the computation of thermodynamical properties.
- $*$ N Number of frequencies to be neglected

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

- A **NOECKART** Eckart conditions not imposed to the Hessian.

- 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 **NOUSESYMM** Symmetry is removed, the space group is set to P_1

• 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 (one displacement for each atom along each cartesian direction) 2 Central-difference formula: h(x)=(g(x+t)-g(x-t))/2t t=0.001 Å (two displacements for each atom along each cartesian direction)

• 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 (MPa) [1,0.101325,0.101325]

• A **PRINT**	Extended printing active (Hessian and Hessian eigenvectors)

• A **RAMANEXP**	Takes into account experimental conditions (temperature, incoming laser) in the calculation of Raman intensities, according to Eq. 8.9 (see page 218).
• ∗ T,FREQ	2 reals, T is the temperature, FREQ is the frequency (in nm) of the incoming laser.

• A **RESTART**	Restart frequency calculation from a previous run. See page 214.

• A **STEPSIZE**	Modify the step size of displacement along each cartesian axis
• ∗ STEP	step (Å) for numerical derivatives [0.003]

• 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 **FREQSCAL**	Perform a scaling of the harmonic frequencies by a factor equal to SCALING
• ∗ SCALING	Scaling factor for harmonic frequencies calculation, reasonable values can be taken from [256]

• A **TEST[FREQ]**	Frequency test run

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

Partition of the modes into Building Unit contributions

The external portion of the motion of the b-th Building Unit (BU) in mode i is quantified by means of

$$\epsilon_{bi} = \frac{\left(\mathbf{P}^{(b)}\mathbf{e}_i\right)^2}{N_{bi}}, \tag{8.1}$$

where matrix $\mathbf{P}^{(b)} \equiv P^{(b)}_{A\alpha,B\beta}$, with atoms A, B belonging to the b-th BU, is the projector onto the roto-translational degrees of freedom of the unit taken as an isolated fragment and $N_{bi} = \sum_{A\in b}\sum_{\alpha=1}^{3} e^2_{i,A\alpha}$ is a normalization factor. Accordingly, the corresponding internal contribution is given by $1 - \epsilon_{bi}$.
The contribution of the b-th BU to mode i is computed as

$$\xi_{bi} = \sum_{\alpha=1}^{3}\sum_{A\in b} \frac{e^2_{i,A\alpha}}{m_A}, \tag{8.2}$$

where m_A is the number of BUs to which atom A belongs. If the mode vector is normalized Eq. (8.2) ensures that $\sum_b \xi_{bi}$ gives the portion of the mode covered by the partition (1 if a full partition is considered). By adopting such partitions, the internal and external contributions (per cent) of the b-th BU to mode i are given by $\Gamma^{(n)}_{bi} = (1 - \epsilon_{bi}) \times \xi_{bi} \times 100$ and $\Gamma^{(x)}_{bi} = \epsilon_{bi} \times \xi_{bi} \times 100$, respectively. In some cases one may be interested in considering the overall external contribution under a given BU partition. This is given by

$$\Xi_i = \frac{\sum_b \Gamma^{(x)}_{bi}}{\sum_b \xi_{bi}}, \tag{8.3}$$

where $\sum_i \Xi_i$ is not 100%, but the percentage of the structure covered by the BUs considered in the partition.
Though this analysis may be somehow arbitrary, the resulting indices provide a systematic and clear description of most of the significant features of the vibrational modes of the system under study.

The keyword **BUNITSDECO** performs a building unit decomposition of the vibrational modes. The vibrational modes are decomposed in terms of internal and external motions of some units defined by input. The latter correspond to rotations and translations of the units behaving like rigid, while the former to the relative motions of the constitutive atoms.

BUNITSDECO	perform a building unit decomposition of the vibrational modes.
• ∗ NBDNGUNIT	number of building units irreducible by symmetry considered (the units symmetry-equivalent are automatically generated)
• ∗ MBDNGUNIT(I), I=1,NBDNGUNIT	number of atoms of each unit. The sum defines NATOMS, the total number of atoms considered
• ∗ (LBDNGUNIT(JA,IU), JA=1,4),IU=1,NATOMS	identification of the atoms: for each atom, the sequence number and three cell indexes, in the order given in MBDNGUNIT

8.2 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 formatted - information on the part of the hessian already computed.

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

fort.28 (binary) Data for restart of IR intensities calculation through Berry phase approach.

fort.80 (binary) localized Wannier functions (if IR intensities through Wannier cunctions are computed).

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

IR intensities calculation using Berry phase or Wannier functions must be present in the first frequency calculation, it can not be inserted in restart only. It is however possible to do so with CPHF analytical intensities.
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** 211).

8.3 IR Intensities

Calculation of IR intensities is invoked by the keyword **INTENS** in **FREQCALC** input block.

Three different techniques can be adopted:

- IR intensities through Berry phase - keyword **INTPOL** [default]
- IR intensities through Wannier functions - keyword **INTLOC**
- IR intensities through CPHF approach - keyword **INTCPHF**

The first two approaches imply numerical differentiations, while the latter is entirely analytical. Atomic Born tensors are the key quantities for the calculation of the IR intensities, the oscillator strengths, the LO/TO splitting and the static dielectric tensor. Such quantities are written in the external formatted unit BORN.DAT.
In order to compute the LO/TO splitting, the high frequency dielectric tensor must be provided. See keyword **DIELTENS**, page 210.

The integrated IR intensity I_p for the $p-th$ mode is computed according to:

$$I_p = \frac{\pi}{3}\frac{N_A}{c^2}\cdot d_p \cdot |\vec{Z}_p|^2 \tag{8.4}$$

where N_A is the Avogadro's number, c is the speed of light, d_p is the degeneracy of the mode, $\vec{Z}_p$ is the mass-weighted effective mode Born charge vector.
The oscillator strength tensor f_p for the $p-th$ mode is computed according to:

$$f_{p,ij} = \frac{4\pi}{\Omega}\frac{\vec{Z}_{p,i}\vec{Z}_{p,j}}{\nu_p^2} \tag{8.5}$$

here Ω is the unit cell volume, ν_p is the mode TO frequency and $\vec{Z}_p$ is again the mass-weighted effective mode Born charge vector.

8.3.1 IR intensities through Berry phase [default]

Calculation of IR intensities through Berry Phase approach, keyword **INTPOL**, is the default choice.

This is possible for 3D, 2D, 1D and 0D systems, but only for insulating system.

The Berry phase approach consists in evaluating the atomic Born tensors, that is the derivative of the dipole moment with respect to the atomic displacements, as polarization differences between the central and the distorted geometries: the polarization difference is then equal to the time-integrated transient macroscopic current that flows through the insulating sample

during the vibrations.

The scheme operates on the crystalline-orbital eigenfunctions in the reciprocal space. As a consequence of that, the accuracy of IR intensities might be sensitive to the density of the Monkhorst net.

There are no additional keywords related to this method.

8.3.2 IR intensities through Wannier functions

Calculation of IR intensities through Wannier functions is invoked by the keyword **INTLOC** following **INTENS** in FREQCALC input block.

Many keywords are related to the Wannier functions calculation that should be used by developers and very experienced users only.

If keyword **INTLOC** is activated in **FREQCALC** input block, IR intensities, atomic Born tensors and LO/TO splitting are evaluated through the Wannier functions, obtained by localizing the crystalline orbitals. This is possible for insulators only.

IR intensities calculation through localization is very demanding, in terms of memory allocation. **NOINTENS**, default choice, avoids intensity calculation, when not necessary.

As regards the computation of the IR intensities, they are obtained by means of the Wannier Function (WF) 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. The system must be an insulator. By default the dipole moment in the non central points are computed with Wannier Functions that are the projection onto the occupied space of the current point of those obtained by localization at the central point. If **RELOCAL** is requested these WFs are relocalized at each point

This procedure leads not to real WFs, but to an approximation contained into a cyclic space. In the mapping (unfolding) that permits to convert cyclic to real WFs, 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 the tested cases, 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 WFs, 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).

A set of keywords can be used to modify the localization process (see ***properties*** input, keyword **LOCALWF**, page 327) They are entered after the **DIPOMOME** keyword. Modification of default choices is not recommended, it should be restricted to developers only.

The keyword **DIPOMOME** defines an input block (closed by **END**) with keywords allowing modification of the localization process.

To be modified by developers only.

• A **DIPOMOME**	Calculation of the dipole moment - see Localisation part (properties, keyword **LOCALWF**, page 327 *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 330
• A **CAPTURE**	see **LOCALWF**, page 333
• A **WANDM**	see **LOCALWF**, page 337
• A **FULLBOYS**	see **LOCALWF**, page 337
• A **MAXCYCLE**	see **LOCALWF**, page 330
• A **CYCTOL**	see **LOCALWF**, page 328
• A **ORTHNDIR**	see **LOCALWF**, page 337
• A **CLUSPLUS**	see **LOCALWF**, page 336
• A **PHASETOL**	see **LOCALWF**, page 328
• A **RESTART**	see **LOCALWF**, page 328
• A **IGSSBNDS**	see **LOCALWF**, page 332
• A **IGSSVCTS**	see **LOCALWF**, page 332
• A **IGSSCTRL**	see **LOCALWF**, page 332

8.3.3 IR intensities through CPHF/CPKS

Calculation of IR intensities through Coupled-Perturbed Hartree–Fock/Kohn–Sham approach, keyword **INTCPHF**, allows for completely analytical calculation of Born charges and, hence, IR intensities.

A few optional keywords are available:

- **IRREA**: reads the tensor of Born charges from file **TENS_IR.DAT**. Note that if a frequency restart is performed, the tensor is automatically retrieved from file **FREQINFO.DAT** unless the keyword **DOINTCPHF** is used (see below).
- **IRSPEC**: The IR spectrum is produced using Lorentzian broadening and stored in file **IRSPEC.DAT**

The **INTCPHF** keyword opens a CPHF input block, that must be closed by **END**. Here all keywords proper to the **CPHF** keyword (page 265) can be adopted.

all keywords are optional	II
• A **FMIXING**	see **CPHF**, page 265
• A **RESTART**	see **CPHF**, page 265
• A **TOLALPHA**	see **CPHF**, page 265
• A **MAXCYCLE**	see **CPHF**, page 265

Alternatively to the **INTCPHF** keyword, users can use **DOINTCPHF**. It has the same meaning and optional keywords, but in a case of a frequency restart it forces the calculation of Born (or Raman, see below) tensor instead of reading it from FREQINFO.DAT

8.4 Raman Intensities

The calculation of analytical Raman intensities [181, 182] can be activated through the **INTRAMAN** keyword. If the Raman tensor is not already available (on unit TENS_RAMAN.DAT or inside FREQINFO.DAT) it must be computed by inserting the keyword **INTCPHF**(page 217). As for IR intensities, **INTCPHF** opens a CPHF block that

must be closed by the **END** keyword. Be careful that the **INTRAMAN** should be always used *together with the* **INTENS** *keyword*. A Coupled-Perturbed Hartree–Fock/Kohn–Sham approach calculation up to fourth-order (CPHF2) will be performed, prior to the calculation of frequencies. The Raman tensor, after a successful calculation, is written in a formatted unit named **TENS_RAMAN.DAT**. The same information is stored in the file FREQINFO.DAT The simplest possible input of the frequency block is, then:

```
FREQCALC
INTENS
INTRAMAN
INTCPHF
END
END
```

For an oriented single-crystal the Raman Stokes scattering intensity associated with, for instance, the xy component of the polarizability tensor corresponding to the i-vibrational mode of frequency ω_i may be calculated as:

$$I^i_{xy} \propto \left(\frac{\alpha_{xy}}{\partial \mathcal{Q}_i}\right)^2 \tag{8.6}$$

where $\mathcal{Q}_i$ is the the normal mode coordinate for mode i.
While the intensity of the transverse optical (TO) modes is straightforwardly computed once the appropriate polarizability derivative is obtained, the corresponding calculation for longitudinal optical (LO) modes requires a correction[157, 278] due to $\chi^{(2)}_{bcd}$:

$$\begin{aligned}\left.\frac{\partial \alpha_{b,c}}{\partial \mathcal{R}^A_a}\right|_{\mathcal{R}_0} =& \frac{1}{V}\left.\frac{\partial^3 E^{TOT}}{\partial \mathcal{R}^A_a \partial \mathcal{E}_b \partial \mathcal{E}_c}\right|_{\mathcal{E}=0,\mathcal{R}_0} \\ &- 2\sum_{b'} Z^*_{b'aA} \sum_{d'} \epsilon^{-1}_{b',d'} \chi^{(2)}_{bcd'}\end{aligned} \tag{8.7}$$

In Eq. (8.7) ϵ^{-1} is the inverse of the high-frequency (*i.e.* pure electronic) dielectric tensor. $\chi^{(2)}$ is defined as in Eq. (69) of Ref. [103]. All these quantities ($Z^*_{b'aA}$, ϵ^{-1} and $\chi^{(2)}_{bcd'}$) are saved in FREQINFO.DAT file for restart, since they are obtained as a byproduct of the Raman intensities calculation. Otherwise, the first and second-order dielectric tensors can be provided in input through the keywords **DIELTENS** and **CHI2TENS**, respectively.
Finally, as commonly done in the reporting of experimental data, the intensities are normalized to the highest peak, arbitrarily set to 1000.00.
A few optional keywords are available:

- **RAMANEXP**: In the case the user desires to reproduce experimental conditions, Eq. 8.6 is substituted by

  $$I^i_{xy} \propto C\left(\frac{\alpha_{xy}}{\partial \mathcal{Q}_i}\right)^2 \tag{8.8}$$

  The prefactor C depends[278] on the laser frequency ω_L and the temperature T:

  $$C \sim (\omega_L - \omega_i)^4 \frac{1 + n(\omega_i)}{30\omega_i} \tag{8.9}$$

  with the Bose occupancy factor $n(\omega_i)$ given by

  $$1 + n(\omega_i) = \left[1 - \exp\left(-\frac{\hbar\omega_i}{K_B T}\right)\right]^{-1} \tag{8.10}$$

  The polycrystalline (powder) spectrum can be computed by averaging over the possible orientations of the crystallites as described in Eq. (4) and (5) of Ref. [233], which we used for our implementation.

- **RAMANREA**: reads the Raman tensor from file **TENS_RAMAN.DAT**.Note that if a frequency restart is performed, the tensor is automatically retrieved from file **FREQINFO.DAT** unless the keyword **DOINTCPHF** is used (see below)
- **NORENORM**: Turns off renormalization of the Raman intensity to the highest peak
- **RAMSPEC**: The Raman spectrum is produced using Lorentzian broadening and stored in file **RAMSPEC.DAT**
- **TENSONLY**: After the calculation of the IR and Raman tensors, and their storage in files **TENS_IR.DAT** and **TENS_RAMAN.DAT**, the calculation stops
- **ROTRAMAN**
  θ, φ, ψ
  The Raman tensor is rotated according to the three angles (in degrees) provided in input. For the meaning of such angles and see the effect of the rotations see `http://www.cryst.ehu.es/cryst/transformtensor.html`

Here all keywords proper to the **CPHF** keyword (page 265) can be adopted.
Note that in the case of a DFT calculation the CPHF defaults are set to `FMIXING=FMIXING2=60`, `MAXCYCLE=MAXCYCLE2=200`.

all keywords are optional II

- A **FMIXING2** see **CPHF**, page 265
- A **TOLGAMMA** see **CPHF**, page 265
- A **MAXCYCLE2** see **CPHF**, page 265

Alternatively to the **INTCPHF** keyword, users can use **DOINTCPHF**. It has the same meaning and optional keywords, but in a case of a frequency restart it forces the calculation of Born and Raman tensor instead of reading it from FREQINFO.DAT

8.4.1 Vibrational contribution to SHG and dc-Pockels tensors

rec	variable	meaning
• *A*	**BETAVIB**	Activates the calculation of the vibrational contribution to the SHG and Pockels tensor
∗	BVIB_FRQ	Wavelength (in nm) of the frequency ω_1

The double harmonic vibrational $[\mu\alpha]^{(0,0)}$ (or ionic) contribution to the SHG or Pockels tensors (see Section about DYNAMIC CPHF on page 268 for the electronic part of the tensor) is obtained by a simple generalization of the molecular formula [28, 238]:

$$d^{vib}_{tuv}(-\omega_\sigma;\omega_1,\omega_2) = \frac{1}{2}\frac{2\pi}{V}\sum_{i=1}^{3N-6} P_{\frac{t}{-\omega_\sigma},\frac{u}{+\omega_1},\frac{v}{+\omega_2}} \frac{\left(\frac{\partial\mu_t}{\partial Q_i}\right)\left(\frac{\partial\alpha_{uv}}{\partial Q_i}\right)}{\omega_i^2-\omega_\sigma^2} \tag{8.11}$$

where μ_t and α_{uv} are dipole moment and polarizability components, while ω_i and Q_i are phonon frequencies and normal modes of the i-th vibration at the Γ-point. In SHG, $\omega_1 = \omega_2$ and $\omega_\sigma = 2\omega_1$, while in Pockels, $\omega_\sigma = -\omega_1$ and $\omega_2 = 0$ The operator $P_{\frac{t}{-\omega_\sigma},\frac{u}{+\omega_1},\frac{v}{+\omega_2}}$ which permutes the pairs $(t/-\omega_\sigma)$, $(u/+\omega_1)$, and $(v/+\omega_2)$ has been discussed, for example, by Bishop[27])
The terms $\frac{\partial\mu_t}{\partial Q_i}$ and $\frac{\partial\alpha_{uv}}{\partial Q_i}$ are the same tensors needed for the evaluation of infrared and Raman intensities and are computed as described in Refs. [180] and [183], respectively. As a consequence, this keyword can be activated only if INTENS and INTRAMAN keywords are present in the same input.

8.5 Scanning of geometry along selected normal modes

Scanning of geometry along selected normal modes is invoked by the keyword **SCANMODE** in **FREQCALC** input block. Preliminary frequency calculation is required to single out the selected mode.

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 16).
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).

8.5.1 Example 1 - Methane molecule

First run: optimization of the geometry (full input at page 386).

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 1-3, translational mode; modes 4-6, rotational modes).

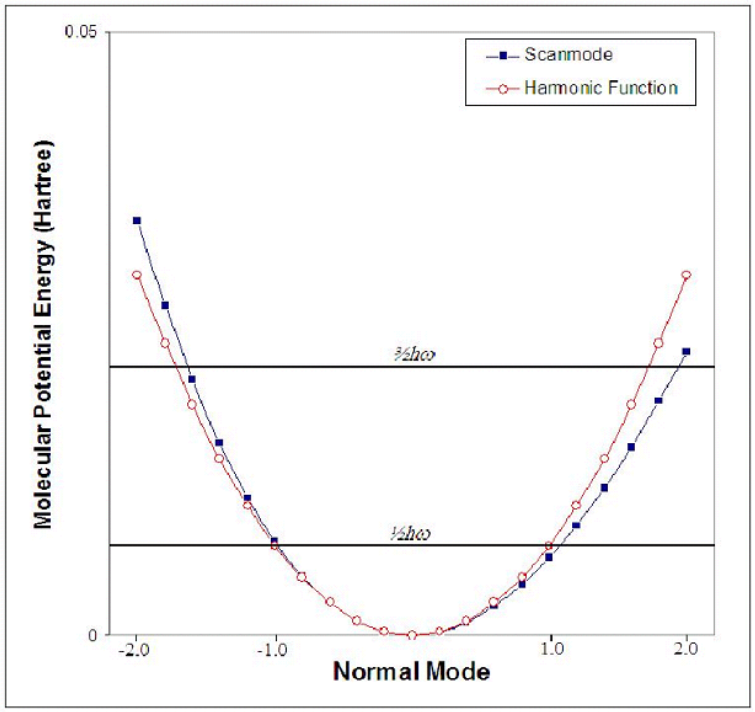
MODES	EIGV (HARTREE**2)	FREQUENCIES (CM**-1)	(THZ)	IRREP	IR	RAMAN
1- 3	-0.1863E-11	-0.2995	-0.0090	(F2)	A	A
4- 6	0.7530E-07	60.2270	1.8056	(F1)	I	I
7- 9	0.4821E-04	1523.8308	45.6833	(F2)	A	A
10- 11	0.6302E-04	1742.3056	52.2330	(E)	I	A
12- 12	0.2099E-03	3179.3763	95.3153	(A)	I	A
13- 15	0.2223E-03	3272.4193	98.1047	(F2)	A	A

Third run: Scanning of a selected mode.
To explore the 12th normal mode, corresponding to C-H symmetric stretching, 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.

Figure 8.1: Scanning of the energy along normal mode 12, ν=3179.3763 cm^{-1}, corresponding to C–H symmetric stretching



FREQ(CM**-1) 3179.38

AT. 1 C X 0.0000
Y 0.0000
Z 0.0000
AT. 2 H X 0.0560
Y 0.0560
Z 0.0560
AT. 3 H X -0.0560
Y -0.0560
Z 0.0560
AT. 4 H X 0.0560
Y -0.0560
Z -0.0560
AT. 5 H X -0.0560
Y 0.0560
Z -0.0560

8.5.2 Example 2 - $PbCO_3$

The space group of this carbonate, as it can be found in the literature [ICSD database], is *Pmcn* (orthorhombic lattice).

First run: full optimization of the geometry in *Pmcn* space group (full input at page 386).

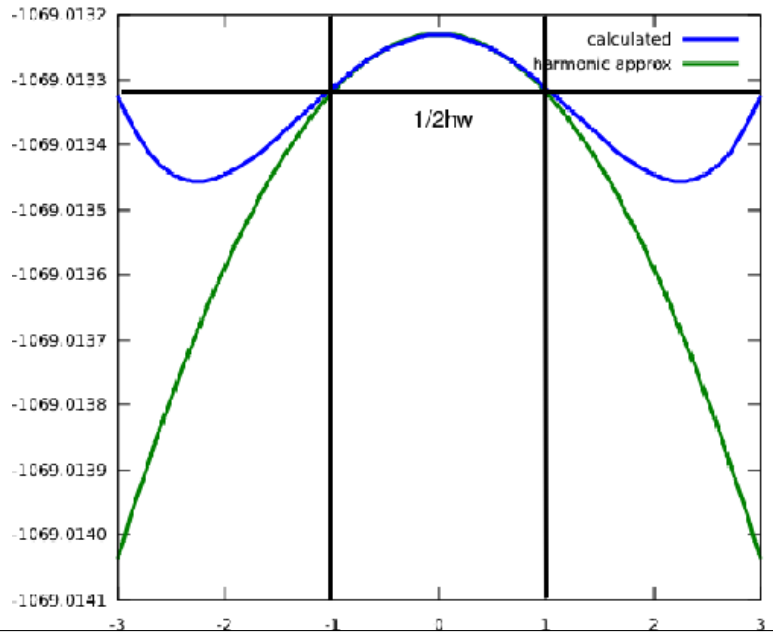
Second run: frequency calculation. The output would look as follows:

MODES		EIGV (HARTREE**2)	FREQUENCIES (CM**-1)	(THZ)	IRREP	IR	RAMAN
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 [201]); 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

Figure 8.2: Scanning of the energy along normal mode 1, corresponding to a frequency of -39.3362 cm^{-1} (L. Valenzano, unpuplished results)



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
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 16) from one of the files written during the scan, copied as file fort.34. For instance SCAN1_-39.3361_DISP_-2.400 (scan of mode 1, frequency -39.3361 cm^{-1}, displacement -2.4 the classical amplitude).
Please refer to the standard script for running CRYSTAL09 to handle input/output file names.
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.

8.6 Calculation of the infrared spectrum

Keyword **IRSPEC**, inserted in the **FREQCALC** input block, activates the calculation of the infrared spectra. Prior calculation of IR intensities is required (keyword **INTENS**, page 215) and definition of the dielectric tensor matrix (keyword **DIELTENS**, page 210) or of the isotropic dielectric constant (keyword **DIELISO**, page 210).

The keyword **IRSPEC** defines an input block (closed by **END**).

The key quantity is the complex dielectric tensor $\epsilon(\nu)$, which is computed for each inequivalent polarization direction on the basis of a classical Drude-Lorentz model:

$$\epsilon_{ii}(\nu) = \epsilon_{\infty,ii} + \sum_p \frac{f_{p,ii}\nu_p^2}{\nu_p^2 - \nu^2 - i\nu\gamma_p} \tag{8.12}$$

where ii indicates the polarization direction, ϵ_∞ is the optical dielectric tensor, ν_p, f_p and γ_p are the TO frequency, oscillator strength and damping factor for the p^{th} vibrational mode, respectively. The real and imaginary parts of $\epsilon(\nu)$ are computed; the maxima of the latter function correspond to the TO frequencies. The imaginary part of $1/\epsilon(\nu)$ is computed as well, whose maxima correspond to the LO frequencies.

The real and imaginary parts of the refractive index $n(\nu)$ are obtained for each inequivalent polarization direction by exploiting the relations:

$$\begin{aligned} \{Re[n_{ii}(\nu)]\}^2 - \{Im[n_{ii}(\nu)]\}^2 &= Re[\epsilon_{ii}(\nu)] \\ 2 \cdot Re[n_{ii}(\nu)] \cdot Im[n_{ii}(\nu)] &= Im[\epsilon_{ii}(\nu)] \end{aligned} \tag{8.13}$$

The reflectance curve $R(\nu)$ is computed for each inequivalent polarization direction through:

$$R_{ii}(\nu) = \left| \frac{\sqrt{\epsilon_{ii}(\nu) - \sin^2(\theta)} - \cos(\theta)}{\sqrt{\epsilon_{ii}(\nu) - \sin^2(\theta)} + \cos(\theta)} \right|^2 \tag{8.14}$$

where θ is the angle between the incident beam and the normal to the crystal surface.

The absorbance spectrum $A(\nu)$ is calculated according to four different models:

1. raw superposition of Lorentzian (or Gaussian) peaks; this is the only quantity not requiring the dielectric function, as it only requires the TO frequencies ν_p and the integrated intensities I_p computed as described from page 215 on:

$$A_{raw,Lor}(\nu) = \sum_p \frac{I_p}{\pi} \frac{\gamma_p/2}{(\nu - \nu_p)^2 + \gamma_p^2/4} \tag{8.15}$$

$$A_{raw,Gau}(\nu) = \sum_p 2\sqrt{\frac{\ln 2}{\pi}} \frac{I_p}{\gamma_p} \exp\left[-\frac{4 \ln 2 (\nu - \nu_p)^2}{\gamma_p^2} \right] \tag{8.16}$$

2. classical absorption formula, averaged over the polarization directions 1, 2, and 3 (according to Voigt notation, page 406):

$$A_{clas}(\nu) = \frac{1}{3} \sum_{ii=1}^{3} \frac{4\pi}{\lambda\rho} Im[n_{ii}(\nu)] \tag{8.17}$$

   where λ is the wavelength of the incident beam and ρ is the crystal density;

3. Rayleigh approximation of spherical particles [30], averaged over the directions 1,2 and 3:

$$A_{sph}(\nu) = \frac{1}{3} \sum_{ii=1}^{3} \frac{2\pi}{\lambda\rho} Im\left[\frac{\epsilon_{ii}(\nu) - 1}{\epsilon_{ii}(\nu) + 2} \right] \tag{8.18}$$

4. Rayleigh approximation of particles as continuous distribution of ellipsoids (CDE) [30], averaged over the directions 1,2 and 3:

$$A_{CDE}(\nu) = \frac{1}{3} \sum_{ii=1}^{3} \frac{2\pi}{\lambda\rho} Im\left[\frac{2\epsilon_{ii}(\nu)}{\epsilon_{ii}(\nu) - 1} \log \epsilon_{ii}(\nu) \right] \tag{8.19}$$

Reflectance spectrum, dielectric function and refractive index are at the moment limited to:

- frequency-independent optical dielectric tensor (CPHF) and frequency-dependent vibrational (IR) contributions
- directions 1, 2 and 3 only (Voigt notation, $\epsilon_{11}, \epsilon_{22}, \epsilon_{33}$)

Generalization is in progress.

IRSPEC data are written in files IRSPEC.DAT, IRREFR.DAT and IRDIEL.DAT and can be directly plotted with *gnuplot* (http://www.gnuplot.info, appendix D, page 448).

Once the vibrational frequencies, IR intensities and dielectric tensor are calculated, the infrared spectra can be calculated at almost zero computational time with a **RESTART** in the **FREQCALC** input block (the FREQINFO.DAT file is required):

```
FREQCALC
RESTART
INTENS
DIELTENS or DIELISO
```

```
[input dielectric tensor or constant]
IRSPEC
[optional keywords for IRSPEC]
END           [end IRSPEC]
END           [end FREQCALC]
```

By default, only the absorption and reflectance spectra are generated.

If the dielectric tensor is omitted, only the raw absorption spectrum is computed.

To enable the generation of the dielectric function and refractive index, specific keywords are required:

rec	variable	meaning
•	**REFRIND**	Activates the generation of the refractive index (same directions than reflectance spectrum).
•	**DIELFUN**	Activates the generation of the dielectric function (same directions than reflectance spectrum).

There are several keywords which allow to modify the **default** values.

rec	variable	meaning
•	A **LENSTEP**	
•	∗ X	Step size used for the spectra [1 cm^{-1}].
•	A **NUMSTEP**	
•	∗ N	Number of points used for the spectra [no default]. To be used as an alternative to LENSTEP.
•	A **RANGE**	
•	∗ X1,X2	Frequency range used for the spectra Default: [0., ν_{max} + 400.], where ν_{max} is the frequency of the highest calculated mode.
•	**GAUSS**	Gaussian line shape is adopted for the raw absorbance spectrum, instead of the Lorentzian one [which is the default].
•	A **DAMPFAC**	
•	∗ GAMMA	Damping factor (related to peak width) used for the spectra [8.0].
•	A **ANGLE**	
•	∗ ALPH	Angle of incidence used for the reflectance spectrum (degrees) [10.0].

8.7 Calculation of the Raman spectra

Keyword **RAMSPEC**, inserted in the **FREQCALC** input block, activates the calculation of the Raman spectra. Prior calculation of Raman intensities is required (keyword **INTENS** with **INTRAMAN** option, page 215).

The keyword **RAMSPEC** defines an input block (closed by **END**).

The Raman spectra $A(\nu)$ are computed for both the cases of polycrystalline powder (total intensity, parallel polarisation, perpendicular polarisation) and single crystal (xx, xy, xz, yy, yz, zz polarisations). They are constructed by using the Transverse Optical (TO) modes and by adopting a pseudo-Voigt functional form:

$$A(\nu) = \eta \cdot L(\nu) + (1 - \eta) \cdot G(\nu) \tag{8.20}$$

with:

$$L(\nu) = \sum_p \frac{I_p}{\pi} \frac{\gamma_p/2}{(\nu - \nu_p)^2 + \gamma_p^2/4} \tag{8.21}$$

$$G(\nu) = \sum_p 2\sqrt{\frac{\ln 2}{\pi}} \frac{I_p}{\gamma_p} \exp\left[-\frac{4\ln 2(\nu - \nu_p)^2}{\gamma_p^2}\right] \tag{8.22}$$

where ν_p and I_p are the computed TO frequencies and Raman intensities, respectively, for each mode p; γ_p is the full width at half maximum; η is the Lorentz factor, with values in the range $0 - 1$. Raman intensities in CRYSTAL are normalized, so that the largest value is set to 1000 (non-normalized intensities can be obtained by inserting the **NORENORM** sub-keyword). Note that η values of 1 and 0 correspond to pure Lorentzian and pure Gaussian functional forms, respectively.

RAMSPEC data are written in file RAMSPEC.DAT and can be directly plotted with *gnuplot* (http://www.gnuplot.info, appendix D, page 449).

Once the vibrational frequencies and Raman intensities are calculated, the Raman spectra can be calculated at almost zero computational cost with a **RESTART** in the **FREQCALC** input block (the FREQINFO.DAT and TENS_RAMAN.DAT files are required):

```
FREQCALC
RESTART
INTENS
INTRAMAN
RAMANREA
RAMSPEC
[optional keywords for RAMSPEC]
END           [end RAMSPEC]
END           [end FREQCALC]
```

There are several keywords which allow to modify the **default** values.

rec	variable	meaning
•	A **LENSTEP**	
•	∗ X	Step size used for the spectra [1 cm^{-1}].
•	A **NUMSTEP**	
•	∗ N	Number of points used for the spectra [no default]. To be used as an alternative to LENSTEP.
•	A **RANGE**	
•	∗ X1,X2	Frequency range used for the spectra [0., 1.3 ν_{max}], where ν_{max} is the frequency of the highest calculated mode.
•	A **VOIGT**	
•	∗ ETA	Lorentz factor, determining mixing between Lorentzian and Gaussian contributions [1.0].
•	A **DAMPFAC**	
•	∗ FWHM	Full width at half maximum used for the spectra [8.0].

8.8 Phonon dispersion

The keyword **DISPERSION**, inserted in the **FREQCALC** input block, activates the calculation of vibration frequencies and normal modes on a set of points in reciprocal space, defined in terms of a direct-space supercell to be generated in input with the **SCELPHONO** keyword (page 68).

In the presence of interatomic interactions the static (non-vibrating) crystal assumes an equilibrium configuration $\mathcal{R}_0 \equiv \{\dots\ ,\ [(\mathbf{R}_0)_a + \mathbf{g}]\ ,\ \dots\}$ that is unambiguously defined by the equilibrium positions $\{(\mathbf{R}_0)_a\}$ of the N atoms of the cell ($a = 1, \dots, N$); the lattice vector $\mathbf{g} = \sum_{m=1}^{3} l_m^g\, \mathbf{a}_m$ identifies the general crystal cell where $\mathbf{a}_m$ are the direct lattice basis vectors: within Born von Kármán periodic boundary conditions the integers l_m^g run from 0 to $L_m - 1$. The parameters $\{\ L_m\}$ define the size and shape of a supercell (SC) in direct space (to be generated with the **SCELPHONO** keyword).
When nuclear motion (due to Heisenberg principle, finite temperature or other external perturbations) is considered, the atomic equilibrium positions become the static average positions of the atoms displaced by $\mathbf{u}_a^{\mathbf{g}}$ which define the general configuration $\mathcal{R} \equiv \{\dots\ ,\ [(\mathbf{R}_0)_a + \mathbf{g} + \mathbf{u}_a^{\mathbf{g}}]\ ,\ \dots\}$.
Let us introduce the $L = L_1 \times L_2 \times L_3$ Hessian matrices $\{\mathbf{H}^{\mathbf{g}}\}$ whose elements are the second derivatives of the total energy per cell with respect to the atomic displacements:

$$H^{\mathbf{g}}_{ai,bj} = \left(\frac{\partial^2 E}{\partial u^{\mathbf{0}}_{ai} \partial u^{\mathbf{g}}_{bj}} \right) \tag{8.23}$$

where atom a in the reference cell is displaced along the i-th Cartesian direction and atom b in cell $\mathbf{g}$ is displaced, along with all its periodic images in the crystal, along the j-th Cartesian

direction. First derivatives are computed analytically while second derivatives are computed numerically.
The set of L Hessian matrices $\{\mathbf{H^g}\}$ can be Fourier transformed into a set of *dynamical matrices* $\left\{\mathbf{W^k}\right\}$ each one associated with a wavevector $\mathbf{k} = \sum_{n=1}^{3} (\kappa_n / L_n)\, \mathbf{b}_n$ where $\mathbf{b}_n$ are the reciprocal lattice vectors and the integers κ_n run from 0 to $L_n - 1$:

$$\mathbf{W^k} = \sum_{\mathbf{g}=1}^{L} \mathbf{M}^{-\frac{1}{2}} \mathbf{H^g} \mathbf{M}^{-\frac{1}{2}} \, \exp(\imath \mathbf{k} \cdot \mathbf{g}) \, , \tag{8.24}$$

where $\mathbf{M}$ is the diagonal matrix of the nuclear masses. The diagonalization of the dynamical matrices

$$(\mathbf{U^k})^\dagger \, \mathbf{W^k} \, \mathbf{U^k} \;=\; \mathbf{\Lambda^k} \quad \text{with} \quad (\mathbf{U^k})^\dagger \mathbf{U^k} = \mathbf{I} \, . \tag{8.25}$$

provides with the *vibration frequencies* $\nu_i^{\mathbf{k}} = \sqrt{\lambda_i^{\mathbf{k}}}$ (from the eigenvalues) and the *normal modes* (from the columns of $\mathbf{U^k}$):

$$\mathbf{q^k} = \mathbf{M}^{\frac{1}{2}} (\mathbf{U^k})^\dagger \mathbf{u^k} \quad \text{with} \quad \mathbf{u^k} = \frac{1}{\sqrt{L}} \sum_{\mathbf{g}=1}^{L} \mathbf{u^g} \, \exp(\imath \mathbf{k} \cdot \mathbf{g}) \, .$$

To each **k**-point in the first Brillouin zone, $3N$ oscillators (*i.e.* phonons) are associated which are labeled by a *phonon band index* i ($i = 1, \ldots, 3N$).
In principle, equation (8.24) can be used to compute, and then diagonalize according to equation (8.25), the dynamical matrices of just the L **k**-points defined above. However, if the energy second derivatives $\{\mathbf{H^g}\}$ vanish within the supercell (SC) defined by the keyword **SCELPHONO** then such an expression can be used to construct the dynamical matrices of a denser set of **k**-points represented by three parameters $\{L'_m \geq L_m\}$, with $m = 1, \ldots, D$ (where D is the dimensionality of the system: 1, 2, 3 for 1D, 2D, 3D periodic systems). The quantum contributions to the second derivatives of the total energy usually vanish within a SC of radius $\approx$ 10-15 Å. Such an *interpolation technique*, can be easily activated with the **INTERPHESS** keyword (to be inserted after **DISPERSION**) and can be quite effective in the case of a fully covalent crystal without long-range electrostatic contribution to the total energy.
On the contrary, when such electrostatic contributions become relevant (as in ionic crystals), they have to be explicitly accounted for with appropriate corrections (see the **WANG** keyword below). In this case one should use a SC large enough to contain all relevant quantum contributions and the two keywords **INTERPHESS** and **WANG**.
Phonon bands and phonon density of states can be computed with sub-keywords **BANDS** and **PDOS**, respectively.
A detailed description of these options follows:

- A **NOKSYMDISP** The dynamical matrices $\left\{\mathbf{W^k}\right\}$ in reciprocal space are not factorized according to the irreducible representations of the Space Group when diagonalized. Phonons are not labeled according to symmetry.

• A	**BANDS**	Phonon bands calculation for plotting purposes. Phonon bands are sampled along given directions in reciprocal space. Data is stored in files fort.25 (Crgra format) and PHONBANDS.DAT (xmgrace format). Note: **BANDS** implies **NOKSYMDISP**.
• ∗	ISS	Shrinking factor in terms of which the coordinates of the extremes of the segments are expressed.
	NSUB	Total number of k points along each of the lines in the path.
∗	NLINE	Number of lines in reciprocal space to be explored.
		if ISS > 0 then
		add NLINE records
• ∗	I1,I2,I3	Integers that define the starting point of the line (I1/ISS $\mathbf{b}_1$+I2/ISS $\mathbf{b}_2$+I3/ISS $\mathbf{b}_3$), with $\mathbf{b}_1$,$\mathbf{b}_2$,$\mathbf{b}_3$ reciprocal lattice vectors.
	J1,J2,J3	Integers that define the final point of the line (J1/ISS $\mathbf{b}_1$+J2/ISS $\mathbf{b}_2$+J3/ISS $\mathbf{b}_3$) with $\mathbf{b}_1$,$\mathbf{b}_2$,$\mathbf{b}_3$ reciprocal lattice vectors.
		if ISS = 0 then
		add NLINE records
• ∗	LABELA	Label of the starting point of the line (see tables 14.1 and 14.2 below for a legend).
	LABELB	Label of the final point of the line (see tables 14.1 and 14.2 below for a legend).

• A	**INTERPHESS**	Activates the Hessian Fourier interpolation that permits the calculation of vibration frequencies on a denser set of points in reciprocal space. This option has to be used only if the starting supercell (SC) is sufficiently large (radius ≈ 10-15 Å) so that any quantum contribution to the energy second derivatives vanishes within it. This keyword has to be combined with the keyword **WANG** if long-range electrostatic contributions to the energy exist (ionic crystals).
• ∗	L'_m $m = 1, \ldots, D$	where D is the dimensionality of the system: 1, 2, 3 for 1D, 2D, 3D periodic systems. These three integers are the expansion parameters of the starting SC.
∗	IPRINT	0 Output printings disabled for each **k**-point. 1 Output printings active for each **k**-point.

• A	**WANG**	For 3D systems only. Used in combination with **INTERPHESS**, **BANDS** and **PDOS** when dealing with polar materials, activates the correction to the dynamical matrices to take into account the long range Coulomb interactions. The mixed-space approach as proposed by Wang and coworkers [283] is implemented. Born tensor charges are read from the external file BORN.DAT which is obtained by performing a Γ only vibrational frequencies calculation with IR intensities active (see **INTENS** subkeyword, page 215).
• ∗	TENS(1:9)	Dielectric tensor matrix TENS (3x3 elements, input by rows: 9 reals (3D). Can be obtained with a CPHF calculation (see page 265).

8.9 Anisotropic Displacement Parameters (ADP)

The keyword **ADP**, inserted in the **FREQCALC** block, allows to compute the anisotropic displacement parameters (ADP) as 3×3 tensors $\mathbf{U}_A$ associated to each atom A of the cell. Such tensors can be used for computing Debye-Waller thermal factors and obtaining dynamic structure factors (see the **XFAC** keyword). Each 3×3 atomic tensor $\mathbf{U}_A$ can be diagonalized as $\mathbf{U}_A\mathbf{E}_A = \mathbf{e}_A\mathbf{E}_A$ where $\mathbf{e}_A$ is the diagonal matrix of the eigenvalues. If the three eigenvalues

are positive, then the surfaces of constant probability are *ellipsoids* enclosing some definite probability for atomic displacement. The lengths of the principal axes of the ellipsoids are proportional to the eigenvalues e_1, e_2 and e_3 of $\mathbf{U}_A$ which are usually expressed in 10^{-4} Å^2 . The orientation of the ellipsoid with respect to the reference Cartesian frame is given by the eigenvectors $\mathbf{E}_A$.
Let us consider the set of L (number of cells in the cyclic crystal) $3n \times 3n$ (with n number of atoms per cell) Hessian matrices $\mathbf{H}(\mathbf{g})$ whose elements are the second derivatives of the energy E with respect to the atomic displacements $\mathbf{u}$:

$$H_{A\mathbf{0}i,A'\mathbf{g}i'}(\mathbf{g}) = \frac{\partial^2 E}{\partial u_{A\mathbf{0}i} \partial u_{A'\mathbf{g}i'}}$$

where $\mathbf{g}$ labels a crystal cell and i a Cartesian component of $\mathbf{u}$. The usual way of dealing with lattice dynamics is considering the set of L $3n \times 3n$ dynamical matrices $\mathbf{D}(\mathbf{k})$ defined as the (mass-weighted) Fourier transform of the $\mathbf{H}(\mathbf{g})$:

$$\mathbf{D}(\mathbf{k}) = \sum_{\mathbf{g}}^{L} \mathbf{M}^{-\frac{1}{2}} \mathbf{H}(\mathbf{g}) \mathbf{M}^{-\frac{1}{2}} e^{i\mathbf{k}\cdot\mathbf{g}}$$

where $\mathbf{M}$ is a $3n \times 3n$ real symmetric (diagonal) matrix with the nuclear masses on the diagonal. To each $\mathbf{k}$, $3n$ oscillators (phonons) are associated whose label is $s = 1, \ldots, 3n$ and whose frequencies and eigenvectors are obtained by diagonalizing $\mathbf{D}(\mathbf{k})$:

$$\mathbf{D}(\mathbf{k})\mathbf{W}(\mathbf{k}) = \mathbf{\Omega}(\mathbf{k})\mathbf{W}(\mathbf{k}) \quad \text{with} \quad \mathbf{W}^\dagger(\mathbf{k})\mathbf{W}(\mathbf{k}) = \mathbf{I}$$

To each phonon $k \equiv (\mathbf{k}s)$, an atomic displacement vector $\mathbf{u}_k$ can be associated. Let us introduce the frequency-scaled normal coordinate proper of each phonon k: $\xi_k = \mathbf{\Omega}^{\frac{1}{4}}(\mathbf{k})\mathbf{W}^\dagger(\mathbf{k})\mathbf{M}^{\frac{1}{2}}\mathbf{u}_k$. The Boltzman probability density function $p(\{\xi_k\}; T)$ for the nuclei, at a given temperature T can be expressed as the product of independent probabilities, associated to the different phonons $p(\{\xi_k\}; T) = \prod_k p(\xi_k; T)$. Let us introduce the so-called *Gaussian approximation* that consists in expressing such probabilities as Gaussian functions with standard deviations $\sigma_{k,T}$ (k_B is Boltzman's constant):

$$p(\xi_k; T) \approx G(\xi_k; \sigma_{k,T}) \quad \text{with} \quad (\sigma_{k,T})^2 = |\xi_k\rangle\langle\xi_k| = \left[\frac{1}{e^{\frac{\omega_k}{k_B T}} - 1} + \frac{1}{2}\right]$$

Within such an approximation, $p(\{\xi_k\}; T) = e^{-\sum_{\mathbf{k}} \langle \mathbf{u}_\mathbf{k} | \mathbf{U}(\mathbf{k};T)^{-1} | \mathbf{u}_\mathbf{k} \rangle}$ with

$$\mathbf{U}(\mathbf{k}; T) = \mathbf{M}^{-\frac{1}{2}} \mathbf{W}(\mathbf{k}) \mathbf{\Omega}^{-\frac{1}{4}}(\mathbf{k}) \mathbf{\Xi}(T) \mathbf{\Omega}^{-\frac{1}{4}}(\mathbf{k}) \mathbf{W}^\dagger(\mathbf{k}) \mathbf{M}^{-\frac{1}{2}}$$

where $\mathbf{\Xi}(T)$ is a $3n \times 3n$ diagonal matrix with $(\sigma_{k,T})^2$ as diagonal elements, so that we can define the total mean square displacement tensors $\mathbf{U}(\mathbf{g}; T) = 1/L \sum_{\mathbf{k}} \mathbf{U}(\mathbf{k}; T) e^{-i\mathbf{k}\cdot\mathbf{g}}$. The 3×3 diagonal blocks of $\mathbf{U}(\mathbf{g} = \mathbf{0}; T)$ are the ADPs $\mathbf{U}_A$.
The default temperature at which the ADPs are computed is 298.15 K. Different temperatures can be defined with the keyword **TEMPERAT**. The ADPs are saved (in atomic units) in the external formatted unit ADP.DAT that can be used by the keyword **XFAC** of PROPERTIES for computing Debye-Waller thermal factors for dynamic structure factors.

Users of this option are kindly requested to cite the following paper [91]:

A. Erba, M. Ferrabone, R. Orlando and R. Dovesi, *J. Comput. Chem.*, **34**, 346 (2013). *Accurate dynamical structure factors from ab initio lattice dynamics: The case of crystalline silicon.*

rec		variable	value	meaning
•	∗	NTYP	0	Recommended algorithm for the computation of the ADPs
		NNEGL		Number of additional modes at the Γ-point to be neglected (others than rotations+translations that are already neglected) in the computation of the ADPs

8.10 Phonon Density-of-States (and Inelastic Neutron Scattering Spectra)

Knowledge of the vibration frequencies of a system also allows to compute the phonon density-of-states (PDOS) and to simulate the results of inelastic neutron scattering (INS) experiments. The total PDOS $g(\omega)$ is defined by the equation:

$$g(\omega) = \frac{1}{V_{\rm BZ}} \int_{\rm BZ} \sum_{p=1}^{3N} \delta(\omega_{\mathbf{k}p} - \omega) d\mathbf{k} \ , \tag{8.26}$$

where $\omega_{\mathbf{k}p}$ is the vibration frequency of a phonon with wave-vector $\mathbf{k}$, of branch p, and $V_{\rm BZ}$ is the volume of the Brillouin zone, and the integration is performed over it. From equation (8.26), the PDOS is normalized to $3N$, being N the number of atoms per cell ($\int g(\omega)d\omega = 3N$). The total PDOS can be partitioned into atomic contributions $g(\omega) = \sum_a g_a(\omega) x_a$ where the sum runs over the atomic species of the system, x_a is the fraction of atomic species a with respect to N, and

$$g_a(\omega) = \frac{1}{n_{\mathbf{k}}} \sum_{p,\mathbf{k}} |\mathbf{e}_{p,\mathbf{k};a}|^2 \delta(\omega_{\mathbf{k}p} - \omega) \ , \tag{8.27}$$

where $\mathbf{e}_{p,\mathbf{k}}$ are the eigenvectors of the dynamical matrices $W^{\mathbf{k}}$ defined in equation (8.24) and the integral in equation (8.26) has been replaced by the sum over the sampled points within the FBZ.
From atomic projected PDOS, a neutron-weighted phonon density-of-states (NW-PDOS) may be defined, which can be compared to the outcomes of INS experiments:

$$g^{\rm NW}(\omega) = C \sum_a \frac{\sigma_a}{M_a} g_a(\omega) x_a \ , \tag{8.28}$$

where C is a normalization factor such that $\int g^{\rm NW}(\omega)d\omega = 3N$, and the weight of each atomic species a is given by the ratio of the atomic scattering cross-section σ_a and the atomic mass M_a [200, 174]. Depending on whether the inelastic scattering is coherent or incoherent, different cross-sections have to be considered, which are tabulated and available on-line [258, 148].
The phonon density-of-state can be computed also from a Γ-only calculation but the cell on which the vibration frequencies are computed should be large in that case.

Users of this module are kindly asked to cite the following reference:
J. Baima, M. Ferrabone, R. Orlando, A. Erba and R. Dovesi, *Phys. Chem. Minerals*, **43**, 137-149 (2016)

•	A **PDOS**	Phonon density-of-states is computed.
•	∗ NUMA	Maximum frequency considered (value in $\rm cm^{-1}$).
	NBIN	Number of intervals in which the frequency range $0 < \omega <$ NUMAX is partitioned for the representation of the PDOS [recommended value > 50].
	∗ LPRO	0 No projected atomic DOS. 1 Projected atomic DOS enabled.

• A	**INS**	Neutron-weighted phonon density-of-states is computed, which can be compared with Inelastic neutron scattering spectra (INS)
• ∗	NUMA	Maximum frequency considered (value in cm^{-1}).
	NBIN	Number of intervals in which the frequency range $0 < \omega <$ NUMAX is partitioned for the representation of the PDOS [recommended value > 50].
∗	NWTYPE	0 Coherent cross-section. 1 Incoherent cross-section. 2 Coherent + Incoherent cross-section.

Input examples are as follows:

```
FREQCALC
DISPERSI      [optional: to run a phonon dispersion calculation rather than a Gamma-only one]
PDOS
2500 250      [Range: 0-2500 cm^-1 subdivided into 250 intervals for histogram representation]
1             [Atomic partitioning of the PDOS active]
END
```

Another one

```
FREQCALC
DISPERSI      [optional: to run a phonon dispersion calculation rather than a Gamma-only one]
INS
3000 300      [Range: 0-3000 cm^-1 subdivided into 300 intervals for histogram representation]
2             [Coherent + Incoherent cross-section in neutron-weighting]
END
```

8.11 Anharmonic calculation of frequencies of X-H (X-D) bond stretching

Anharmonic calculation of frequencies of X-H, or X-D, bond stretching (where H and D stand for hydrogen and deuterium and X for any element) is invoked by the keyword **ANHARM** that has to be inserted outside from the **FREQCALC** block (a **FREQCALC** run is not needed at all for this purpose).

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 [172]. 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

TEST[ANHA] 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
HF (molecule)	W_{01}	4358.6	4359.0
	W_{02}	8607.3	8608.1
	W_e	4468.6	4468.8
	W_eX_e	55.0	54.9
$Be(OH)_2$ (bulk)	W_{01}	3325.3	3325.8
	W_{02}	6406.3	6407.4
	W_e	3569.5	3569.9
	W_eX_e	122.1	122.1
$Ca(OH)_2$ (bulk)	W_{01}	3637.2	3637.5
	W_{02}	7111.4	7111.9
	W_e	3800.3	3800.7
	W_eX_e	81.5	81.6

8.12 Anharmonic Terms of the PES and Anharmonic States from VSCF and VCI

In CRYSTAL23 we have implemented algorithms for the evaluation of high-order terms of the potential energy surface (PES) and for the vibrational self-consistent field (VSCF) and vibrational configuration interaction (VCI) calculation of anharmonic vibrational states.
In this section, we present the implemented algorithms and briefly introduce some formal aspects of anharmonic vibrational states of molecules and solids, where, in the case of solids, we restrict our attention to Γ-point vibration modes. However, let us note that, by working in terms of a supercell of the primitive one, vibration modes of solids proper of different **k**-points can be folded back to the Γ-point.
Please, keep in mind that tight computational parameters (i.e. tight convergence thresholds for the SCF and geometry optimization steps) are often necessary for this type of calculations.

Please, do not forget to cite this papers when using the options discussed below:

A. Erba, J. Maul, M. Ferrabone, P. Carbonniére, M. Rérat and R. Dovesi, *J. Chem. Theory Comput.*, **15**, 3755-3765 (2019) *Anharmonic Vibrational States of Solids from DFT Calculations. Part I: Description of the Potential Energy Surface*

A. Erba*, J. Maul, M. Ferrabone, R. Dovesi, M. Rérat and P. Carbonniére, *J. Chem. Theory Comput.*, **15**, 3766-3777 (2019) *Anharmonic Vibrational States of Solids from DFT Calculations. Part II: Implementation of the VSCF and VCI Methods*

8.12.1 Evaluation of High-Order Terms of the PES

The starting point of our anharmonic vibrational description is represented by the harmonic approximation according to which the nuclear dynamics of the system is described in terms of a set of M independent quantum harmonic oscillators, whose corresponding normal coordinates are $Q_1, Q_2, \cdots, Q_M \equiv \mathbf{Q}$.
The numerical description of high-order terms of the PES represents the most delicate and computationally expensive step in the anharmonic treatment of vibrational states of materials. In our strategy, the PES is truncated to quartic order and contains one-, two- and three-mode interatomic force constants. Four different numerical approaches have been implemented, all based on a grid representation of the PES in the basis of the normal coordinates, that require different ingredients (energy and/or forces) to be evaluated at each point (i.e. nuclear configuration) of the grid. Different algorithms have been explored to compute the high-order energy derivatives: energy fitting and finite differences. The numerical stability and relative computational efficiency of the various schemes has been discussed in the references given above. Figure 8.3 provides a graphical representation of the grids used with the four schemes.

Within the Born-Oppenheimer approximation, vibrational states are determined by solving the nuclear Schrödinger equation, which, in terms of normal coordinates, reads:

$$\mathcal{H}\Psi_s(\mathbf{Q}) = E_s\Psi_s(\mathbf{Q}) \; , \tag{8.29}$$

where $\Psi_s(\mathbf{Q})$ is the vibrational wavefunction of the s-th vibrational state and E_s the corresponding energy. By setting the rotational angular momentum to zero and by neglecting rotational coupling effects, the Hamiltonian operator in Eq. (8.29) can be written as:

$$\mathcal{H} = \sum_{i=1}^{M} -\frac{1}{2}\frac{\partial^2}{\partial Q_i^2} + V(\mathbf{Q}) \; , \tag{8.30}$$

where $V(\mathbf{Q})$ is the usual Born-Oppenheimer potential energy surface (PES) in the basis of mass-weighted normal coordinates. As discussed above, the description of the potential term in the Hamiltonian is a computationally challenging task. Here, we expand the PES in a Taylor's series centered at the equilibrium nuclear configuration as follows:

$$\begin{aligned} V(\mathbf{Q}) &= \frac{1}{2}\sum_{i=1}^{M}\omega_i^2 Q_i^2 + \frac{1}{3!}\sum_{i,j,k=1}^{M}\eta_{ijk}Q_iQ_jQ_k + \\ &+ \frac{1}{4!}\sum_{i,j,k,l=1}^{M}\eta_{ijkl}Q_iQ_jQ_kQ_l \\ &+ \frac{1}{5!}\sum_{i,j,k,l,m=1}^{M}\eta_{ijklm}Q_iQ_jQ_kQ_lQ_m + \cdots \; , \end{aligned} \tag{8.31}$$

where ω_i is the harmonic frequency of the i-th vibration normal mode and where η_{ijk}, η_{ijkl} and η_{ijklm} are cubic, quartic and fifth-order force constants, respectively:

$$\eta_{ijk} = \left(\frac{\partial^3 E}{\partial Q_i \partial Q_j \partial Q_k}\right) \tag{8.32}$$

$$\eta_{ijkl} = \left(\frac{\partial^4 E}{\partial Q_i \partial Q_j \partial Q_k \partial Q_l}\right) \tag{8.33}$$

$$\eta_{ijklm} = \left(\frac{\partial^5 E}{\partial Q_i \partial Q_j \partial Q_k \partial Q_l \partial Q_m}\right) \; . \tag{8.34}$$

The inclusion of anharmonic (i.e. higher than quadratic) terms in the potential (8.31) therefore implies the evaluation of high-order energy derivatives with respect to atomic displacements. These high-order energy derivatives are computed numerically, which makes the description of the PES a computationally demanding task. For this reason, it proves crucial to devise: i) effective strategies to truncate the expansion of the PES in Eq. (8.31) so as to include only those terms contributing significantly to the description of the vibrational states of the system, and ii) efficient algorithms for the numerical evaluation of the high-order energy derivatives in Eqs. (8.32) - (8.34).

The Truncation of the PES

We include only terms up to fourth-order in the PES (namely, we use a 4T representation of the potential). Within a 4T representation, the PES can be further truncated by considering only those force constants involving a maximum of n distinct modes (namely, a nM representation of the potential). By combining the two truncation strategies introduced above, a 1M4T representation of the PES would require the evaluation of these force constants:

$$\eta_{iii}, \; \eta_{iiii} \quad \forall \; i \in M \; . \tag{8.35}$$

This representation of the PES neglects two-mode couplings and almost always results in a wrong description of the vibrational states. A popular representation of the potential is the

2M4T one, which includes all two-mode coupling force constants:

$$\begin{aligned} \eta_{iii},\ \eta_{iiii} \quad & \forall \quad i \in M \\ \eta_{ijj},\ \eta_{iij},\ \eta_{iiij},\ \eta_{ijjj},\ \eta_{iijj} \quad & \forall \quad i < j \in M\ . \end{aligned} \tag{8.36}$$

Analogously, the 3M4T representation of the PES includes the following terms:

$$\begin{aligned} \eta_{iii},\ \eta_{iiii} \quad & \forall \quad i \in M \\ \eta_{ijj},\ \eta_{iij},\ \eta_{iiij},\ \eta_{ijjj},\ \eta_{iijj} \quad & \forall \quad i < j \in M \\ \eta_{ijk},\ \eta_{iijk} \quad & \forall \quad i < j < k \in M\ . \end{aligned} \tag{8.37}$$

Here, we work in terms of 2M4T and 3M4T representations of the PES as given in Eqs. (8.36) and (8.37), respectively.

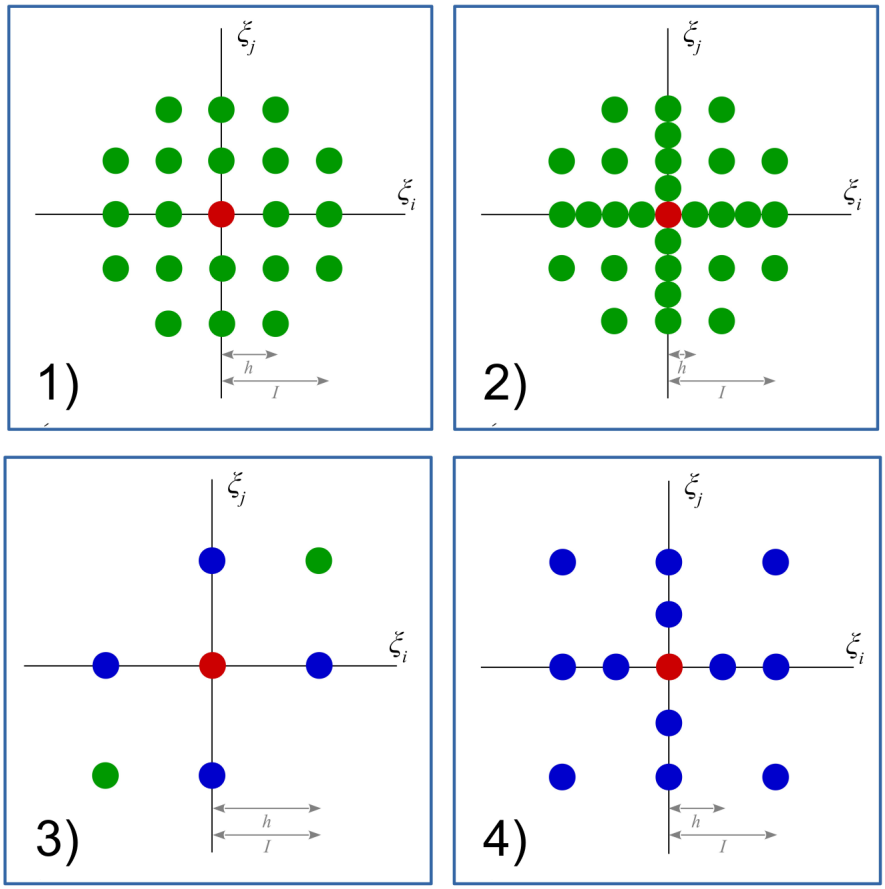


Figure 8.3: 2D grid of points defining the nuclear configurations that need to be considered in the evaluation of the adiabatic PES in its 2M4T representation for the four different numerical schemes implemented. Different colors correspond to different quantities computed for each nuclear configuration: only energy (green), energy and forces (blue), energy, forces and Hessian (red).

Numerical Evaluation of High-Order Force Constants

We have developed and implemented four different numerical approaches to compute those terms of the PES required to get a 2M4T representation, which we are going to discuss into detail below. Different approaches are characterized by a different numerical stability, accuracy

and computational cost. In order to get two-mode terms, for each pair of modes (Q_i, Q_j), a grid of points is needed where the energy (and forces, for some approaches) are computed. The shape of this grid is illustrated in Figure 8.3 for the four different schemes. The first two schemes only require the evaluation of the energy at each displaced nuclear configuration, while the last two combine information from the energy and forces.
Formal details on the four schemes are given in the references above. We recommend using the scheme 3 as it offers an optimal balance between cost and accuracy. Therefore, here we just discuss scheme 3 into detail.
By computing the analytical gradients at some configurations (only those where atoms are displaced along one normal coordinate at a time), an effective finite difference scheme has been devised, which is called EGH from the different ingredients it requires: energy, gradients and Hessian. Figure 8.3 3) shows the points needed for each pair of modes (Q_i, Q_j), where some nuclear configurations only require the energy to be evaluated while others require energy and gradients. The Hessian matrix is computed just at the equilibrium nuclear configuration to get the harmonic normal modes and frequencies. For each pair of modes, all the terms of the 2M4T representation of the PES in Eq. (8.36) can be obtained from the following finite difference relations:

$$\eta_{iii} = \frac{1}{s_i^2}\left(G^i_{-1} - 2G^i_0 + G^i_1\right) \tag{8.38}$$

$$\eta_{iij} = \frac{1}{s_i^2}(G^j_{-1,0} - 2G^j_{0,0} + G^j_{1,0}) \tag{8.39}$$

$$\eta_{iiii} = \frac{3}{s_i^3}\left(G^i_1 - 2s_i\omega_i - G^i_{-1}\right) \tag{8.40}$$

$$\eta_{iiij} = \frac{3}{s_i^3}(G^j_{1,0} - G^j_{-1,0}) \tag{8.41}$$

$$\begin{aligned}\eta_{iijj} = & -\frac{1}{2s_i^2 s_j^2}(8E_{0,0} - 4E_{-1,-1} - 4E_{1,1} + \\ & -s_j G^j_{0,-1} + s_j G^j_{0,1} - s_i G^i_{-1,0} + s_i G^i_{1,0} + \\ & -4s_j G^j_{-1,0} + 4s_j G^j_{1,0} - 4s_i G^i_{0,-1} + 4s_i G^i_{0,1} + \\ & +2s_i^2\omega_i + 2s_j^2\omega_j)\ ,\end{aligned} \tag{8.42}$$

where $s_i = h/\sqrt{\omega_i}$ and $s_j = h/\sqrt{\omega_j}$ are the adaptive steps among the points of the grid along the Q_i and Q_j normal coordinates (see the references above for more details on how the step h is defined) and where $E_{a,b}$ is the energy computed at a nuclear configuration displaced by $as_iQ_i + bs_jQ_j$ from the equilibrium one. For those terms of the PES involving only one mode, a more compact notation is used where E_a is the energy of a nuclear configuration displaced by as_iQ_i from the equilibrium one. $G^i_{a,b}$ is the gradient with respect to Q_i computed at a nuclear configuration displaced by $as_iQ_i + bs_jQ_j$ from the equilibrium one (analogously, $G^j_{a,b}$ is the gradient with respect to Q_j computed at the same nuclear configuration). For those terms of the PES involving only one mode, a more compact notation is used where G^i_a is the gradient with respect to Q_i of a nuclear configuration displaced by as_iQ_i from the equilibrium one.

Keywords

In order to activate the calculation of those cubic and quartic anharmonic terms of the PES discussed above, the **ANHAPES** keyword has to be used, within the **FREQCALC** input block. Indeed, first harmonic frequencies must be computed in order to get the normal modes of the system, which constitute the starting point of the anharmonic treatment.
The keyword **ANHAPES** allows the user to manually define the M modes to be considered for the anharmonic treatment. That is, the user can decide to treat anharmonically just a sub-set of the whole set of normal modes of the system. The keyword **ANHAPES** also allows the user to select which of the four numerical schemes to be used in the evaluation of the PES and to set the step h.
An input example reads like:

```
...Optimized Geometry Input...
FREQCALC
[Optional keywords of the harmonic frequency calculation]
ANHAPES
9
7 8 9 10 11 12 13 14 15
3 0.9
END   [End of FREQCALC block]
END   [End of Geometry block]
... Basis Set...
```

This input will run the harmonic frequency calculation (building and diagonalization of the Hessian matrix) followed by the numerical evaluation of the anharmonic terms of the PES. In the input above, 9 modes have been selected for the anharmonic calculation (modes 7 8 9 10 11 12 13 14 15 in the list obtained at the end of the harmonic calculation). The numerical scheme 3 has been selected and a step of $h = 0.9$ has been set (this is the recommended value).

Note: translational (and rotational for molecular systems) modes must not be included in the list of modes to be considered for the anharmonic treatment (i.e. the first 3 modes for 3D systems, up to the first 6 modes for non-linear 0D systems).

In case the harmonic frequencies have been computed in a previous run and all the necessary files have been saved for a restart (including the FREQINFO.DAT file), the anharmonic calculation can be run by reading the info about the previously computed Hessian as:

```
...Optimized Geometry Input...
FREQCALC
RESTART
ANHAPES
9
7 8 9 10 11 12 13 14 15
3 0.9
END   [End of FREQCALC block]
END   [End of Geometry block]
... Basis Set...
```

With the input above, the harmonic Hessian is read from a previous run and the calculation will jump to the evaluation of the anharmonic terms of the PES.

Info on the computed anharmonic terms of the PES is stored on the external file SCANPES.DAT, which has to be used should the calculation be restarted. For instance, assume your **ANHAPES** job was interrupted. You could restart it by providing the FREQINFO.DAT file (with the harmonic Hessian) and the SCANPES.DAT file (with those anharmonic terms of the PES already computed in the previous run before it got interrupted) and by using the **RESTPES** keyword:

```
...Optimized Geometry Input...
FREQCALC
RESTART
RESTPES
ANHAPES
9
7 8 9 10 11 12 13 14 15
3 0.9
END   [End of FREQCALC block]
END   [End of Geometry block]
... Basis Set...
```

8.12.2 The VSCF and VCI Approaches

Different approaches (of increasing complexity and accuracy) can be used to solve Eq. (8.29) numerically. In particular, several methods have been developed to progressively take into account the correlation among vibration modes, through mode-mode couplings, which are formally analogous to the hierarchy of wavefunction-based methods in electronic structure theory. The vibrational analog of the Hartree-Fock (HF) method is known as vibrational self-consistent field (VSCF) approach: a mean-field scheme where each vibrational degree of freedom interacts with an average potential over the other modes. In analogy to the definition of dynamical electron correlation, the vibrational correlation among modes is defined as the difference between the exact vibrational states and VSCF ones. In electronic structure theory, the HF solution can be used as a starting point to improve the description of the electronic wave-function, passing from a single-determinantal to a multi-determinantal representation by using either perturbative (MP2, MP4, etc.) or variational (CC, CI, etc.) approaches. In the vibration theory, starting from the reference VSCF state, the analog of the electronic Møller-Plesset perturbation theory is known as vibrational perturbation theory truncated at n-th order (VPTn), the analog of the coupled-cluster family of methods is the vibrational coupled-cluster approach (VCC), and the analog of the configuration-interaction methodology is the vibrational configuration-interaction (VCI), where mode-mode couplings are treated exactly (at least in the full-VCI limit). In CRYSTAL23 we have implemented the VSCF and VCI methods for both molecules and solids.

VSCF: Theoretical Remarks and Options

Vibrational modes are distinguishable so that the M-mode wavefunction of a given vibrational configuration $\mathbf{n}$ does not need to be antisymmetrized and can be written as a Hartree product of one-mode functions (modals):

$$\Phi^{\mathbf{n}}(Q_1, Q_2, \cdots, Q_M) \equiv \Phi^{\mathbf{n}}(\mathbf{Q}) = \prod_{i=1}^{M} \phi_i^{n_i}(Q_i) \, , \tag{8.43}$$

where $\mathbf{n} = (n_1, n_2, \cdots, n_i, \cdots, n_M)$ is the vibrational configuration vector of the quantum numbers of the M one-mode functions. For each given vibrational configuration $\mathbf{n}$, the VSCF method consists in looking for the variationally best form of the corresponding M one-mode functions. This is achieved by requiring that the expectation value of the full Hamiltonian is stationary:

$$E^{\mathbf{n}} = \langle \Phi^{\mathbf{n}} | \mathcal{H} | \Phi^{\mathbf{n}} \rangle \quad \text{where} \quad \mathcal{H} = \sum_{i=1}^{M} T_i + V(\mathbf{Q}) \, , \tag{8.44}$$

with $T_i = -1/2(\partial^2/\partial Q_i^2)$ being the one-mode kinetic energy operator.

In CRYSTAL, after the calculation of the anharmonic terms of the PES, a VSCF calculation can be activated by use of the **VSCF** keyword. The following input refers to a calculation where the harmonic frequencies, anharmonic terms of the PES and VSCF are all performed in a single run:

```
...Optimized Geometry Input...
FREQCALC
[Optional keywords of the harmonic frequency calculation]
ANHAPES
12
4 5 6 7 8 9 10 11 12 13 14 15
3 0.9
VSCF
END   [End of FREQCALC block]
END   [End of Geometry block]
... Basis Set...
```

In the example above, the anharmonic treatment was performed on 12 modes, listed in the input. If instead one wants to run a VSCF calculation on a previously computed harmonic and anharmonic PES (from files FREQINFO.DAT and SCANPES.DAT), the input would read like:

```
...Optimized Geometry Input...
FREQCALC
RESTART
RESTPES
ANHAPES
12
4 5 6 7 8 9 10 11 12 13 14 15
3 0.9
VSCF
END   [End of FREQCALC block]
END   [End of Geometry block]
... Basis Set...
```

The results of the VSCF calculation are printed in the output file below this header:

```
VVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV
VVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV
V                                                                               V
V                VIBRATIONAL SELF-CONSISTENT FIELD METHOD (VSCF)                V
V                        M. Rerat, M. Ferrabone, A. Erba                        V
V                                                                               V
VVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV
VVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV
```

For each vibrational configuration $\mathbf{n}$, an iterative self-consistent field process is performed to determine its energy $E^{\mathbf{n}}$. The evolution of the energy along the VSCF of each configuration is reported in the output, as well as the final energy and the difference with respect to that of the fundamental vibrational state (in cm^{-1}):

```
(  1,  0,  0,  0,  0,  0,  0,  0,  0)

 CYC:   0 ENERGY:  13426.7911169 CM-1 DELTA:   13426.7911169 CM-1
 CYC:   1 ENERGY:  13376.0933948 CM-1 DELTA:     -50.6977221 CM-1
 CYC:   2 ENERGY:  13374.9297536 CM-1 DELTA:      -1.1636412 CM-1
 CYC:   3 ENERGY:  13374.8546140 CM-1 DELTA:      -0.0751396 CM-1
 CYC:   4 ENERGY:  13374.8467135 CM-1 DELTA:      -0.0079005 CM-1
 CYC:   5 ENERGY:  13374.8459743 CM-1 DELTA:      -0.0007393 CM-1

 CONVERGENCE ACHIEVED AT CYCLE:   5 ENERGY:  13374.8459743 CM-1
 FUNDAMENTAL TRANSITION ENERGY:    1669.7802066 CM-1
```

The example above refers to a vibrational configuration where all modes are in their fundamental states but the first one that is singly exited.

The VSCF iterative process converges when the difference between the energies at two subsequent cycles is below a given threshold (default value of $1{\cdot}10^{-3}$ cm^{-1}). This threshold can be set by use of the **VSCFTOL** keyword. The VSCF iterative process has a Fock-mixing scheme active where at each cycle the VSCF matrix is mixed with 25% of that from the previous iteration. Such mixing fraction can be tuned by use of the **VSCFMIX** keyword. See the example below for the input syntax of these two options:

```
...Optimized Geometry Input...
FREQCALC
RESTART
RESTPES
ANHAPES
12
4 5 6 7 8 9 10 11 12 13 14 15
3 0.9
VSCF
VSCFTOL
5
VSCFMIX
40
END   [End of FREQCALC block]
END   [End of Geometry block]
... Basis Set...
```

In the input above, the convergence threshold is set to $1{\cdot}10^{-5}$ cm^{-1} and the mixing fraction to 40%.

VCI: Theoretical Remarks and Options

In the VCI method, the wave-function of each vibrational state s is written as a linear combination of M-mode wave-functions of different vibrational configurations in the form of Hartree products of modals as in Eq. (8.43):

$$\Psi_s(\mathbf{Q}) = \sum_{n=1}^{N_{\text{conf}}} A_{n,s}\Phi^{\mathbf{n}}(\mathbf{Q}) \,, \tag{8.45}$$

where the sum runs over N_{conf} configurations, each characterized by a vibrational configuration vector $\mathbf{n}$. The selection of the N_{conf} configurations determines the truncation of the VCI expansion. For each vibrational state s, the corresponding VCI wave-function and energy are obtained by solving the corresponding Schrödinger equation $\mathcal{H}\Psi_s = E_s\Psi_s$. The VCI method can be expressed in matrix form as follows: $\mathbf{HA} = \mathbf{AE}$, where $\mathbf{A}$ is the squared matrix containing, column-wise, the coefficients $A_{n,s}$ of the eigenvectors, $\mathbf{E}$ is the diagonal matrix of the eigenvalues and $\mathbf{H}$ is the VCI Hamiltonian matrix (of size $N_{\text{conf}} \times N_{\text{conf}}$), whose elements are:

$$H_{m,n} = \langle\Phi^{\mathbf{m}}|\mathcal{H}|\Phi^{\mathbf{n}}\rangle \,. \tag{8.46}$$

The VCI method therefore reduces to the construction and diagonalization of the VCI Hamiltonian matrix, from which all vibrational states are simultaneously determined. VSCF solutions can be used to express the modals in the VCI method (according to the so-called VCI@VSCF approach).

The VCI method relies on the expansion of the wave-function of each vibrational state in terms of N_{conf} Hartree product functions describing different vibrational configurations, as introduced in Eq. (8.45). The number N_{conf} of functions used in the VCI expansion is of critical importance with regard to both the accuracy and computational cost of the method. Indeed, the larger N_{conf} the better the description of the vibrational state but also the larger the size of the VCI Hamiltonian matrix in Eq. (8.46) to be diagonalized. In particular, this latter aspect is the main limiting factor to the application of standard VCI to the study of those systems where more than just a few vibration modes need to be coupled.

Therefore, it is crucial to devise effective schemes to reduce as much as possible the configurational space used in the VCI expansion. We have implemented two such schemes, to be introduced below. The following strategies can be used:

1. The first strategy for the truncation of the VCI expansion consists in including only those vibrational configurations where there are a maximum of N_{quanta} excitation quanta involved. Formally, we can express this strategy as follows to say that only those configurations satisfying the next condition are included in the expansion:

$$\sum_{i=1}^{M} n_i \leq N_{\text{quanta}} \; ; \tag{8.47}$$

2. A second strategy that we use to truncate the VCI expansion consists in setting a maximum number of modes N_{modes} that can be simultaneously excited in a given configuration. In other words, only those vibrational configurations where there are a maximum of N_{modes} with $n_i \neq 0$ are used.

The effect of these two schemes on the truncation of the VCI expansion is documented in the references above. As regards N_{modes}, let us note that, when working in terms of 2M4T or 3M4T representations of the PES, the VCI description converges for $N_{\text{modes}} = 3$ and $N_{\text{modes}} = 4$, respectively.

In CRYSTAL, after the calculation of the anharmonic terms of the PES, a VCI calculation can be activated by use of the **VCI** keyword. By activating a VCI calculation, a VSCF calculation will be automatically performed before it. The **VCI** keyword requires the setting of three mandatory parameters. The following input refers to a calculation where the harmonic frequencies, anharmonic terms of the PES, VSCF and VCI are all performed in a single run:

```
...Optimized Geometry Input...
FREQCALC
[Optional keywords of the harmonic frequency calculation]
ANHAPES
12
4 5 6 7 8 9 10 11 12 13 14 15
3 0.9
VCI
6 3
1
END   [End of FREQCALC block]
END   [End of Geometry block]
... Basis Set...
```

Note that even without insertion of the **VSCF** keyword, a VSCF calculation will be performed anyway before the VCI one. In this way, the VSCF solutions can be used as an initial guess of the VCI treatment. As mentioned above, the **VCI** keyword requires the setting of three mandatory parameters: the first parameter in the first row below the keyword sets N_{quanta} (6 in the example above), the second parameter in the first row below the keyword sets N_{modes} (3 in the example above), the only parameter in the second row below the keyword can be either 0 or 1 and governs the initial guess for the VCI treatment (0 implies a harmonic initial guess, 1 implies a VSCF initial guess). Use of the VCI@VSCF approach improves the convergence of the VCI expansion.

If instead one wants to run a VCI calculation on a previously computed harmonic and anharmonic PES (from files FREQINFO.DAT and SCANPES.DAT), the input would read like:

```
...Optimized Geometry Input...
FREQCALC
RESTART
RESTPES
ANHAPES
12
4 5 6 7 8 9 10 11 12 13 14 15
3 0.9
VCI
6 3
1
END   [End of FREQCALC block]
END   [End of Geometry block]
... Basis Set...
```

The results of the VCI calculation are printed in the output file below this header:

```
 ********************************************************************************
 ********************************************************************************
 *                                                                              *
 *                  VIBRATIONAL CONFIGURATION INTERACTION (VCI)                 *
 *                A. Erba, Ph. Carbonniere, M. Ferrabone, M. Rerat              *
 *                                                                              *
 ********************************************************************************
 ********************************************************************************
```

For each vibrational state s, its energy (relative to that of the fundamental vibrational state) is given in the output file, $E_s - E_0$ (in cm^{-1}). The four most dominant vibrational configurations $\mathbf{n}$ involved in each state are listed along with their respective coefficients $A_{n,s}$:

```
VCI STATE (     2)  ENE - ZPE:    1674.00 CM-1

 MOST IMPORTANT CONFIGURATIONS:

 CONFs:              3              2             27              4
 COEFs:   0.704041E+00   0.696658E+00   0.834075E-01   0.177008E-01
```

Chapter 9

Harmonic and Quasi-Harmonic Thermodynamics

Within the frame of standard quantum-chemical methods, temperature can be modeled by explicitly treating the lattice dynamics. To do so, a harmonic approximation (HA) to the lattice potential is commonly adopted, which proves extremely effective in predicting vibration properties (such as Infrared and Raman spectroscopic features, for instance) of a variety of materials, particularly so when very light atoms are not involved.
However, when the lattice dynamics of a crystal is solved within the purely HA, vibration frequencies are described as independent of interatomic distances and the corresponding vibrational contribution to the internal energy of the crystal turns out to be independent of volume. It follows that, within such an assumption, a variety of physical properties would be wrongly described: thermal expansion would be null, elastic constants would not depend on temperature, constant-pressure and constant-volume specific heats would coincide with each other, thermal conductivity would be infinite as well as phonon lifetimes, etc.
A simple and powerful approach for correcting most of the above mentioned deficiencies of the HA is represented by the so-called Quasi-Harmonic Approximation (QHA), which introduces the missing volume dependence of phonon frequencies by retaining the harmonic expression for the Helmholtz free energy. This approach allows for the natural combination of temperature and pressure effects. Depending on the temperature and pressure ranges to be explored, the lattice dynamics has to be solved at few volumes in the expansion and compression regime, respectively.
Harmonic vibration frequencies and harmonic thermodynamic functions are computed by using the **FREQCALC** keyword since the CRYSTAL03 version of the program.
A fully-automated algorithm has been implemented in the CRYSTAL17 program for computing a wealth of quasi-harmonic thermal properties of solids, which is activated by use of the keyword **QHA** (to be inserted at the end of the geometry input block), which opens a sub-block where optional keywords can be inserted and that must be closed by a keyword **END**. Default values are set for all computational parameters so that the simplest input reads as:

```
Title
Geometry input block
QHA                      # Keyword to perform a Quasi-harmonic calculation
[Optional sub-keywords]
END                      # End of the Quasi-harmonic input block
END                      # End of the Geometry input block
Basis set block
...
```

A direct-space approach is used in CRYSTAL for sampling phonon-dispersion, which requires supercells to be used. While harmonic thermodynamic properties generally require large supercells to be used for their convergence, quasi-harmonic properties show a much faster convergence. However, in some cases (particularly when small cell systems are studied), also

quasi-harmonic properties must be computed on supercells of the original primitive one, to make sure of their convergence, by use of the keyword **SCELPHONO** to be inserted right before the **QHA** one:

```
Title
Geometry input block
SCELPHONO               # Keyword to build a supercell
2 0 0                   # supercell expansion matrix
0 2 0                   # supercell expansion matrix
0 0 2                   # supercell expansion matrix
QHA                     # Keyword to perform a Quasi-harmonic calculation
[Optional sub-keywords]
END                     # End of the Quasi-harmonic input block
END                     # End of the Geometry input block
Basis set block
...
```

In the example above, an isotropic 2×2×2 supercell is considered. With a quasi-harmonic calculation, a wealth of quasi-harmonic physical quantities are computed:

- Isotropic and anisotropic thermal expansion coefficients;
- P-V-T equation-of-state;
- Temperature-dependent elastic bulk modulus (both isothermal and adiabatic);
- Difference between constant-pressure and constant-volume specific heats;
- Combined effect of pressure and temperature on structural, thermodynamic and average elastic properties;
- Grüneisen parameters;

References to be cited when using this module:

A. Erba, *J. Chem. Phys.*, **141**, 124115 (2014)
A. Erba, M. Shahrokhi, R. Moradian and R. Dovesi, *J. Chem. Phys.*, **142**, 044114 (2015)
A. Erba, J. Maul, M. De la Pierre and R. Dovesi, *J. Chem. Phys.*, **142**, 204502 (2015)
A. Erba, J. Maul, R. Demichelis and R. Dovesi, *Phys. Chem. Chem. Phys.*, **17**,11670-11677 (2015)

9.1 A Few Theoretical Remarks

9.1.1 Phonon Dispersion

The calculation of the harmonic thermodynamic properties of crystals requires the knowledge of phonon modes over the complete first Brillouin zone (FBZ) of the system; phonons at points different from Γ can be obtained by building a supercell (SC) of the original unit cell, following a direct-space approach.
Given the usual reciprocity between direct and reciprocal spaces, the size of the adopted SC (within the direct space approach) corresponds to the sampling of the FBZ. Use of the primitive cell would allow for the description of Γ modes only while a $3 \times 3 \times 3$ SC would correspond to a $3 \times 3 \times 3$ mesh of **k**-points in the FBZ, for instance. Increasing the size of the SC simply corresponds to increasing the sampling of the phonon dispersion in the FBZ in reciprocal space.

9.1.2 Harmonic Lattice Dynamics

To each **k**-point in the first Brillouin zone, 3M harmonic oscillators (i.e. phonons) are associated, which are labeled by a phonon band index p ($p = 1, \ldots, 3M$) and whose energy levels are given by the usual harmonic expression:

$$\varepsilon_m^{p,\mathbf{k}} = \left(m + \frac{1}{2} \right) \omega_{\mathbf{k}p} \; , \tag{9.1}$$

where m is an integer, $\omega_{\mathbf{k}p} = 2\pi\nu_{\mathbf{k}p}$ and M is the number of atoms per primitive cell. The overall vibrational canonical partition function of a crystal, $Q_{vib}(T)$, at a given temperature T, can be expressed as follows:

$$Q_{vib}(T) = \prod_{\mathbf{k}} \prod_{p=1}^{3M} \sum_{m=0}^{\infty} \exp\left[-\frac{\varepsilon_m^{p,\mathbf{k}}}{k_B\, T} \right] \; , \tag{9.2}$$

where k_B is Boltzmann's constant. According to standard statistical mechanics, thermodynamic properties of crystalline materials such as entropy, $S(T)$, and thermal contribution to the internal energy, $\mathcal{E}(T)$, can be expressed as:

$$S(T) = k_B T \left(\frac{\partial \log(Q_{vib})}{\partial T} \right) + k_B \log(Q_{vib}) \; , \tag{9.3}$$

$$\mathcal{E}(T) = k_B T^2 \left(\frac{\partial \log(Q_{vib})}{\partial T} \right) \; . \tag{9.4}$$

From the above expression for $\mathcal{E}(T)$, the constant-volume specific heat, $C_V(T)$, can also be computed according to $C_V(T) = \partial\mathcal{E}(T)/\partial T$. By casting equation (9.2) into equations (9.3) and (9.4) one gets the following harmonic expressions:

$$S(T) = k_B \sum_{\mathbf{k}p} \left[\frac{\hbar\omega_{\mathbf{k}p}}{k_B T \left(e^{\frac{\hbar\omega_{\mathbf{k}p}}{k_B T}} - 1 \right)} - \log(1 - e^{-\frac{\hbar\omega_{\mathbf{k}p}}{k_B T}}) \right] \tag{9.5}$$

$$\mathcal{E}(T) = \sum_{\mathbf{k}p} \hbar\omega_{\mathbf{k}p} \left[\frac{1}{2} + \frac{1}{e^{\frac{\hbar\omega_{\mathbf{k}p}}{k_B T}} - 1} \right] \tag{9.6}$$

$$C_V(T) = \sum_{\mathbf{k}p} \frac{(\hbar\omega_{\mathbf{k}p})^2}{k_B T^2} \frac{e^{\frac{\hbar\omega_{\mathbf{k}p}}{k_B T}}}{\left(e^{\frac{\hbar\omega_{\mathbf{k}p}}{k_B T}} - 1 \right)^2} \tag{9.7}$$

An explicit harmonic expression of Helmholtz's free energy, $F(T)$, can also be derived (see below).

9.1.3 Quasi-harmonic Quantities

The limitations of the HA are well-known: zero thermal expansion, temperature independence of elastic constants and bulk modulus, equality of constant-pressure and constant-volume specific heats, infinite thermal conductivity and phonon lifetimes, etc. The simplest way for correcting most of the above mentioned deficiencies of the HA is represented by the QHA, according to which the Helmholtz free energy of a crystal is written retaining the same harmonic expression but introducing an explicit dependence of vibration phonon frequencies on volume:

$$F^{QHA}(T,V) = U_0(V) + F_{vib}^{QHA}(T,V) \; , \tag{9.8}$$

where $U_0(V)$ is the zero-temperature internal energy of the crystal without any vibrational contribution (a quantity commonly accessible to standard *ab initio* simulations via volume-constrained geometry optimizations) and the vibrational part reads:

$$F_{vib}^{QHA}(T,V) = E_0^{ZP}(V) + k_B T \sum_{\mathbf{k}p} \left[\ln\left(1 - e^{-\frac{\hbar\omega_{\mathbf{k}p}(V)}{k_B T}} \right) \right] \; , \tag{9.9}$$

where $E_0^{ZP}(V) = \sum_{\mathbf{k}p} \hbar\omega_{\mathbf{k}p}(V)/2$ is the zero-point energy of the system. The equilibrium volume at a given temperature T, $V(T)$, is obtained by minimizing $F^{QHA}(V;T)$ with respect to volume V, keeping T as a fixed parameter. A volumetric thermal expansion coefficient $\alpha_V(T)$ can be defined as:

$$\alpha_V(T) = \frac{1}{V(T)} \left(\frac{\partial V(T)}{\partial T} \right)_{P=0} . \tag{9.10}$$

For cubic crystals, a linear thermal expansion coefficient $\alpha_l(T)$ is commonly considered which is simply $\alpha_l(T) = \alpha_V(T)/3$. Directional thermal expansion coefficients $\alpha_x(T) = 1/x(T)[\partial x(T)/\partial T]$ can also be computed where x can be the a, b or c lattice parameter. The anisotropy of the thermal expansion is obtained by optimizing the lattice parameters as a function of the purely internal energy E at different volumes. A more accurate optimization of the lattice parameters with respect to the free energy F could be considered, which, however, would require a much larger set of calculations to be performed.
The temperature-dependent isothermal bulk modulus of the system, $K_T(T)$, can be obtained as an isothermal second derivative of equation (9.8) with respect to the volume:

$$K_T(T) = V(T) \left(\frac{\partial^2 F^{QHA}(V;T)}{\partial V^2} \right)_T . \tag{9.11}$$

Some experimental techniques for measuring the bulk modulus of a crystal involve elastic waves and are characterized by very short time-scales that prevent the system from reaching a thermal equilibrium; in these cases, an adiabatic bulk modulus, K_S, is measured. Adiabatic and isothermal bulk moduli do coincide with each other at zero temperature only, K_S always being larger than K_T at any finite temperature. The QHA also offers a way to compute the adiabatic bulk modulus from the isothermal one, given that:

$$K_S = K_T + \frac{\alpha_V^2 V T K_T^2}{C_V} , \tag{9.12}$$

where the dependence of all quantities on temperature is omitted for clarity sake.
One of the powerful advantages of the QHA is that of allowing for a natural combination of pressure and temperature effects on structural and elastic properties of materials. By differentiating equation (9.9) with respect to the volume and changing sign, the thermal pressure is obtained:

$$P(V;T) = -\frac{\partial F^{QHA}(V;T)}{\partial V} . \tag{9.13}$$

The description of the isothermal bulk modulus of the system at simultaneous high-temperatures and high-pressures, $K_T(P,T)$, can be obtained as an isothermal second derivative of equation (9.9) with respect to the volume and by exploiting relation (9.13):

$$K_T(P,T) = V(P,T) \left(\frac{\partial^2 F^{QHA}(V(P,T);T)}{\partial V(P,T)^2} \right)_T . \tag{9.14}$$

Let us finally recall that the QHA allows for computing the difference between constant-pressure and constant-volume specific heats as follows:

$$C_P(T) - C_V(T) = \alpha_V^2(T) K(T) V(T) T . \tag{9.15}$$

9.1.4 Grüneisen Simplified Model

Grüneisen formalism assumes a linear dependence of vibration phonon frequencies on volume, in the vicinity of the zero temperature equilibrium volume V_0. The key quantity is here represented by the dimensionless mode Grüneisen parameter, which for each phonon reads:

$$\gamma_{\mathbf{k}j} = -\frac{V_0}{\omega_{\mathbf{k}j}(V_0)} \left(\frac{\partial \omega_{\mathbf{k}j}(V)}{\partial V} \right)_{V=V_0} . \tag{9.16}$$

An overall thermal Grüneisen parameter, $\gamma(T)$ can be defined as the weighted average of the various mode Grüneisen parameters in terms of the corresponding mode contribution to the specific heat:

$$\gamma(T) = \sum_{\mathbf{k}j} \frac{\gamma_{\mathbf{k}j}\, C_{V,\mathbf{k}j}(T)}{C_{V,\mathbf{k}j}(T)} \ . \tag{9.17}$$

Within Grüneisen approach, the volumetric thermal expansion is given by:

$$\alpha_V^{gru}(T) = \frac{1}{KV_0} \sum_{\mathbf{k}j} \gamma_{\mathbf{k}j}\, C_{V,\mathbf{k}j}(T) \ . \tag{9.18}$$

9.2 The Automated Algorithm: Some Computational Parameters

A fully-automated algorithm for computing quasi-harmonic thermal properties of solids has been implemented in the CRYSTAL program, which relies on the direct fitting of individual phonon frequencies, $\omega_{\mathbf{k}p}(V)$, versus volume. Once phonon frequencies are known continuously as a function of volume, all other properties can indeed be analytically derived through the expressions reported above. In order to uniquely determine the continuity of phonon frequencies among different volumes, correctly accounting for possible crossings, scalar products of the corresponding normal modes are performed. The definition of both the explored volume range and the number of considered volumes has been devised so as to make it as black-box as possible.
Here, this fully-automated computational procedure is briefly sketched and the main computational parameters one can play with are introduced:

- The starting geometrical structure of the system must be the zero-temperature, zero-pressure one. One can either start from a previously fully-optimized structure or use the **PREOPTGEOM** sub-keyword, within the **QHA** input block to activate an accurate initial optimization as concerns both lattice parameters and atomic positions. The zero-temperature, zero-pressure equilibrium volume V_0 is determined (note that zero-point motion is neglected at this stage);

- A volume range is defined from a -s% compression to a +$2s$% expansion with respect to V_0, where N_V equidistant volumes are considered (possible values for N_V are 4, 7 and 13, corresponding to equidistances of s%, $s/2$% and $s/4$%, respectively). At each volume, the structure is fully relaxed via a volume-constrained geometry optimization and phonon frequencies are computed. By default, the step s is set to 3% and N_V to 4, which means that vibration frequencies are actually computed at 4 volumes covering a volume range from a -3% compression to a +6% expansion. The step amplitude s can be modified with the **STEP** sub-keyword and the number of considered volumes for frequency calculation N_V with the **POINTS** sub-keyword as:

```
Title
Geometry input block
QHA                     # Keyword to perform a Quasi-harmonic calculation
STEP                    # Keyword to change volume step
3.5                     # New step [3.5%]
POINTS                  # Keyword to change number of volumes
7                       # New number of volumes
END                     # End of the Quasi-harmonic input block
END                     # End of the Geometry input block
Basis set block
...
```

This is a graphical representation of the standard definition of the volume interval and number of volumes where vibration frequencies are actually computed:

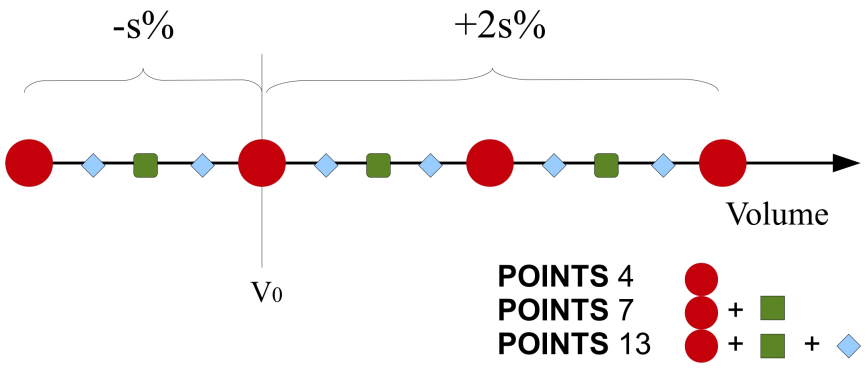


Figure 9.1: Definition of volume step and number

It is recommended to use the above standard definition for the explored volume range. However, if different volume ranges have to be explored (for instance more oriented towards compression than expansion), the sub-keyword **VRANGE** can be used, which allows for a more customized definition (the syntax is the same as for the same option in the **EOS**, equation-of-state, block);

- Once all volumes have been considered, purely electronic internal energies as a function of volume are fitted to various equations-of-state. The third-order Birch-Murnaghan is the one to be used for subsequent thermodynamic purposes. The continuity of phonon frequencies on volume is determined before individually fitting them as a function of volume with polynomial functions of different orders, up to third order (a second-order polynomial fit is used as a default). For each phonon mode p, a coefficient of determination R_p^2 is computed, which measures the goodness-of-fit (the closer R^2 to 1, the better the fitting). An average index over all modes $\overline{R}^2$ is also defined which gives an overall measure of the regularity of all computed frequencies on volume. If low values of R_p^2 are found for some modes, warnings will be printed in the output as one can not expect a good quasi-harmonic description from irregular vibration frequencies as a function of volume;

- From fitted frequencies, at any considered temperature T the Helmholtz free energy is evaluated through equations (9.8) and (9.9) at several volumes, minimized for determining the corresponding equilibrium volume $V(T)$ and fitted for getting $K(T)$ from its second derivative. By fitting $V(T)$ data to a polynomial function and by taking its temperature derivative, the thermal expansion coefficient $\alpha_V(T)$ of equation (9.10) can then be computed numerically.

All computed thermal properties are evaluated at a series of temperature values to be defined by means of the **TEMPERAT** sub-keyword (by default 50 temperatures would otherwise be considered from 10 K to 300 K). Given that some properties are obtained as numerical derivatives with respect to temperature (as the thermal expansion coefficient), at least 50 temperature values must be considered. The desired set of temperatures can be modified as:

```
Title
Geometry input block
QHA                     # Keyword to perform a Quasi-harmonic calculation
TEMPERAT                # Keyword to change temperature range
100 10 2000             # 100 temperature values from 10 K to 2000 K
END                     # End of the Quasi-harmonic input block
END                     # End of the Geometry input block
Basis set block
...
```

9.2.1 Optional Keywords

This is a list of some possible (optional) sub-keywords to be inserted after **QHA**. This is just a partial list since the main sub-keywords have been described above.

- A **RESTART** Activates a restart from a previous incomplete run. The file EOSINFO.DAT from the previous run must be provided.

- A **RESTART2** Activates a restart from a previous complete run. The file EOSINFO.DAT from the previous run must be provided. Useful when the temperature range needs to be changed, which can be done *a posteriori* at almost no additional cost.

- A **NEGLEFRE**
- ∗ NFREQ Number of frequencies to be neglected
- ∗ FREQ(1:NFREQ) List of frequencies to be neglected

9.3 QHA: A Guided Tour of the Output

In order to illustrate how to prepare an Input file for the calculation of quasi-harmonic properties with CRYSTAL and how to extract the corresponding information from the Output file, let us consider the simple crystal of Al_2O_3 corundum. The geometry block of the Input file reads as follows:

```
CORUNDUM
CRYSTAL
0 0 0
167
4.82860168 13.11488905
2
13  3.333333333333E-01 -3.333333333333E-01  1.955179055198E-02
8  -3.617448850731E-01 -3.333333333333E-01 -8.333333333333E-02
SCELPHONO
1 -1  0
0  1 -1
1  1  1
QHA
TEMPERAT
100 10 1200
END
END
...
```

The keyword **QHA** must be inserted at the end of the geometry input block and in turn opens a sub-block that must be terminated with an **END** keyword. Note that the structure given in input is assumed to be fully optimized. If not, one can perform a preliminary geometry optimization by inserting the **PREOPTGEOM** keyword after **QHA**. The calculation in this case is performed on the conventional super-cell.

9.3.1 The Equation-of-State

At 4 (by default) volumes (the equilibrium one, a compressed one and two expanded ones) the structure of the system is relaxed by means of volume-constrained geometry optimizations. At

each optimized structure, phonon vibration frequencies are computed. Once the four vibration frequency calculations have been performed, the actual quasi-harmonic elaboration starts. The internal energy of the system as a function of volume is first fitted to various equations of state and the corresponding information reported in the output (no thermal contributions are included at this stage, not even zero-point terms). See the EOS Tutorial for more details on this part:

```
EQUATION OF STATE             VOL(A^3)        E(AU)        BM(GPa)   BM PRIME
-------------------------------------------------------------------------------
MURNAGHAN 1944                264.8935     -4263.44395040  230.70     4.08
BIRCH-MURNAGHAN 1947          264.8959     -4263.44395044  230.78     4.02
POIRIER-TARANTOLA 1998        264.8978     -4263.44395047  230.83     3.98
VINET 1987                    264.8968     -4263.44395046  230.81     4.00

POLYNOMIAL FITTING            VOL(A^3)        E(AU)        BM(GPa)
-------------------------------------------------------------------------------
THIRD ORDER POLYNOMIAL        264.9108     -4263.44395072  231.21
```

For each EOS, the athermal P-V equation-of-state is given (at 0 K, with no zero-point term). This is the Birch-Murnaghan third-order one, for instance:

```
+++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
THERMODYNAMIC FUNCTIONS OBTAINED AT 0 K WITH EOS: BIRCH-MURNAGHAN 1947
+++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

V = VOLUME, P = PRESSURE, E = ENERGY, G = GIBBS FREE ENERGY, B = BULK MODULUS

V (A^3)   P (GPa)        E (a.u.)            G (a.u.)      B (GPa)

 256.87      7.56     -4263.43717037     -4262.99197328     260.73
 257.88      6.54     -4263.43880796     -4263.05211322     256.74
 258.89      5.54     -4263.44021102     -4263.11133295     252.82
 259.91      4.56     -4263.44138406     -4263.16964689     248.95
 260.92      3.60     -4263.44233147     -4263.22706920     245.14
 261.93      2.65     -4263.44305757     -4263.28361377     241.40
 262.95      1.73     -4263.44356658     -4263.33929423     237.71
 263.96      0.82     -4263.44386265     -4263.39412397     234.08
 264.97     -0.07     -4263.44394982     -4263.44811612     230.50
 265.99     -0.94     -4263.44383209     -4263.50128358     226.98
 267.00     -1.80     -4263.44351335     -4263.55363902     223.51
 268.01     -2.64     -4263.44299741     -4263.60519485     220.10
 269.03     -3.46     -4263.44228803     -4263.65596329     216.74
 270.04     -4.27     -4263.44138888     -4263.70595632     213.43
 271.05     -5.06     -4263.44030355     -4263.75518571     210.17
 272.07     -5.84     -4263.43903558     -4263.80366301     206.96
 273.08     -6.61     -4263.43758843     -4263.85139958     203.80
 274.10     -7.36     -4263.43596549     -4263.89840657     200.68
 275.11     -8.09     -4263.43417008     -4263.94469491     197.61
 276.12     -8.81     -4263.43220547     -4263.99027538     194.59
 277.14     -9.52     -4263.43007487     -4264.03515853     191.61
 278.15    -10.21     -4263.42778140     -4264.07935474     188.68
 279.16    -10.89     -4263.42532814     -4264.12287421     185.79
 280.18    -11.56     -4263.42271812     -4264.16572696     182.94
 281.19    -12.22     -4263.41995429     -4264.20792284     180.14
```

9.3.2 Phonon Modes Continuity on Volume

The QHA results are printed after the following header:

```
*****************************************************************************
*****************************************************************************
*****************************************************************************

                          QUASI-HARMONIC APPROXIMATION
                                 A. Erba (2014)

      Please do cite the following references when using this module:

A. Erba, J. Chem. Phys., 141, 124115 (2014)
A. Erba, M. Shahrokhi, R. Moradian, R. Dovesi, J. Chem. Phys., 142, 044114 (2015)
A. Erba, J. Maul, M. De la Pierre, R. Dovesi, J. Chem. Phys., 142, 204502 (2015)
A. Erba, J. Maul, M. Itou, R. Dovesi, Y. Sakurai, Phys. Rev. Lett., 115, 117402 (2015)

*****************************************************************************
*****************************************************************************
*****************************************************************************
```

As a first step, the continuity of phonon modes on volume is established by computing scalar products of the corresponding eigenvectors. The evolution of each vibration frequency on volume is reported. If the program fails at establishing the required continuity, warnings are printed:

```
ESTABLISHING CONTINUITY OF NORMAL MODES ON VOLUME

ANALYSIS OF NORMAL MODE SCALAR PRODUCTS:
WHEN TWO OR MORE SCALAR PRODUCTS ARE CLOSER THAN 0.4 THEY ARE REPORTED
AS THEY MAY CAUSE A MISMATCH.
OTHERWISE, THE CONTINUITY IS CONSIDERED AS SAFE


FOUND CONTINUITY OF FREQUENCIES WITH VOLUME:


FREQUENCY #     1
   VOL (A^3)   FREQ (cm^-1)   IRREP

    256.8676      0.0000       4
    264.8122      0.0000       4
    272.9187      0.0000       4
    281.1886      0.0000       4

FREQUENCY #     2
   VOL (A^3)   FREQ (cm^-1)   IRREP

    256.8676      0.0000       3
    264.8122      0.0000       3
    272.9187      0.0000       3
    281.1886      0.0000       3
```

```
FREQUENCY #     3
   VOL (A^3)   FREQ (cm^-1)   IRREP

    256.8676      0.0000       3
    264.8122      0.0000       3
    272.9187      0.0000       3
    281.1886      0.0000       3

FREQUENCY #     4
   VOL (A^3)   FREQ (cm^-1)   IRREP

    256.8676    154.0503       3
    264.8122    146.8375       3
    272.9187    139.3457       3
    281.1886    131.6465       3

FREQUENCY #     5
   VOL (A^3)   FREQ (cm^-1)   IRREP

    256.8676    154.0503       3
    264.8122    146.8375       3
    272.9187    139.3457       3
    281.1886    131.6465       3

FREQUENCY #     6
   VOL (A^3)   FREQ (cm^-1)   IRREP

    256.8676    154.0503       1
    264.8122    146.8375       1
    272.9187    139.3457       1
    281.1886    131.6465       1

FREQUENCY #     7
   VOL (A^3)   FREQ (cm^-1)   IRREP

    256.8676    154.0503       1
    264.8122    146.8375       1
    272.9187    139.3457       1
    281.1886    131.6465       1

FREQUENCY #     8
   VOL (A^3)   FREQ (cm^-1)   IRREP

    256.8676    264.1931       4
    264.8122    256.7193       4
    272.9187    249.7118       4
    281.1886    242.5474       4

FREQUENCY #     9
   VOL (A^3)   FREQ (cm^-1)   IRREP

    256.8676    264.1931       2
    264.8122    256.7193       2
    272.9187    249.7118       2
    281.1886    242.5474       2
    ...
```

Each frequency is then fitted as a function of volume with different polynomial functions and a coefficient of determination is computed to measure the goodness-of-fit (the closer to 1, the better). If low values of this parameter are obtained, warnings are printed in the output.

```
 FITTING FREQUENCIES VERSUS VOLUME FOR MAX. NUMBER OF POINTS
 THE COEFFICIENT OF DETERMINATION R^2 IS REPORTED
 FOR DIFFERENT POLYNOMIAL FITTING ORDERS.
 ONLY ONE FREQ PER MODE IS REPORTED.

 FREQ #    R^2 LINEAR   R^2 QUADRATIC   R^2 CUBIC

      4     0.999968       0.999999     1.000000
      6     0.999968       0.999999     1.000000
      8     0.999580       0.999921     1.000000
      9     0.999580       0.999921     1.000000
     10     0.999960       0.999998     1.000000
     12     0.999960       0.999998     1.000000
     14     0.997897       0.999994     1.000000
     15     0.997897       0.999994     1.000000
     16     0.999307       1.000000     1.000000
     17     0.999916       0.999997     1.000000
     19     0.999208       0.999990     1.000000
     20     0.999208       0.999990     1.000000
     21     0.997294       0.999999     1.000000
     22     0.999971       0.999998     1.000000
     24     0.999790       0.999994     1.000000
     26     0.999790       0.999994     1.000000
     28     0.998933       0.999980     1.000000
....
....
     84     0.999068       0.999992     1.000000
     85     0.997886       0.999999     1.000000
     87     0.998882       1.000000     1.000000
     88     0.998882       1.000000     1.000000
     89     0.998838       0.999979     1.000000
     90     0.998838       0.999979     1.000000

 AVERAGE    0.998794       0.999971     1.000000
```

In this case, excellent coefficients of determination are obtained (to second order; to third order they are ideal just because we fit over 4 points), which means that vibration frequencies are varying very regularly with volume, a necessary prerequisite to a good quasi-harmonic description of the system.

9.3.3 Grüneisen Model

Some thermal properties are then printed as obtained from the simplified Grüneisen model. The thermal expansion coefficient, the overall Grüneisen parameter, and the difference between constant-pressure and constant-volume specific heat are reported as a function of temperature:

```
 ******************************************************************************
 THERMAL PROPERTIES FROM GRUNEISEN PARAMETERS
 ******************************************************************************


 NUMBER OF POINTS: 4

    ORDER OF POLYNOMIALS: 2
```

```
   T (K)     ALPHA (1/K)     GRUNEISEN    Cp-Cv (J/mol*K)

   10.00      0.44781E-12        1.662         0.0000
   22.02      0.94330E-08        1.661         0.0000
   34.04      0.12057E-06        1.650         0.0000
   46.06      0.37164E-06        1.622         0.0002
   58.08      0.73774E-06        1.582         0.0012
   70.10      0.12272E-05        1.540         0.0039
   82.12      0.18530E-05        1.502         0.0104
   94.14      0.26122E-05        1.473         0.0236
  106.16      0.34862E-05        1.450         0.0475
  118.18      0.44475E-05        1.433         0.0860
  130.20      0.54659E-05        1.419         0.1432
  142.22      0.65134E-05        1.409         0.2221
  154.24      0.75658E-05        1.401         0.3249
  166.26      0.86042E-05        1.395         0.4530
  178.28      0.96141E-05        1.389         0.6065
  190.30      0.10585E-04        1.385         0.7848
  202.32      0.11512E-04        1.381         0.9867
  214.34      0.12389E-04        1.378         1.2108
  226.36      0.13216E-04        1.376         1.4550
  238.38      0.13992E-04        1.373         1.7175
  250.40      0.14718E-04        1.371         1.9963
  262.42      0.15396E-04        1.370         2.2894
  274.44      0.16029E-04        1.368         2.5951
  286.46      0.16618E-04        1.367         2.9115
  298.48      0.17167E-04        1.366         3.2373
  310.51      0.17677E-04        1.364         3.5710

   ...
   ...
```

9.3.4 Main Results: Thermal Properties from Helmholtz Free Energy

Then, the main results of the quasi-harmonic calculation are reported: thermal properties as derived from the Helmholtz free energy. Several properties are reported as a function of temperature: volume (V), isothermal bulk modulus (Bt), thermal expansion coefficient (ALPHA), difference between constant-pressure and constant-volume specific heats (Cp-Cv), adiabatic bulk modulus (Bs), difference between constant-pressure and constant-volume entropy (Sp-Sv):

```
 ***************************************************************************
 THERMAL PROPERTIES FROM HELMHOLTZ FREE ENERGY
 ***************************************************************************


 ZERO-POINT ENERGY INCLUDED


 NUMBER OF POINTS: 4

    ORDER OF POLYNOMIALS: 2
```

T (K)	V (Ang^3)	Bt (GPa)	ALPHA (1/K)	Cp-Cv (J/mol*K)	Bs (GPa)	Sp-Sv (J/mol*K)
10.00	267.4132	225.407	0.00000E+00	0.000	225.407	0.000
22.02	267.4132	225.406	0.24715E-07	0.000	225.407	0.004
34.04	267.4132	225.405	0.24684E-06	0.000	225.411	0.046
46.06	267.4148	225.395	0.49375E-06	0.000	225.405	0.139
58.08	267.4164	225.378	0.93795E-06	0.002	225.402	0.274
70.10	267.4196	225.350	0.13823E-05	0.005	225.387	0.456
82.12	267.4259	225.300	0.20736E-05	0.013	225.364	0.689
94.14	267.4323	225.241	0.29126E-05	0.029	225.341	0.972
106.16	267.4434	225.155	0.37520E-05	0.054	225.292	1.302
118.18	267.4577	225.046	0.47896E-05	0.098	225.240	1.670
130.20	267.4736	224.923	0.58747E-05	0.163	225.181	2.064
142.22	267.4942	224.774	0.69609E-05	0.250	225.104	2.479
154.24	267.5197	224.602	0.81466E-05	0.370	225.020	2.906
166.26	267.5467	224.419	0.92323E-05	0.512	224.926	3.338
178.28	267.5784	224.216	0.10268E-04	0.679	224.814	3.772
190.30	267.6134	223.998	0.11354E-04	0.886	224.703	4.206
202.32	267.6515	223.769	0.12242E-04	1.094	224.567	4.636
214.34	267.6928	223.527	0.13178E-04	1.341	224.433	5.061
226.36	267.7357	223.281	0.14064E-04	1.612	224.298	5.479
238.38	267.7834	223.018	0.14950E-04	1.916	224.158	5.895
250.40	267.8327	222.751	0.15788E-04	2.242	224.016	6.304
262.42	267.8851	222.475	0.16524E-04	2.572	223.858	6.709
274.44	267.9391	222.194	0.17164E-04	2.899	223.688	7.107
286.46	267.9963	221.905	0.17849E-04	3.269	223.526	7.503
...						
...						

Directional thermal expansion coefficients are then reported, as well as the a, b and c lattice parameters as a function of temperature:

```
*******************************************************************************
LINEAR THERMAL EXPANSION OF CONVENTIONAL LATTICE PARAMETERS
*******************************************************************************
```

T (K)	a (Ang)	ALPHA_a (1/K)	b (Ang)	ALPHA_b (1/K)	c (Ang)	ALPHA_c (1/K)
10.00	4.8438	0.00000E+00	4.8438	0.00000E+00	13.1605	0.00000E+00
22.02	4.8438	0.79935E-08	4.8438	0.79935E-08	13.1605	0.87281E-08
34.04	4.8438	0.79835E-07	4.8438	0.79835E-07	13.1605	0.87171E-07
46.06	4.8439	0.15969E-06	4.8439	0.15969E-06	13.1605	0.17436E-06
58.08	4.8439	0.30336E-06	4.8439	0.30336E-06	13.1605	0.33123E-06
70.10	4.8439	0.44709E-06	4.8439	0.44709E-06	13.1606	0.48816E-06
82.12	4.8439	0.67066E-06	4.8439	0.67066E-06	13.1607	0.73223E-06
94.14	4.8440	0.94204E-06	4.8440	0.94204E-06	13.1608	0.10285E-05
106.16	4.8440	0.12136E-05	4.8440	0.12136E-05	13.1610	0.13248E-05
118.18	4.8441	0.15492E-05	4.8441	0.15492E-05	13.1613	0.16912E-05
130.20	4.8442	0.19002E-05	4.8442	0.19002E-05	13.1615	0.20742E-05
142.22	4.8443	0.22517E-05	4.8443	0.22517E-05	13.1619	0.24575E-05
154.24	4.8445	0.26353E-05	4.8445	0.26353E-05	13.1623	0.28759E-05
166.26	4.8446	0.29867E-05	4.8446	0.29867E-05	13.1628	0.32588E-05
178.28	4.8448	0.33218E-05	4.8448	0.33218E-05	13.1634	0.36240E-05
190.30	4.8450	0.36734E-05	4.8450	0.36734E-05	13.1640	0.40069E-05
202.32	4.8452	0.39610E-05	4.8452	0.39610E-05	13.1646	0.43198E-05
214.34	4.8455	0.42640E-05	4.8455	0.42640E-05	13.1653	0.46493E-05

```
226.36   4.8457   0.45512E-05   4.8457   0.45512E-05    13.1661   0.49614E-05
238.38   4.8460   0.48382E-05   4.8460   0.48382E-05    13.1669   0.52732E-05
250.40   4.8463   0.51098E-05   4.8463   0.51098E-05    13.1678   0.55679E-05
262.42   4.8466   0.53486E-05   4.8466   0.53486E-05    13.1687   0.58267E-05
274.44   4.8469   0.55561E-05   4.8469   0.55561E-05    13.1696   0.60514E-05
286.46   4.8473   0.57785E-05   4.8473   0.57785E-05    13.1706   0.62921E-05
298.48   4.8476   0.59855E-05   4.8476   0.59855E-05    13.1716   0.65159E-05
...
...
```

9.3.5 Thermal Pressure: The P-V-T Equation-of-State

While previous properties refer to zero pressure, from now on pressure will be included. The P-V-T EOS is explicitly given (for each T, the corresponding P-V relation is reported):

```
*****************************************************************************
PRESSURE-VOLUME-TEMPERATURE RELATION
*****************************************************************************

  T (K)     V (Ang^3)     P (GPa)        F (Ha/cell)           G (Ha/cell)

  10.00       256.868       9.828    -0.4263336025683E+04   -0.4262756991487E+04
  10.00       257.706       8.973    -0.4263337833759E+04   -0.4262807436702E+04
  10.00       258.545       8.132    -0.4263339478696E+04   -0.4262857251925E+04
  10.00       259.384       7.303    -0.4263340963054E+04   -0.4262906445377E+04
  10.00       260.222       6.488    -0.4263342289350E+04   -0.4262955025140E+04
  10.00       261.061       5.686    -0.4263343460057E+04   -0.4263002999190E+04
  10.00       261.900       4.896    -0.4263344477610E+04   -0.4263050375359E+04
  10.00       262.738       4.118    -0.4263345344399E+04   -0.4263097161377E+04
  10.00       263.577       3.353    -0.4263346062779E+04   -0.4263143364837E+04
  10.00       264.416       2.599    -0.4263346635064E+04   -0.4263188993224E+04
  10.00       265.254       1.857    -0.4263347063528E+04   -0.4263234053927E+04
  10.00       266.093       1.127    -0.4263347350410E+04   -0.4263278554182E+04
  10.00       266.931       0.408    -0.4263347497912E+04   -0.4263322501161E+04
  ...
  ...

  22.02       256.868       9.828    -0.4263336025695E+04   -0.4262756991317E+04
  22.02       257.706       8.973    -0.4263337833771E+04   -0.4262807436523E+04
  22.02       258.545       8.132    -0.4263339478709E+04   -0.4262857251732E+04
  22.02       259.384       7.303    -0.4263340963068E+04   -0.4262906445182E+04
  22.02       260.222       6.488    -0.4263342289364E+04   -0.4262955024933E+04
  22.02       261.061       5.686    -0.4263343460072E+04   -0.4263002998975E+04
  22.02       261.900       4.896    -0.4263344477626E+04   -0.4263050375125E+04
  22.02       262.738       4.118    -0.4263345344416E+04   -0.4263097161136E+04
  22.02       263.577       3.353    -0.4263346062797E+04   -0.4263143364578E+04
  22.02       264.416       2.599    -0.4263346635082E+04   -0.4263188992958E+04
  22.02       265.254       1.857    -0.4263347063548E+04   -0.4263234053632E+04
  22.02       266.093       1.127    -0.4263347350431E+04   -0.4263278553875E+04
  22.02       266.931       0.408    -0.4263347497933E+04   -0.4263322500838E+04
  ...
  ...

  34.04       256.868       9.828    -0.4263336026331E+04   -0.4262756985744E+04
  34.04       257.706       8.973    -0.4263337834427E+04   -0.4262807430739E+04
  34.04       258.545       8.132    -0.4263339479387E+04   -0.4262857245733E+04
```

```
 34.04     259.384       7.304    -0.4263340963767E+04   -0.4262906438954E+04
 34.04     260.222       6.488    -0.4263342290087E+04   -0.4262955018473E+04
 34.04     261.061       5.686    -0.4263343460818E+04   -0.4263002992271E+04
 34.04     261.900       4.896    -0.4263344478396E+04   -0.4263050368173E+04
 34.04     262.738       4.118    -0.4263345345211E+04   -0.4263097153918E+04
 34.04     263.577       3.353    -0.4263346063618E+04   -0.4263143357091E+04
 34.04     264.416       2.599    -0.4263346635930E+04   -0.4263188985184E+04
 34.04     265.254       1.858    -0.4263347064423E+04   -0.4263234045569E+04
 34.04     266.093       1.127    -0.4263347351335E+04   -0.4263278545503E+04
 34.04     266.931       0.408    -0.4263347498868E+04   -0.4263322492144E+04
 ...
 ...

 ...
 ...

 ...
 ...
```

After that, for various values of pressure, some thermal properties are reported (volume, thermal expansion coefficient and isothermal bulk modulus):

```
*****************************************************************************
 THERMAL PROPERTIES FROM HELMHOLTZ FREE ENERGY AT P =   4.37
*****************************************************************************


   T (K)    V (Ang^3)    ALPHA (1/K)     B (GPa)

  10.00     262.467     0.23273E-07     242.677
  22.02     262.467     0.29239E-07     242.677
  34.04     262.468     0.17305E-06     242.675
  46.06     262.468     0.39004E-06     242.668
  58.08     262.470     0.72023E-06     242.653
  70.10     262.473     0.11652E-05     242.628
  82.12     262.477     0.17278E-05     242.589
  94.14     262.483     0.24072E-05     242.533
 106.16     262.492     0.31923E-05     242.459
 118.18     262.503     0.40635E-05     242.365
 130.20     262.517     0.49970E-05     242.251
 142.22     262.535     0.59686E-05     242.116
 154.24     262.555     0.69562E-05     241.963
 166.26     262.579     0.79417E-05     241.792
 178.28     262.605     0.89103E-05     241.606
 190.30     262.635     0.98513E-05     241.404
 202.32     262.668     0.10757E-04     241.190
 214.34     262.703     0.11623E-04     240.964
 226.36     262.741     0.12445E-04     240.727
 238.38     262.782     0.13224E-04     240.481
 250.40     262.825     0.13959E-04     240.227
 262.42     262.870     0.14650E-04     239.966
 274.44     262.917     0.15300E-04     239.698
 286.46     262.967     0.15910E-04     239.424
 298.48     263.018     0.16483E-04     239.145
 ...
 ...
```

9.4 Quasi-Harmonic Thermo-Elasticity

The elastic linear response of a three-dimensional continuous body - such as a crystalline system - is completely determined by the elastic stiffness tensor $\boldsymbol{C}$ (a 6×6 matrix). Its components, C_{vu}, are the elastic stiffness constants of the system. The QHA provides a couple of schemes to compute the thermal dependence of such quantities: the so-called thermo-elasticity of the material.
The static elastic constants are expressed as:

$$C_{vu} = \frac{1}{V_0} \left. \frac{\partial^2 E}{\partial \eta_v \, \partial \eta_u} \right|_{\boldsymbol{\eta}=\mathbf{0}} \tag{9.19}$$

where E is the internal static energy and V_0 is the volume of the system after a standard geometry optimization - which gives the structure corresponding to the minimum of E. In CRYSTAL, a fully automated algorithm for the determination of the static elastic constants can be exploited by means of the **ELASTCON** keyword.
Instead, the expression for the thermo-elastic constants is the following one:

$$C^T_{vu}(T) = \frac{1}{V(T)} \left. \frac{\partial^2 F}{\partial \eta_v \, \partial \eta_u} \right|_{\boldsymbol{\eta}=\mathbf{0}} \tag{9.20}$$

This makes use of the Helmholtz free energy F (which of course includes the thermal contributions) and takes into account the thermal expansion of the system, $V(T)$. Equation (9.20) yields the thermo-elastic constants referring to isothermal conditions - that's why they are labelled $C^T_{vu}(T)$. In order to compute them, the Quasi-Harmonic Approximation (QHA) may be used: the phonon frequencies are computed at the harmonic level but at several unit cell volumes and strained lattice configurations.
Because the phonon frequencies are computationally expensive to determine, the full QHA model above can become quite computationally expensive. By contrast, a plain static approach as in Eq. (9.19) is not able to predict any thermal dependence of the elastic constants. A compromise to this problem is offered by the so-called Quasi-Static Approximation (QSA):

$$C^T_{vu}(T) = \frac{1}{V(T)} \left. \frac{\partial^2 E}{\partial \eta_v \, \partial \eta_u} \right|_{\boldsymbol{\eta}=\mathbf{0}} \tag{9.21}$$

This model assumes that most of the thermal effects on the elastic response can be taken into account by considering the thermal expansion of the system only; the lattice potential energy is still approximated with the internal static energy E - which does not include any thermal contribution. The result is a quick answer to the thermal dependence of the elasticity. Although this may not be as accurate as the full QHA, it is certainly much better with respect to a merely static description.
In order to implement Equation (9.20) two ingredients are needed:

- The thermal expansion of the compound, $V(T)$;
- The second-order derivative of the Helmholtz free energy with respect to the strain, $\frac{\partial^2 F}{\partial \eta^2}$

The thermal expansion of the system can be determined by an isotropic QHA calculation. This functionality can be exploited by means of the **QHA** keyword discussed above.

9.4.1 Formal Aspects of the Implementation

The stiffness constants express the elastic response of a crystal with respect to a couple of deformations. Therefore, the determination of the whole elastic tensor requires the exploration of the elastic potential energy surface of the system over all the couples of strain directions. Due to the complexity of this task, the problem must be tackled by exploring a specific direction at a time. First of all, the strain vector $\boldsymbol{\eta}$ is expressed as a product of a scalar amplitude η and a vector shape $\boldsymbol{\eta'}$:

$$\boldsymbol{\eta} = \eta \, \boldsymbol{\eta'} \tag{9.22}$$

The shape of the strain, $\boldsymbol{\eta'}$, is then kept fixed; the variable is just its amplitude η. In this way, the elastic potential energy surface is explored only in the direction defined by such strain shape: this"surface" is actually seen as a mono-dimensional curve. Its second-order derivative (which is a directional derivative with respect to the surface) will lead to a specific scalar elastic constant. The scalar elastic constants obtained from the strain direction $\boldsymbol{\eta'}$ is labelled C':

$$C'(T) = \frac{1}{V(T)} \left(\frac{\partial^2 F}{\partial \eta^2} \right)_{\boldsymbol{\eta'}} \Bigg|_{\boldsymbol{\eta}=\mathbf{0}} \tag{9.23}$$

Of course, the full elastic tensor contains the complete information of the spatial dependence of the elasticity phenomenon. Indeed, this particular scalar constant C' is just a projection of the stiffness tensor $\boldsymbol{C}$ over twice the direction $\boldsymbol{\eta'}$. Therefore, such matrix projection can be determined by the following formula (which is analogous to the projection of a vector over another one):

$$C'(T) = \sum_{vu} C_{vu}(T)\, \eta'_v\, \eta'_u \tag{9.24}$$

where the summation runs over the two indexes v and u, both ranging from 1 to 6. That is, this particular elastic constant corresponds to a linear combination of the basic ones (i.e. the elements of the elastic tensor), where the weights depend on the selected direction (strain shape).

9.4.2 Computing Thermo-Elastic Constants with CRYSTAL

In order to take into account thermal effects, a semi-automated algorithm for the thermo-elasticity has been developed, which consists of the following steps (to be run separately in sequence):

1. A full geometry optimization (keyword **OPTGEOM**), which yields the equilibrium structure;
2. An isotropic QHA calculation (keyword **QHA**), which gives the thermal expansion of the material (i.e. the volume-temperature relation $V(T)$);
3. A value of temperature of interest is selected T, and the corresponding volume $V(T)$ determined from the previous step;
4. A constant-volume geometry optimization (keyword **CVOLOPT**) is performed so that the expanded equilibrium structure at this temperature is determined;
5. A strain shape η' is selected and the actual thermo-elastic calculation can be carried out (with the keyword **THERMOELAS** to be illustrated below).

The result of this calculation will be the $C'(T)$ scalar elastic constant relative to the strain shape $\boldsymbol{\eta'}$ at the selected temperature T. By performing several calculations with different strain shapes, it is possible to determine the whole elastic tensor; by performing several calculations with different temperature values, it is possible to obtain the thermal dependence of the elastic response of the system.

The keyword **THERMOELAS** (to be inserted into the **ELASTCON** input block) requires the definition of two records: the selected temperature T (in Kelvin) and the strain shape $\boldsymbol{\eta'}$ (expressed in the Voigt notation as a vector of six integers).
NOTE: the starting geometry of a **THERMOELAS** run must be the one obtained at the end of step 4 above.
An input looks like:

```
Title
Geometry input block
ELASTCON
[Optional sub-keywords of ELASTCON]
THERMOELAS
298
1 0 1 0 0 0
[Optional sub-keywords relevant to the thermo-elastic calculation]
END
END
Basis-set block
```

In the example above, the temperature has been set to 298 K and a strain shape defined as a combination of the first and third fundamental types.
What **THERMOELAS** does is automatically explore different strained lattice configurations (strained according to $\boldsymbol{\eta'}$), relax the atomic positions within each strained configuration, compute harmonic frequencies at each strained configuration, compute the free energy, fit it and determine the corresponding $C'(T)$ elastic constant. A scheme of the explored strained configurations is provided in Figure 9.4.2. As a by-product, quasi-static elastic constants are also provided in the output.

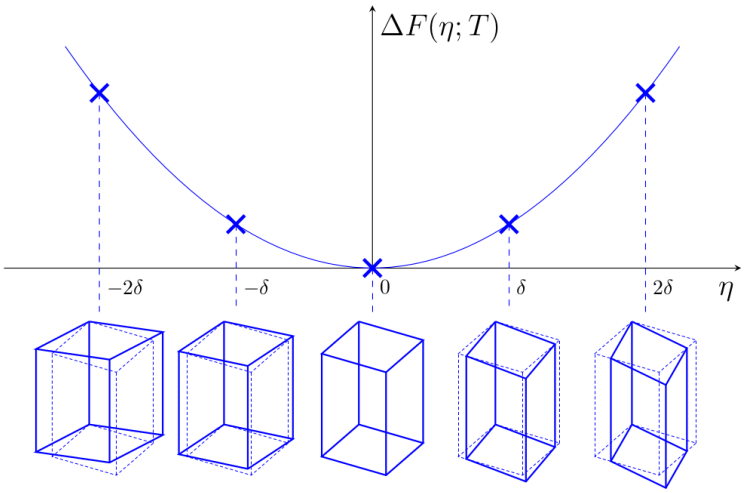


Figure 9.2: Scheme of the lattice strained configurations automatically explored by the **THERMOELAS** option, where harmonic frequencies are computed at each of them.

Users of this option are kindly reminded to cite the following papers:
M. Destefanis, C. Ravoux, A. Cossard and A. Erba, *Minerals*, **9**, 16 (2019) *Thermo-Elasticity of Materials from Quasi-Harmonic Calculations*

J. Maul, D. Ongari, S. M. Moosavi, B. Smit and A. Erba, *J. Phys. Chem. Lett.*, **11**, 8543–8548 (2020) *Thermoelasticity of Flexible Organic Crystals from Quasi-harmonic Lattice Dynamics: The Case of Copper(II) Acetylacetonate*

9.4.3 Computing the Whole Thermo-Elastic Tensor with CRYSTAL

As we said, several strain shapes have to be used in order to determine all elastic constants. Now we will address this point: given that we are interested in a specific elastic constant, how can we choose a proper strain shape?

Diagonal Thermo-Elastic Constants

The generic diagonal elastic constant C_{vv} can be directly determined by choosing the corresponding basic strain shape - the one having only the v-th component set to 1 and all the others set to 0. For instance, C_{22} can be easily obtained by choosing $\boldsymbol{\eta'} = [\,0\,1\,0\,0\,0\,0\,]$. In fact, since only η'_2 is non-zero, only the desired term survives in the summation of Equation (9.24):

$$C' = \sum_{vu} C_{vu}\,\eta'_v\,\eta'_u \equiv C_{22} \tag{9.25}$$

The same holds for C_{44} (by choosing $\boldsymbol{\eta'} = [\,0\,0\,0\,1\,0\,0\,]$, the resulting scalar elastic constant is indeed the desired one) and for all other diagonal elastic constants C_{vv}.

Off-Diagonal Thermo-Elastic Constants

The off-diagonal elastic constants C_{vu} require a bit more effort to be determined, because they can only be computed embedded into a linear combination containing also other constants. The easiest way to accomplish this is to use a composite strain shape consisting of the two involved basic ones - i.e. having the v-th and u-th components set to 1 with all the others set to 0.
For instance, in order to obtain the C_{23} elastic constant, start by choosing $\boldsymbol{\eta'} = [\,0\,1\,1\,0\,0\,0\,]$. According to Eq. (9.24), this will lead to a scalar elastic constant:

$$C' = \sum_{vu} C_{vu}\,\eta'_v\,\eta'_u = C_{22} + C_{23} + C_{32} + C_{33} = C_{22} + 2C_{23} + C_{33} \tag{9.26}$$

This scalar constant clearly contains C_{23}, but in a combination with C_{22} and C_{33}. These are diagonal terms, which are easy to be determined as just described above. Once these have been computed, it is possible to remove their contributions from the linear combination - leading to the desired result:

$$C_{23} = \frac{1}{2}\,(C' - C_{22} - C_{33}) \tag{9.27}$$

Now you get the idea. For instance, the same holds for C_{36}: by choosing $\boldsymbol{\eta'} = [\,0\,0\,1\,0\,0\,1\,]$, the resulting scalar elastic constant is $C' = C_{33} + 2C_{36} + C_{66}$ and once C_{33} and C_{66} are known, C_{36} can be determined by difference.

Symmetry-Preserving Strains

Up until now, we have not yet taken into account the internal symmetry of the system: the rules described so far are completely general and valid for all off-diagonal constants. However, different linear combinations of strain shapes can be identified that are more symmetry-preserving than others: in this way the computation will be much faster.
For instance, let's consider a cubic system (for which $C_{11} = C_{22} = C_{33}$ and $C_{12} = C_{13} = C_{23}$). We already know that the C_{12} may be determined with $\boldsymbol{\eta'} = [\,1\,1\,0\,0\,0\,0\,]$. However this deformation will strain the system along the x and y directions, reducing the symmetry to the tetragonal one. Instead, it is more convenient to choose $\boldsymbol{\eta'} = [\,1\,1\,1\,0\,0\,0\,]$ because it preserves the cubic symmetry. This will lead to a scalar elastic constant given by:

$$C' = \sum_{vu} C_{vu}\,\eta'_v\,\eta'_u = C_{11}+C_{12}+C_{13}+C_{21}+C_{22}+C_{23}+C_{31}+C_{32}+C_{33} = 3\,(C_{11} + 2C_{12}) \tag{9.28}$$

As usual, once C_{11} is known, C_{12} can be determined by difference from this combination. Note that this particular deformation is isotropic. As it turns out, we can completely avoid to perform the calculation having this strain shape. This is because the isotropic QHA calculation

has already been performed (this is a prerequisite of the thermo-elastic one). Therefore, it is possible to make use of the thermal dependence of the bulk modulus (given as a result of the **QHA** calculation), because it is a linear combination of the elastic constants:

$$K = \frac{1}{9}\left(C_{11} + C_{22} + C_{33} + 2C_{12} + 2C_{13} + 2C_{23}\right) \tag{9.29}$$

In other words, the isotropic QHA already provides one of the equations for the linear system involving the thermo-elastic constants.

9.4.4 Optional Sub-Keywords of the Thermo-Elastic Calculation

The thermo-elastic calculation **THERMOELAS** , being a sub-section of the **ELASTCON** keyword, makes use of the corresponding options. With **THERMOELAS** the elasticity is described in terms of Equation (9.20) where the free energy surface is explored along one direction at a time, through Equation (9.23). The derivative in the right-hand-side of Equation (9.23) is computed numerically by applying finite strains to the crystal lattice (see Figure 9.4.2). As we said earlier, these strains have a fixed shape $\boldsymbol{\eta'}$ defined by the user in the input. The set of strain amplitudes $\{\eta_i\}$ is determined according to the number of strained configurations N_s (keyword **NUMDERIV**, with a default value of 5) and the strain step size δ (keyword **STEPSIZE**, with a default value of 0.005). Hence, the actual strain matrices are determined and applied to the lattice, leading to deformed configurations.
Now, for each deformed configuration, atomic position are relaxed through an internal geometry optimization: this accounts for the nuclear contribution to the elasticity of the crystal. At this point the harmonic frequency calculation is carried out: the Hessian of the energy is determined by performing Cartesian atomic displacements of the system. The number of displacements per Cartesian atomic coordinate may be set through the keyword **HNUMDERIV**. Such parameter is 1 by default (meaning that the derivatives are performed by one-sided finite difference expressions) but may also be set to 2 (implying a two-sided finite difference formula). The amplitude of each atomic Cartesian displacement has a pre-defined value of 0.006 Å (for a standard equilibrium harmonic frequency calculation the default is 0.003 Å; however, since the system is expanding, we have found that a higher value is able to provide more stable results) and may be adjusted by using the keyword **HSTEPSIZE**.
From the frequencies of the system, the Helmholtz free energy is determined (through the vibrational partition function). At this point, the free energies of all the strained configurations are gathered. A polynomial regression is then performed over this set of data, leading to an estimation of the free energy as a function of the strain amplitude. By default only the quadratic fit is performed (this is the most stable one). The keyword **PRINT** may be used to request for all the available fits to be printed in output - from 3 regression parameters up to the number of strained configurations. From the polynomial free energy function, its derivative can be determined trivially. By normalizing this value by the volume of the system, the elastic constant is obtained and printed.
All these calculated elastic constants refer to isothermal conditions - indeed, these should be labelled C_{vu}^{T}. Experimental values are often measured under adiabatic conditions C_{vu}^{S}. In order to convert one set of constants to the other, the adiabatic correction has to be applied.

Chapter 10

Dielectric Properties up to Fourth Order via the Coupled Perturbed HF/KS Method

10.1 Response to an Electric Field - Theoretical Framework of the Coupled-Perturbed Treatment

The total energy E of a crystal in a constant static electric field $\boldsymbol{\varepsilon}$ can be expressed as

$$E(\boldsymbol{\varepsilon}) = E(0) - \sum_t \mu_t \varepsilon_t - \frac{1}{2!}\sum_{tu} \alpha_{tu}\varepsilon_t\varepsilon_u + \\ - \frac{1}{3!}\sum_{tuv} \beta_{tuv}\varepsilon_t\varepsilon_u\varepsilon_v - \frac{1}{4!}\sum_{tuvw} \gamma_{tuvw}\varepsilon_t\varepsilon_u\varepsilon_v\varepsilon_w + \dots \tag{10.1}$$

with $E(0)$ the field-free energy and $\boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \dots$ the total energy derivative tensors of order $1, 2, 3, 4 \dots$ with respect to the electric field (Cartesian components indicated by Roman subscripts t, u, v, w):

$$\mu_t = -\left.\frac{\partial E}{\partial \varepsilon_t}\right|_0 \tag{10.2}$$

$$\alpha_{tu} = -\left.\frac{\partial^2 E}{\partial \varepsilon_t \partial \varepsilon_u}\right|_0 \tag{10.3}$$

$$\beta_{tuv} = -\left.\frac{\partial^3 E}{\partial \varepsilon_t \partial \varepsilon_u \partial \varepsilon_v}\right|_0 \tag{10.4}$$

$$\gamma_{tuvw} = -\left.\frac{\partial^4 E}{\partial \varepsilon_t \partial \varepsilon_u \partial \varepsilon_v \partial \varepsilon_w}\right|_0 \tag{10.5}$$

As for the corresponding physical properties, $\boldsymbol{\mu}$ represents the dipole moment, $\boldsymbol{\alpha}$ the polarizability, $\boldsymbol{\beta}$ the first hyperpolarizability and $\boldsymbol{\gamma}$ the second hyperpolarizability.
By default, the perturbative series 10.1 is truncated at the second order and only the second energy derivatives α_{tu} are calculated (the dipole moment μ_t is ill-defined in the reciprocal space) using the expression

$$\alpha_{tu} = -\frac{4}{n_k}\sum_{\vec{k}}^{BZ} \Re\left\{ \sum_{\mu\nu}\sum_{a}^{\text{occ}}\sum_{p}^{\text{virt}} C^{\vec{k}*}_{a\mu}\Omega^{\vec{k},t}_{\mu\nu}C^{\vec{k}}_{\nu p}\mathsf{U}^{\vec{k},u}_{pa} \right\} \tag{10.6}$$

where n_k is the number of $\vec{k}$ points in the first Brillouin Zone (BZ) and the indices a $(b, c \dots)$ and p $(q, r \dots)$ run over the occupied and virtual crystalline orbitals, respectively.

$\mathsf{U}^{\vec{k},u}$ is the unknown anti-Hermitian off-diagonal-block matrix that linearly transforms the unperturbed eigenvectors $C^{\vec{k}}$ under the effect of the electric field perturbation represented by the matrix $\Omega^{\mathbf{k}}$:

$$\mathbb{C}^{\mathbf{k},u}_{\mu i} \equiv \left.\frac{\partial C^{\mathbf{k}}_{\mu i}}{\partial \varepsilon_u}\right|_0 = \sum_j^{\text{all}} C^{\mathbf{k}}_{\mu j} \mathsf{U}^{\mathbf{k},u}_{ji} \tag{10.7}$$

The off-diagonal blocks $\mathsf{U}^{\mathbf{k},u}_{ap}$ are defined as

$$\mathsf{U}^{\mathbf{k},u}_{ap} = \sum_{\mu\nu} \frac{C^{\mathbf{k}*}_{a\mu} \mathbb{F}^{\mathbf{k},u}_{\mu\nu} C^{\mathbf{k}}_{\nu p}}{E^{\mathbf{k}}_p - E^{\mathbf{k}}_a} \tag{10.8}$$

and are functions of both the energy gap $\left(E^{\mathbf{k}}_p - E^{\mathbf{k}}_a\right)$, and the perturbed Fock matrix,

$$\mathbb{F}^{\mathbf{k},u}_{\mu\nu} \equiv \left.\frac{\partial F^{\mathbf{k}}_{\mu\nu}}{\partial \varepsilon_u}\right|_0 = \Omega^{\mathbf{k},u}_{\mu\nu} + \sum_{\mathbf{g}} e^{\imath \mathbf{k}\cdot\mathbf{g}} \sum_{\lambda\rho} \sum_{\mathbf{k}'}^{BZ} \sum_{\mathbf{g}'} e^{-\imath \mathbf{k}'\cdot\mathbf{g}'} \sum_a \left(\mathbb{C}^{\mathbf{k}',u*}_{a\lambda} C^{\mathbf{k}'}_{\rho a} + C^{\mathbf{k}'*}_{a\lambda} \mathbb{C}^{\mathbf{k}',u}_{\rho a}\right) \times$$
$$\times \sum_{\mathbf{g}''} \left[\left(\mu^{\mathbf{0}} \nu^{\mathbf{g}} \left| \lambda^{\mathbf{g}'} \rho^{\mathbf{g}''} \right.\right) - \frac{1}{2}\left(\mu^{\mathbf{0}} \lambda^{\mathbf{g}'} \left| \nu^{\mathbf{g}} \rho^{\mathbf{g}''} \right.\right)\right] \tag{10.9}$$

Hence, according to Eq. (10.7), a Self-Consistent Coupled-Perturbed procedure (SC-CP) is to be carried out. For details about the CPHF/KS method and its implementation see Refs. [104, 105, 106, 237].

10.2 Coupled-Perturbed HF/KS Calculation of Dielectric Properties up to Second Order

The **CPHF** keyword activates the calculation of the the polarizability (and optionally the first and second order hyper-polarizabilities – see below) via the Coupled Perturbed HF/KS method. **CPHF** (or **CPKS**) must be the last keyword in the geometry input block:

```
. . . geometry input . . . .
CPHF
END       ! close CPHF input block
END       ! close geometry input block
```

The density functionals that are currently available for **CPHF** calculations are:

CORRELAT	Correlation Potential (default: no correlation).
LYP	GGA. Lee-Yang-Parr [167]
PBE	GGA. Perdew-Burke-Ernzerhof [204]
PBESOL	GGA. PBE functional revised for solids [206]
PWLSD	LSD. Perdew-Wang parameterization of the Ceperley-Alder free electron gas correlation results [207]
PWGGA	GGA. Perdew-Wang [205]
VWN	LSD. Vosko-Wilk-Nusair parameterization of the Ceperley-Alder free electron gas correlation results [282]
EXCHANGE	Exchange potential (default: Hartree-Fock exchange).
BECKE	GGA. Becke [18]
LDA	LSD. Dirac-Slater [68]
PBE	GGA. Perdew-Becke-Ernzerhof [204]
PBESOL	GGA. PBE functional revised for solids [206]
PWGGA	GGA. Perdew-Wang [205]
SOGGA	second order GGA. [302]
WCGGA	GGA - Wu-Cohen [293]

All combinations of the above mentioned exchange-correlation functionals are allowed as well as corresponding standalone keywords as: SVWN, BLYP, PBEXC, SOGGAXC, ...

HYBRID	Hybrid method
Global hybrids	
B3PW	Becke's 3 parameter functional [19] combined with the non-local correlation PWGGA [203, 208, 209, 207]
B3LYP	Becke's 3 parameter functional [19] combined with the non-local correlation LYP
PBE0	Adamo and Barone [5]
PBESOL0	Same than PBE0 but with PBEsol instead of PBE
B1WC	see [25]
WC1LYP	see [60]
B97H	see [60]
Range-separated hybrids	
RSHXLDA	see [1, 151]
wB97	see [153, 4]
wB97X	see [153, 4]
LC-BLYP	see [152, 271]
CAM-B3LYP	see [270]
SC-BLYP	see [152, 271]
HSE06	see [2, 204]
HSEsol	see [162, 206]
HISS	see [267, 268]
LC-wPBE	see [85]
LC-wPBEsol	see [85]
LC-wBLYP	see [85]
LC-PBE	see [152, 164]

Users of this module are kindly requested to cite the following papers:

M. Ferrero, M. Rérat, R. Orlando and R. Dovesi
Coupled perturbed Hartree-Fock for periodic systems: the role of symmetry and related computational aspects
J. Chem. Phys. 128, Art.N. 014100 (2008)

M. Ferrero, M. Rérat, R. Orlando and R. Dovesi
The calculation of static polarizabilities in 1-3D periodic compounds. The implementation in the CRYSTAL code
J. Comput. Chem. 29, 1450–1459 (2008)

M. Ferrero, M. Rérat, B. Kirtman and R. Dovesi
Calculation of first and second static hyper-polarizabilities of 1-3D periodic compounds. Implementation in the CRYSTAL code
J. Chem. Phys. 129, Art.N. 244110 (2008)

10.2.1 Tools for tuning convergence and accuracy in the Coupled-Perturbed iterations

Starting from CRYSTAL17, the DIIS accelerator is activated by default in the SC-CP iterations. This improves considerably stability with respect to previous implementations. However, difficult situations might arise, in which the user might want to act in order to achieve convergence. Convergence of the SC-CP cycle might be helped and/or tuned using the following optional keywords within the **CPHF** input block:

rec	variable	meaning
•	*A* **DIIS**	activates DIIS convergence accelerator for mixing Fock/KS matrix derivatives [default] (page 96)
•	*A* **NODIIS**	deactivates DIIS convergence accelerator (page 96)
•	*A* **DIISALLK**	DIIS is performed over all k-points of the Brillouin zone
•	*A* **HISTDIIS**	Limits DIIS mixing to the most recent $NCYC$ cycles
•	∗ NCYC	maximum number of previous cycles to be kept
•	*A* **PRTDIIS**	Activates detailed printing for the DIIS procedure
•	*A* **ANDERSON**	mixing of Fock/KS matrix derivatives (page 93)
•	*A* **BROYDEN**	mixing of Fock/KS matrix derivatives (page 95)
	∗ W0	*W0* parameter in Anderson's paper [154]
	∗ IMIX	percent of Fock/KS derivative matrices mixing
	∗ ISTART	SC-CP iteration after which Broyden method is active (minimum 2)
•	*A* **FMIXING**	mixing of Fock/KS matrix derivatives from iterations n and $n-1$
	∗ IPMIX	percentage of cycle $n-1$ [IPMIX=0, no mixing]
•	*A* **MAXCYCLE**	modify the maximum number of SC-CP iterations
	∗ NMAX	maximum number of iterations [default: 100]
•	*A* **SELEDIR**	set the maximum number of iterations along each cartesian direction
•	∗ NX, NY, NZ	maximum number of SC-CP iterations along x, y, z
•	*A* **TOLALPHA**	threshold on α variation between two SC-CP iterations
	∗ ITOL	$\|\Delta\alpha\| < 10^{-ITOL}$ [default: 3]

The keyword TOLALPHA tunes the desired accuracy to be achieved in the SC-CP iterative procedure.
The DIIS procedure requires the storage of error matrices for each cycle of the history. By default the information of all the previous iterations is kept, which can result in a significant occupation of disk space if the procedure goes on for several (50, 100) cycles. The maximum number of cycles to be used as an history can be set by the HISTDIIS keyword.

By default DIIS evaluates errors only in the gamma-point of the Brillouin zone. DIISALLK activates the DIIS procedure in all k-points with a resulting better accuracy – in principle at least – at the price of significantly increased memory/storage requirements. From the experience gathered so far, the default (gamma-point) DIIS proved to work well in most – if not all – cases.

10.2.2 Dynamic (Frequency-dependent) CPHF/KS

The CPHF/KS procedure described above refers to the response to a *static* field. If a *dynamic* polarizability response is desired, that is, dependent on the frequency of the oscillating field, the DYNAMIC keyword has to be inserted in the CPHF block:

rec	variable	meaning
•	*A* **DYNAMIC**	Solution of frequency-dependent SC-CP equations for a series of frequencies
	∗ NSTEPS	Number of frequency steps from FRQ to FRQ2
		if NSTEPS = 1 insert II
	∗ FRQ	Wavelength (in nm)
		else if NSTEPS > 1 insert II
	∗ FRQ	Starting wavelength (in nm)
	∗ FRQ2	Final wavelength (in nm)
•	*A* **DAMPING**	Sets a damping factor (peak broadening) related to the finite lifetime of excited states
	∗ DAMPFAC	Value of the damping factor in Hartree. Values between 0.001 and 0.003 are suggested (≈ 0.03 - 0.1 eV)

If the damping factor is not set, no imaginary part of the polarizability will be observed.
Due to technical reasons, the use of the DYNAMIC keyword in CPHF is restricted, in CRYSTAL17, to the use of non-hybrid functionals. The program will issue an error and stop if

Hartree-Fock or a hybrid functional (such as B3LYP or PBE0) is chosen for the calculation and the DYNAMIC keyword is active.

10.2.3 Static first hyperpolarizability

The **CPHF** calculation can also be extended up to the third perturbative order by including the optional keyword:

rec	variable	meaning
•	*A* **THIRD**	computes energy derivatives up to the third order

THIRD provides third order energy derivatives (see Eq. 10.4) calculated through the $(2n+1)$ scheme:

$$\beta_{tuv} = -\frac{2}{n_k}\sum_{\mathbf{k}}\Re\sum_{a}\sum_{p}\mathcal{P}_{t,u,v}\times \\ \times\left\{\mathsf{U}^{\mathbf{k},t*}_{ap}\left[\sum_{\mu\nu}\left(\sum_{q}C^{\mathbf{k}*}_{p\mu}\mathbb{F}^{\mathbf{k},u}_{\mu\nu}C^{\mathbf{k}}_{\nu q}\mathsf{U}^{\mathbf{k},v}_{qa}-\sum_{b}\mathsf{U}^{\mathbf{k},u}_{pb}C^{\mathbf{k}*}_{b\mu}\mathbb{F}^{\mathbf{k},v}_{\mu\nu}C^{\mathbf{k}}_{\nu a}\right)+\imath\frac{\partial\mathsf{U}^{\mathbf{k},v}_{pa}}{\partial k_u}\right]\right\} \quad (10.10)$$

The operator $\mathcal{P}$ permutes the indices of derivation. The derivative $\partial\mathsf{U}^{\mathbf{k},v}_{pa}/\partial k_u$ depends on the derivative of the coefficients $C^{\mathbf{k}}_{\mu i}$ with respect to $\mathbf{k}$, that is

$$\frac{\partial C^{\mathbf{k}}_{\mu i}}{\partial k_u} = \sum_{j}^{\text{all}} C^{\mathbf{k}}_{\mu j}\mathsf{Q}^{\mathbf{k},\mathbf{u}}_{ji} \quad (10.11)$$

similarly to Eq. (10.7). When $i \neq j$, the expression for matrix $\mathsf{Q}^{\mathbf{k},\mathbf{u}}$ elements is

$$\mathsf{Q}^{\mathbf{k},u}_{ij} = \frac{\sum_{\mu\nu}\left[C^{\mathbf{k}*}_{i\mu}\frac{\partial F^{\mathbf{k}}_{\mu\nu}}{\partial k_u}C^{\mathbf{k}}_{\nu j}-C^{\mathbf{k}*}_{i\mu}\frac{\partial S^{\mathbf{k}}_{\mu\nu}}{\partial k_u}C^{\mathbf{k}}_{\nu j}E^{\mathbf{k}}_{j}\right]}{E^{\mathbf{k}}_{j}-E^{\mathbf{k}}_{i}} \quad (10.12)$$

A threshold value (10^{-6}, by default) is defined below which eigenvalues are considered degenerate. Users can change this threshold using the keyword **TOLUDIK** as follows:

rec	variable	meaning
•	*A* **TOLUDIK**	minimum allowed difference between non-degenerate unperturbed eigenvalues
•	∗ ITOLU	$\lvert E^{\mathbf{k}}_{j}-E^{\mathbf{k}}_{i}\rvert = 10^{-ITOLU}$ [default: 6]

10.2.4 Dynamic first–hyperpolarizability tensors - Pockels and Second-Harmonic Generation

If the keyword **DYNAMIC** is present in input together with **THIRD**, a frequency-dependent field is considered for the third order tensor. The general expression for the first hyperpolarizability of closed-shell periodic systems in the presence of frequency-dependent fields may be written as:

$$\beta_{tuv(-\omega_\sigma;\omega_1,\omega_2)} = -\frac{2}{n_k}\mathcal{K}_{-\omega_\sigma,\omega_1,\omega_2}\sum_{\vec{k}}\Re\sum_{i}\sum_{a}P_{\frac{t}{-\omega_\sigma},\frac{u}{+\omega_1},\frac{v}{+\omega_2}} \\ \left\{U^{(t)*}_{ai(-\omega_\sigma)}(\vec{k})\left[\sum_{b}G^{(u)}_{ab(+\omega_1)}(\vec{k})U^{(v)}_{bi(+\omega_2)}(\vec{k})\right.\right. \\ \left.\left.-\sum_{j}U^{(v)}_{aj(+\omega_2)}(\mathbf{k})G^{(u)}_{ji(+\omega_1)}(\mathbf{k})+\imath\frac{\partial U^{(v)}_{ai(+\omega_2)}(\mathbf{k})}{\partial k_u}\right]\right\} \quad (10.13)$$

where i, j and a, b run over occupied and virtual crystalline orbitals (CO), respectively, and t, u, v indicate directions of the field. $\omega_\sigma = \omega_1 + \omega_2$ and $\mathcal{K}_{-\omega_\sigma;\omega_1,\omega_2}$ depend on the non linear optical process[199] ($\mathcal{K}_{0,0,0} = 1$, $\mathcal{K}_{-2\omega,\omega,\omega} = 1/2$, $\mathcal{K}_{-\omega,0,\omega} = 2$, $\mathcal{K}_{0,\omega,-\omega} = 1/2$, $\mathcal{K}_{\omega,-2\omega,\omega} = 1$, and $\mathcal{K}_{-\omega_\sigma,\omega_1,\omega_2} = 1$ if neither ω_σ, nor ω_1 nor ω_2 are null).
The two situations are considered in the code for each frequency ω:

- Second Harmonic Generation (SHG): $\omega_1 = \omega_2 = \omega$ and $\omega_\sigma = -2\omega$
- dc-Pockels (dc-P): $\omega_1 = \omega$, $\omega_2 = 0$ and $\omega_\sigma = -\omega$

Warning: In the current version of the code the calculation of dynamic hyperpolarizability is implemented only for pure (GGA, LDA) functionals, while it is not possible within hybrid functionals (e.g. PBE0, B3LYP, HSE) or Hartree-Fock.
A vibrational contribution to dc-Pockels and SHG tensors can be computed via the **BETAVIB** keyword – see the frequency calculation section of this manual.

10.3 Fourth-Order CPHF/KS – second hyperpolarizability calculation

Calculation of the second hyperpolarizability tensor $\boldsymbol{\gamma}$ is performed only if required:

rec	variable	meaning
•	*A* **FOURTH**	computes energy derivatives up to fourth order

The keyword **FOURTH** activates a second Self-Consistent Coupled-Perturbed procedure (SC-CP2) to provide matrix $\mathsf{U}^{\mathbf{k},uv}$ such that

$$\mathbb{C}^{\mathbf{k},uv}_{\mu i} \equiv \left.\frac{\partial^2 C^{\mathbf{k}}_{\mu i}}{\partial\varepsilon_u \partial\varepsilon_v}\right|_0 = \sum_j^{all} C^{\mathbf{k}}_{\mu j} \mathsf{U}^{\mathbf{k},uv}_{ji} \tag{10.14}$$

By means of $\mathsf{U}^{\mathbf{k},uv}$, both derivatives (10.4) and (10.5) are defined:

$$\beta_{tuv} = -\frac{1}{n_k}\sum_{\mathbf{k}} \Re\left\{ \mathcal{P}_{t,u,v} \sum_a \sum_p \left(\sum_{\mu\nu} C^{\mathbf{k}*}_{a\mu} \mathbb{F}^{\mathbf{k},t}_{\mu\nu} C^{\mathbf{k}}_{\nu p} \mathsf{U}^{\mathbf{k},uv}_{pa} - \sum_{\mu\nu} C^{\mathbf{k}*}_{a\mu} \mathbb{F}^{\mathbf{k},uv}_{\mu\nu} C^{\mathbf{k}}_{\nu p} \mathsf{U}^{\mathbf{k},t}_{pa} \right) \right\} \tag{10.15}$$

in the $n+1$ formulation (equivalent to Eq. 10.10), and

$$\begin{aligned}\gamma_{tuvw} = -\frac{1}{n_k}\sum_{\mathbf{k}} \Re\Bigg\{ & \mathcal{P}_{t,u,v,w} \sum_a \sum_p \sum_{\mu\nu} \mathsf{U}^{\mathbf{k},t*}_{ap} \times \\ & \times \Bigg[\sum_q C^{\mathbf{k}*}_{p\mu} \left(\mathbb{F}^{\mathbf{k},u}_{\mu\nu} C^{\mathbf{k}}_{\nu q} \mathsf{U}^{\mathbf{k},vw}_{qa} + \frac{1}{2} \mathbb{F}^{\mathbf{k},vw}_{\mu\nu} C^{\mathbf{k}}_{\nu q} \mathsf{U}^{\mathbf{k},u}_{qa} \right) + \\ & - \sum_b \left(\mathsf{U}^{\mathbf{k},vw}_{pb} C^{\mathbf{k}*}_{b\mu} \mathbb{F}^{\mathbf{k},u}_{\mu\nu} + \frac{1}{2} \mathsf{U}^{\mathbf{k},u}_{pb} C^{\mathbf{k}*}_{b\mu} \mathbb{F}^{\mathbf{k},vw}_{\mu\nu} \right) C^{\mathbf{k}}_{\nu a} + \\ & - \sum_b \mathsf{U}^{\mathbf{k},u}_{pb} C^{\mathbf{k}*}_{b\mu} \mathbb{F}^{\mathbf{k},v}_{\mu\nu} \sum_q C^{\mathbf{k}}_{\nu q} \mathsf{U}^{\mathbf{k},w}_{qa} \Bigg] + \imath \mathsf{U}^{\mathbf{k},tw*}_{ap} \frac{\partial \mathsf{U}^{\mathbf{k},u}_{pa}}{\partial k_v} \Bigg\}\end{aligned} \tag{10.16}$$

The off-diagonal blocks of matrix $\mathsf{U}^{\mathbf{k},uv}$,

$$\begin{aligned}\mathsf{U}^{\mathbf{k},uv}_{ap} = \frac{1}{E^{\mathbf{k}}_p - E^{\mathbf{k}}_a} \Bigg\{ & \mathcal{P}_{u,v} \Bigg[\sum_b \sum_{\mu\nu} C^{\mathbf{k}*}_{a\mu} \mathbb{F}^{\mathbf{k},u}_{\mu\nu} C^{\mathbf{k}}_{\nu b} \mathsf{U}^{\mathbf{k},v}_{bp} + \\ & - \sum_q \sum_{\mu\nu} \mathsf{U}^{\mathbf{k},u}_{aq} C^{\mathbf{k}*}_{q\mu} \mathbb{F}^{\mathbf{k},v}_{\mu\nu} C^{\mathbf{k}}_{\nu p} + \imath \frac{\partial \mathsf{U}^{\mathbf{k},v}_{ap}}{\partial k_u} \Bigg] + \sum_{\mu\nu} C^{\mathbf{k}*}_{a\mu} \mathbb{F}^{\mathbf{k},uv}_{\mu\nu} C^{\mathbf{k}}_{\nu p} \Bigg\}\end{aligned} \tag{10.17}$$

depend on themselves through the second derivative of the Fock matrix

$$\mathbb{F}_{\mu\nu}^{\mathbf{k},uv} \equiv \left.\frac{\partial^2 F_{\mu\nu}^{\mathbf{k}}}{\partial\varepsilon_u\partial\varepsilon_v}\right|_0 = \sum_{\mathbf{g}} e^{\imath\mathbf{k}\cdot\mathbf{g}} \sum_{\lambda\rho} \sum_{\mathbf{k}'}^{BZ} \sum_{\mathbf{g}'} e^{-\imath\mathbf{k}'\cdot\mathbf{g}'} \sum_{a} \Big(\mathbb{C}_{a\lambda}^{\mathbf{k}',uv*} C_{\rho a}^{\mathbf{k}'} + \\ + \mathcal{P}_{u,v} \mathbb{C}_{a\mu}^{\mathbf{k}',u*} \mathbb{C}_{\nu a}^{\mathbf{k}',v} + C_{a\lambda}^{\mathbf{k}'*} \mathbb{C}_{\rho a}^{\mathbf{k}',uv} \Big) \left[\left(\mu^{\mathbf{0}} \nu^{\mathbf{g}} \left| \lambda^{\mathbf{g}'} \rho^{\mathbf{g}''} \right. \right) - \frac{1}{2} \left(\mu^{\mathbf{0}} \lambda^{\mathbf{g}'} \left| \nu^{\mathbf{g}} \rho^{\mathbf{g}''} \right. \right) \right] \quad (10.18)$$

and therefore, must be determined iteratively.

10.3.1 Tuning and control of the CPHF2 iterative procedure

Convergence of the SC-CP2 cycle is controlled with the following optional keywords:

rec	variable	meaning
•	*A* **DIIS2**	activates DIIS convergence accelerator for mixing Fock/KS matrix 2nd derivatives *default* (page 96)
•	*A* **NODIIS**	deactivates DIIS convergence accelerator (page 96)
•	*A* **DIIS2ALLK**	DIIS is performed over all k-points of the Brillouin zone
•	*A* **HISTDIIS2**	Limits DIIS mixing to the most recent $NCYC$ cycles
•	∗ NCYC	maximum number of previous cycles to be kept
•	*A* **PRTDIIS2**	Activates detailed printing for the DIIS2 procedure
•	*A* **ANDERSON2**	mixing of Fock/KS matrix 2nd derivatives (page 93)
•	*A* **BROYDEN2**	mixing of Fock/KS matrix 2nd derivatives (page 95)
	∗ W02	*W0* parameter in Anderson's paper [154]
	∗ IMIX2	percentage of Fock/KS second derivative matrices mixing
	∗ ISTART2	SC-CP2 iteration after which Broyden method is activated (minimum 2)
•	*A* **FMIXING2**	mixing of Fock/KS matrix second derivatives from SC-CP2 iterations n and $n-1$
	∗ IPMIX2	percentage of cycle $n-1$ [IPMIX2=0, no mixing]
•	*A* **MAXCYCLE2**	modify the maximum number of SC-CP2 iterations
	∗ NMAX2	maximum number of iterations [default: 100]
•	*A* **SELEDIR2**	set the maximum number of SC-CP2 iterations along each couple of cartesian indices
•	∗ NXX, NXY, NXZ, NYY, NYZ, NZZ	maximum number of SC-CP2 iterations along mixed directions xx, xy, xz, yy, yz, zz
•	*A* **TOLGAMMA**	threshold on $\mathsf{U}^{\mathbf{k},tu}$ variation between two SC-CP2 iterations
	∗ ITOL2	$\lvert\Delta\mathsf{U}^{\mathbf{k},tu}\rvert = 10^{-ITOL2}$ [default: 3]

Also in the case of CPHF2, DIIS is the default since CRYSTAL17. The meaning of optional DIIS keywords is the same as discussed for the first-oder CPHF case.
Dynamic (frequency-dependent) second hyperpolarizability is not yet implemented.

10.4 Restart of first– or second– order CPHF

RESTART
A static CPHF/KS run can be restarted from a previous run (even an incomplete run) by using the **RESTART** keyword. This works for both the first– and second–order perturbed iterative procedures. Every CPHF/KS run writes the necessary information for a restart to file fort.31. This file must be provided as file fort.32 before running the new calculation with the **RESTART** keyword. **GUESSP** (SCF guess from density matrix of a previous run, input block 3, page 109) is not applied by default, but its use is recommended.

In the case of a **DYNAMIC** calculation, restart will not be possible.

Chapter 11

Tools for Studying Solid Solutions

The theoretical modeling of disordered systems and solid solutions relies on obtaining average properties over a number of *configurations*, namely distributions of different species (atoms or vacancies) at a given set of atomic positions. Symmetry plays a key role in this context as shown in the following reference papers:

S. Mustapha, Ph. D'Arco, M. De La Pierre, Y. Noel, M. Ferrabone and R. Dovesi
On the use of symmetry in configurational analysis for the simulation of disordered solids
J. Phys.: Condens. Matter 25, 105401 (2013)

Ph. D'Arco, S. Mustapha, M. Ferrabone, Y. Noel, M. De La Pierre and R. Dovesi
Symmetry and random sampling of symmetry independent configurations for the simulation of disordered solids
J. Phys.: Condens. Matter 25, 355401 (2013)

See also **http://www.crystal.unito.it ⇒ tutorials ⇒ Disordered systems and solid solutions**

Consider, for example, a structure (any dimension) of symmetry group G, characterized by one irreducible crystallographic position d of multiplicity $|D|$. Such a $|D|d$ position (in Wyckoff's notation) is occupied by the atomic species $\mathcal{A}$. Suppose that a different atomic species $\mathcal{X}$ can replace $\mathcal{A}$ in any proportion on d. Then, $|D|+1$ compositions are possible:

$$\mathcal{A}_{|D|-\alpha}\mathcal{X}_{\alpha} \quad , \qquad \alpha = 0 \ldots |D| \tag{11.1}$$

For each composition, there exist

$$|S_{\alpha}| = \left(\begin{array}{c} |D| \\ \alpha \end{array} \right) = \frac{|D|!}{\alpha!(|D|-\alpha)!} \tag{11.2}$$

different possibilities to place atoms $\mathcal{A}$ and $\mathcal{X}$, that are different *configurations*. Overall, we expect a total number of $|S| = |2|^{|D|}$ configurations for $|D|+1$ compositions. Figure 11.1 shows the set of configurations for two atomic species ($\mathcal{A}$ and $\mathcal{X}$) distributed over four positions.
As the group of symmetry G acts on the whole set of configurations (S), the latter is partitioned in $|\Delta(S)|$ classes of equivalence, each one being a symmetry-independent class (SIC). Two configurations belong to the same SIC if there exists at least one element of G that transforms one configuration into the other. Figure 11.2 shows the partitioning of the configurations under the action of C_{4v} group.
All the configurations of a given class are degenerate and share the same properties (composition, symmetry group...). Therefore, in order to fully characterize the system, it is sufficient to determine:

- the number of SIC,

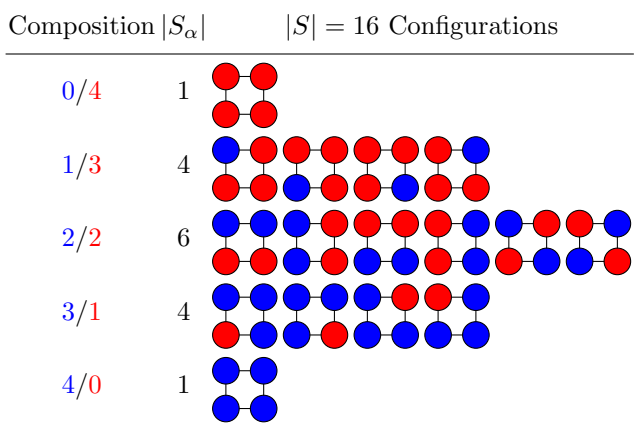


Figure 11.1: Possible configurations for 2 atomic species sitting on 4 positions. The two species are represented by red and blue circles. Configurations are ordered per composition. The number of configurations ($|S_\alpha|$) per composition is indicated.

- the number of configurations per SIC;
- one representative configuration for each SIC.

Such analysis is performed by the CRYSTAL alternative options **CONFCNT** (272) and **CONFRAND** (274). In order to study supercell configurations, the keyword **SCELCONF** (68) is to be coupled with the above mentioned options.

11.1 Counting and Enumerating Configurations

The keyword **CONFCNT** must be inserted in the geometry input block. It allows to calculate the number of classes as a function of the composition and provides a representative for each class.
In the present implementation:

- the number of different atomic species is restricted to two ($\mathcal{A}$ and $\mathcal{X}$);
- substitutions might take place on one or more crystallographic irreducible sites; for each irreducible site, all the symmetry equivalent positions are involved.

The minimal **CONFCNT** input is as follows:

rec	variable	meaning
• *A*	**CONFCNT**	
• ∗	NIS	number of irreducible sites
• ∗	IAT(I),I=1,NIS	atomic label of each irreducible site
• *A*	**END**	end of the **CONFCNT** sub-block

This yields the number of SIC over the full range of (NIS+1) compositions corresponding to NIS irreducible crystallographic positions. For each SIC, a representative configuration is given, along with its multiplicity and the number of symmetry operators of its group (being a subgroup of the group of the original cell). The representative configurations are printed in lexicographic order. By default, the *replacing* species is labeled as 'XX'.
The following optional keywords may be adopted for tuning **CONFCNT** calculations:

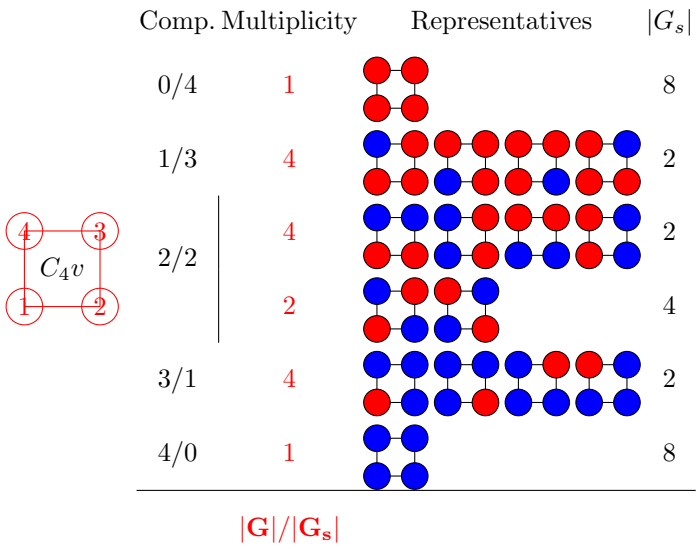


Figure 11.2: The six symmetry independent classes of configurations resulting of the action of the C_{4v}
group on the set of 2-color configurations on 4 positions. The number of configurations per class or multiplicity is indicated. The number of symmetry operators in the group of each configuration is given ($|G_s|$). The multiplicty of each class equals $|G|/|G_s|$.

rec	variable	meaning
•	*A* ATOMB	identifies the species $\mathcal{X}$
•	∗ NATB	atomic number of $\mathcal{X}$ (default 'XX')
•	*A* ONLYCOMP	selects certain compositions
•	∗ NC	number of compositions to be considered
•	∗ TC(I),I=1,NC	number of atoms $\mathcal{X}$ in composition I
•	*A* CONFPRT	prints information about the configurations
•	∗ IP1	integer $0 \div 2$
		0 - number of SIC only
		1 - listing in compact form
		2 - extended output [default]

Further options exist for counting and enumerating two-body interactions.
Two-body interactions are presented as:

label At.1 – label At.2 [I J L]

At.1 is supposed to be in the reference cell [0 0 0] but At.2 can be in another cell whose position is given by [I J L]. Interactions are presented adding the so-called empty and one-body terms, that are related to the cluster expansion formalism. These two terms are configuration independent; they depend only on chemistry.

rec	variable	meaning
•	*A* INTPRT	prints information about the interactions
•	∗ IP2	integer $0 \div 2$
		0 - no information printed [default]
		1 - condensed output (symmetry-independent interactions)
		2 - extended output (full set of interactions)
•	*A* INTMAXDIST	defines the distance cut-off for 2-body interactions
•	∗ RLIM	distance in Å[default 6Å]

11.2 Uniform Random Sampling of Symmetry Independent Configurations

When the number of classes is too large, a symmetry adapted Monte Carlo uniform at random sampling of the symmetry independent classes may be performed. In this case, the keyword **CONFRAND** must be inserted in the geometry input block and terminated by **END**. **CONFRAND** switches on a symmetry-adapted sampling of the SIC at a given composition IX (equivalent to α in equations (11.1) and (11.2)):

rec	variable	meaning
•	*A* **CONFRAND**	
•	∗ NIS	number of irreducible sites
•	∗ IAT(I),I=1,NIS	atomic label of each irreducible site
•	∗ IX	number of positions occupied by the 2nd species
•	*A* **END**	end of the **CONFRAND** sub-block

A file named CONFIGURATIONS.DAT is generated (see **RUNCONFS** at page 275) containing the following information:

- structure and symmetry of the *aristotype* supercell (written the same way as in a regular unit fort.34);
- number of equivalent crystallographic sites involved for substitutions ($|D|$) and number of substitutions (IX);
- labels of the atoms involved for substitutions;
- number of SIC found;
- list of the configurations. Each configuration is identified by its multiplicity and its rank.

For example:

```
. . . fort.34 . . . .
     8           4                    ! 8 sites, 4 substitutions
1    2    3    4    5    6    7    8 ! labels of the involved sites
     4                                 ! 4 SIC found:
   1        8        1                ! label, multip, rank
   2        8        2                !   "      "      "
   3       48        6                !   "      "      "
   4        6       60                !   "      "      "
```

For each configuration found by sampling the SIC space, the corresponding *canonical* configuration (that is the equivalent configuration of minimum rank) is detected as well. One may consider only canonical configurations (and thus save the canonical rank in file CONFIGURATIONS.DAT) by entering the keyword **CANONIC** within the **CONFRAND** block.
By default, the SIC space is sampled entirely and uniformly at random. The user may limit either the number of tries or the number of SIC to be searched by exploiting the **UNIFORM** option (see below). Further optional keywords, namely **SYMONLY** and **SYASYM**, manage a tuned probability distribution according to whether the SIC are symmetric (that is having symmetry operators other than the identity) or not. **SYMONLY**, in particular, sets to 0 the probability of asymmetric SIC. Then, by using a large number of tries NTC, all the simmetric SIC might be found at no significant computational cost. Few runs are enough to establish the exact number of symmetric SIC, the less the larger the total number of SIC.

rec	variable	meaning
•	*A* UNIFORM	uniform at random selection
•	∗ NTC	integer number NTC > 0 number of tries to be performed NTC < 0 number of SIC to be searched; (maximum number of tries set to 10×NTC)

rec	variable	meaning
•	*A* SYMONLY	search for symmetric SIC only
•	∗ NTC	NTC > 0 number of tries to be performed NTC < 0 number of SIC to be found
•	*A* SYASYM	search for symmetric SIC first; then for any SIC
•	∗ NTCS,NTCA	> 0, NTCS tries on symmetric SIC and NTCA tries on any other SIC < 0, search NTCS symmetric SIC and NTCA other the SIC

N.B. In case you are interested in studying supercell configurations, the keyword **SCELCONF** (68) must be used, and it must be inserted before the **CONFRAND** sub-block.

11.3 Calculations on Predefined Configurations

The keyword **RUNCONFS** must be inserted in the geometry input block. It opens the following minimal input sub-block:

rec	variable	meaning
•	*A* RUNCONFS	
•	*A* ATOMSUBS	identify the species involved for substitutions
•	∗ IZA IZB	atomic numbers
•	*A* END	

By default, once specified the two elements involved for substitutions, **RUNCONFS** builds a set of configurations from the list of ranks written in file CONFIGURATIONS.DAT (274) and performs single point calculations. Geometry optimizations are also allowed, and may be activated with the keyword **OPTGEOM** (for the relative options see page 182).
The basis set complete with the functions for atom IZB must be given in input after the SCF block: the keyword **BASE** opens a new input block where a standard basis set must be inserted (Section 3.2).
Independent calculations on different configurations might be carried on simultaneously exploiting the option

rec	variable	meaning
•	*A* **MULTITASK**	perform multiple tasks simultaneously
•	∗ N	number of tasks

Further optional keywords are:

rec	variable	meaning
•	*A* **SYMORDER**	sort SIC from file CONFIGURATIONS.DAT in order of increasing multiplicity
•	*A* **INICO**	select the first configuration to be considered
•	∗ IB	1st configuration label
•	*A* **IFICO**	select the last configuration to be considered
•	∗ IE	last configuration label

The order of execution follows that of the list of configurations written in file CONFIGURATIONS.DAT unless the option **SYMORDER** is introduced which rearranges the SIC in order of increasing multiplicity (that is with a decreasing number of symmetry operators).
The options **INICO** and **IFICO** allow to perform "multi-step" calculations by consecutively selecting different subsets of configurations from file CONFIGURATIONS.DAT or from the list rearranged by **SYMORDER**. This possibility might be useful not only to manage wall-time

queues, but also to restart **RUNCONFS** from the configuration closest to the point where it eventually failed.

11.4 Exploring the Neighborhood of a Configuration

The keyword **CONFNEIG** allows to explore the potential energy surface around one particular SIC. The minimal input sub-block is the following:

rec	variable	meaning
•	*A* CONFNEIG	search for neighboring configurations
•	∗ IRANK	rank of the SIC of interest
•	*A* END	

The only requirement is reading a unit CONFIGURATIONS.DAT with the same aristotype structure. Therefore, it necessary to perform a **CONFRAND** preliminarly (see 274). Then, **CONFNEIG** writes a file named CONFNEIGHBORS.DAT, which contains a list of canonical configurations generated from the specified SIC (i.e., IRANK), by exchange of a couple of atoms of different colors. The unit CONFNEIGHBORS.DAT has the same file format as CONFIGURATIONS.DAT. So, it can be read by the option **RUNCONFS** for running calculations on the configurations listed therein (275).
Suppose you have already run calculations on a certain number of SICs among those found by **CONFRAND** in the first place. You may want to exclude these SICs from the subsequent listing performed by **CONFNEIG**, in order to avoid making duplicate calculations. The optional keyword **RANKLIST** (to be inserted in the **CONFNEIG** sub-block) serves the purpose, when the list of SICs to be avoided is provided in a file named RANK.LST. The structure of RANK.LST must be the following:

rec	variable	meaning
•	∗ NRNK	number of SICs listed
		add NRNK records
•	∗ IRNK	rank of each SIC

Chapter 12

Equations of State

The program can perform an automated scan over the volume in order to compute energy E *vs.* volume V curves that are then fitted to various Equations of State (EOS) as Murnaghan's, third-order isothermal Birch-Murnaghan's, "universal" logarithmic Poirier-Tarantola's and exponential Vinet's, in order to compute equilibrium properties such as bulk modulus B_0 and its first derivative with respect to the pressure B_0' and volume/pressure dependence of the energy, enthalpy and bulk modulus. For each volume, a full V-constrained geometry optimization is performed. This option is activated by inserting the keyword **EOS** at the end of the Geometry input block, which opens a sub-block that must be closed with a keyword **END**:

```
. . . geometry input . . . .
EOS
[Optional keywords]
END   close EOS input block
END   close Geometry input block
```

Users of this module are kindly asked to cite the following reference:

A. Erba, A. Mahmoud, D. Belmonte and R. Dovesi, *J. Chem. Phys.*, **140**, 124703 (2014)

A volume range and a number of volumes N_V can be defined in input (default values are used otherwise) with the **RANGE** sub-keyword. The initial geometry is assumed to be fully optimized; if not, the **PREOPTGEOM** sub-keyword must be used to perform a preliminary optimization of both lattice parameters and atomic positions. For each considered volume V_i, a V_i-constrained optimization is performed (in fractional coordinates only) and the corresponding minimum energy E_i determined. The set of N_V data points $\{V_i, E_i\}$ is fitted to the various EOSs implemented. The fitted energy, enthalpy and bulk modulus are printed at the end of the calculation, for each EOS, as a function of volume and pressure at various points. These volume/pressure ranges can be defined with sub-keywords **VRANGE** and **PRANGE**, respectively. Typically, this kind of final information analysis can be performed with a complete restart of the calculation, using the **RESTART2** sub-keyword. A partial restart from a previous incomplete run can be activated with the sub-keyword **RESTART**. These two restart options require the external restart file EOSINFO.DAT.

12.1 A few theoretical remarks

The equilibrium bulk modulus B_0 of a crystal can be defined as follows:

$$B_0 = -V \left(\frac{\partial P}{\partial V} \right)_T . \tag{12.1}$$

A dimensionless parameter B_0' can then be defined as its first derivative with respect to the pressure, at constant temperature T:

$$B_0' = \left(\frac{\partial B_0}{\partial P}\right)_T \ . \tag{12.2}$$

Let us recall that the pressure P may be written as a function of the volume V as:

$$P(V) = -\left(\frac{\partial E}{\partial V}\right)_S \ . \tag{12.3}$$

According to equation (12.3) we can redefine the bulk modulus in equation (12.1) as the second energy derivative with respect to the volume:

$$B(V) = V\left(\frac{\partial^2 E}{\partial V^2}\right)_{T,S} \ . \tag{12.4}$$

We can now define the enthalpy H (coinciding with Gibbs' free energy G at $T = 0$ K) as a function of the volume V simply as:

$$H(V) = E(V) + P(V) \times V \ . \tag{12.5}$$

Several $E(V)$ equations of state have been proposed. We have implemented four among them.

1. In 1944, Murnaghan proposed his famous equation of state:

$$E(V) = E_0 + \frac{B_0 V}{B_0'}\left[\left(\frac{V_0}{V}\right)^{B_0'}\frac{1}{B_0'-1} + 1\right] - \frac{B_0 V_0}{B_0'-1} \ , \tag{12.6}$$

   where V_0 and E_0 are the equilibrium volume and energy, at zero pressure. Application of equation (12.3) to equation (12.6), gives $P(V)$ Murnaghan's EOS:

$$P(V) = \frac{B_0}{B_0'}\left[\left(\frac{V_0}{V}\right)^{B_0'} - 1\right] \ . \tag{12.7}$$

2. The third-order Birch-Murnaghan isothermal equation of state, published in 1947, reads like:

$$E(V) = E_0 + \frac{9V_0B_0}{16}\left\{\left[\left(\frac{V_0}{V}\right)^{\frac{2}{3}} - 1\right]^3 B_0' + \left[\left(\frac{V_0}{V}\right)^{\frac{2}{3}} - 1\right]^2\left[6 - 4\left(\frac{V_0}{V}\right)^{\frac{2}{3}}\right]\right\} \ . \tag{12.8}$$

   Again, according to equation (12.3), we can get $P(V)$ third-order Birch-Murnaghan's EOS:

$$P(V) = \frac{3B_0}{2}\left[\left(\frac{V_0}{V}\right)^{\frac{7}{3}} - \left(\frac{V_0}{V}\right)^{\frac{5}{3}}\right]\left\{1 + \frac{3}{4}(B_0' - 4)\left[\left(\frac{V_0}{V}\right)^{\frac{2}{3}} - 1\right]\right\} \ . \tag{12.9}$$

3. The third-order "universal" Poirier-Tarantola logarithmic equation of state, proposed in 1998, is:

$$E(V) = E_0 + \frac{B_0V_0}{2}\left[\ln\left(\frac{V_0}{V}\right)\right]^2 + \frac{B_0V_0}{6}\left[\ln\left(\frac{V_0}{V}\right)\right]^3 (B_0' - 2) \ , \tag{12.10}$$

   while $P(V)$ Poirier-Tarantola's EOS is:

$$P(V) = B_0\frac{V_0}{V}\left[\ln\left(\frac{V_0}{V}\right) + \frac{(B_0' - 2)}{2}\left[\ln\left(\frac{V_0}{V}\right)\right]^2\right] \ . \tag{12.11}$$

4. The exponential Vinet's equation of state, published in 1987, reads:

$$E(V) = E_0 + \frac{2B_0V_0}{(B_0'-1)^2}\left\{2 - \left[5 + 3\left(\frac{V}{V_0}\right)^{\frac{1}{3}}(B_0'-1) - 3B_0'\right] \times \right.$$
$$\left. \times \exp\left[-\frac{3}{2}(B_0'-1)\left[\left(\frac{V}{V_0}\right)^{\frac{1}{3}} - 1\right]\right]\right\} \quad (12.12)$$

According to equation (12.3), we get $P(V)$ Vinet's EOS:

$$P(V) = 3B_0\left(\frac{V}{V_0}\right)^{-\frac{2}{3}}\left[1 - \left(\frac{V}{V_0}\right)^{\frac{1}{3}}\right]\exp\left[-\frac{3}{2}(B_0'-1)\left[\left(\frac{V}{V_0}\right)^{\frac{1}{3}} - 1\right]\right] . \quad (12.13)$$

12.2 Keywords, options and defaults

Users of this module are kindly asked to cite the following reference:

A. Erba, A. Mahmoud, D. Belmonte and R. Dovesi, *J. Chem. Phys.*, **140**, 124703 (2014)

A default value is chosen for all computational parameters. The SCF energy convergence threshold is set to 10^{-8}. Optional keywords are (in any order):

rec	variable	meaning
•	*A* **RANGE**	keyword to specify the range of volumes and number of points in the $E(V)$ curve where optimizations have to be performed.
	∗ VOL1	minimum (compression) variation of the initial volume [$default$ = 0.92]
	∗ VOL2	maximum (expansion) variation of the initial volume [$default = 1.08$]
	∗ NPOINTS	number N_V of points in the selected range [$default = 6$]

The interval is specified as the minimum (compression) and maximum (expansion) variation of the volume of the initial geometry. The set of volumes is then defined according to the number of points in the selected range. For instance, default settings correspond to:

```
EOS
RANGE
0.92 1.08 6
END
```

i.e., to 6 equally-spaced volume points between $0.92\times V_{eq}$ and $1.08\times V_{eq}$, where V_{eq} is the volume of the equilibrium geometry given as input (assumed to be the fully optimized structure) or as obtained after a preliminary geometry optimization within the EOS option. Note that the equilibrium values V_{eq} (or V_0) and E_0 are always included in the final data for fitting (*i.e.* 7 points are used for fitting in the example above).

rec	variable	meaning
•	*A* **VRANGE**	Defines the volume range and the number of points where fitted values of pressure, energy, enthalpy and bulk modulus are printed at the end of the calculation. This keyword is for output purposes only.
	∗ VMIN	minimum volume
	∗ VMAX	maximum volume
	∗ NVOL	number of points in the selected range
•	*A* **PRANGE**	Defines the pressure range and the number of points where fitted values of volume, energy, enthalpy and bulk modulus are printed at the end of the calculation. This keyword is for output purposes only.
	∗ PMIN	minimum pressure
	∗ PMAX	maximum pressure
	∗ NPRE	number of points in the selected range
•	*A* **PRINT**	This option turns on a higher level of diagnostic printing and will generally not be necessary for the typical user. Fitting of the E(V) points is also performed for increasing set of points from 4 to the maximum number of points. (*default* is minimal printing)

rec	variable	meaning
•	*A* **RESTART**	Allows for partial restart from a previous incomplete run, using file EOSINFO.DAT.
•	*A* **RESTART2**	Allows for complete restart from a previous complete run, using file EOSINFO.DAT. This option can be used when one wants to explore different ranges of volume/pressure at the end of the calculation, using keywords **VRANGE** and **PRANGE** (see example below).
•	*A* **PREOPTGEOM**	A preliminary geometry optimization of cell and atomic positions is performed before starting the $E(V)$ curve calculation.

Geometry optimization is performed at convergence criteria tighter than the ones given in **OPTGEOM**. Some values can be modified by inserting the following keywords:

•	*A* **TOLDEG**	EOS default [0.0003] - see **OPTGEOM**, page 188
•	*A* **TOLDEX**	EOS default [0.00062] - see **OPTGEOM**, page 188
•	*A* **TOLDEE**	EOS default [8] - see **OPTGEOM**, page 188
•	*A* **MAXCYCLE**	see **OPTGEOM**, page 190
•	*A* **NOTRUSTR**	see **OPTGEOM**, page 188
•	*A* **TRUSTRADIUS**	see **OPTGEOM**, page 188
•	*A* **MAXTRADIUS**	see **OPTGEOM**, page 188

12.3 Output Information

Let us consider the case of α-quartz, with the following input:

```
EOS
RANGE
0.90 1.05 10
PRANGE
-5 10 20
END
```

At the end of the calculation, the 11 volume/energy data points obtained after the geometry optimizations are sorted and printed as follows:

```
SORTING VOLUMES/ENERGIES

VOLUME (A^3)    ENERGY (a.u.)

105.093052   -1.319745599801E+03
106.955671   -1.319747774744E+03
108.858269   -1.319749524843E+03
110.762118   -1.319750838191E+03
112.687177   -1.319751759039E+03
114.634069   -1.319752307968E+03
116.602831   -1.319752502563E+03
116.803375   -1.319752503067E+03
118.593687   -1.319752358314E+03
120.606573   -1.319751889709E+03
122.641225   -1.319751107890E+03
```

The following table is then reported with the fitted values of the minimum volume, energy, bulk modulus B_0 and its first derivative B_0':

```
+++++++ FITTING USING ALL POINTS +++++++
```

EQUATION OF STATE	VOL(A^3)	E(AU)	BM(GPA)	BM PRIME
MURNAGHAN 1944	116.7247	-1319.75250331	43.81	3.78
BIRCH-MURNAGHAN 1947	116.7202	-1319.75250331	43.83	3.88
POIRIER-TARANTOLA 1998	116.7172	-1319.75250341	43.85	3.95
VINET 1987	116.7186	-1319.75250334	43.84	3.92

Additionally, for each EOS, the following fitted data are reported (for instance Vinet's ones) in the pressure range (from -5 GPa to 10 GPa in this case) defined by input with the **PRANGE** sub-keyword (the **VRANGE** sub-keyword could be used instead for defining an explored volume range):

++
THERMODYNAMIC FUNCTIONS OBTAINED WITH EOS: VINET 1987
++

V = VOLUME, P = PRESSURE, E = ENERGY, G = GIBBS FREE ENERGY, B = BULK MODULUS

V (A^3)	P (GPa)	E (a.u.)	G (a.u.)	B (GPa)
136.37	-5.00	-1319.73967343	-1319.89603452	21.96
131.93	-4.21	-1319.74437012	-1319.87175596	25.86
128.22	-3.42	-1319.74762603	-1319.84821045	29.53
125.02	-2.63	-1319.74985141	-1319.82528837	33.03
122.21	-1.84	-1319.75129893	-1319.80290843	36.40
119.70	-1.05	-1319.75213646	-1319.78100994	39.65
117.43	-0.26	-1319.75248193	-1319.75954364	42.81
115.35	0.53	-1319.75242166	-1319.73847018	45.89
113.45	1.32	-1319.75202088	-1319.71775646	48.90
111.68	2.11	-1319.75133026	-1319.69737458	51.86
110.04	2.90	-1319.75038992	-1319.67730070	54.75
108.51	3.69	-1319.74923235	-1319.65751494	57.61
107.06	4.47	-1319.74788404	-1319.63799827	60.41
105.71	5.26	-1319.74636708	-1319.61873479	63.18
104.42	6.05	-1319.74470011	-1319.59971153	65.91
103.20	6.84	-1319.74289869	-1319.58091375	68.61
102.04	7.63	-1319.74097644	-1319.56233179	71.28
100.94	8.42	-1319.73894489	-1319.54395405	73.91
99.89	9.21	-1319.73681430	-1319.52577266	76.53
98.88	10.00	-1319.73459329	-1319.50777698	79.11

Chapter 13

Calculation of Elastic, Piezoelectric and Photoelastic Tensors

The program can compute elastic, piezoelectric and photoelastic (elasto-optic) tensors with a fully-automated procedure by means of keywords **ELASTCON**, **PIEZOCON** (or **PIEZOCP**) and **PHOTOELA**, to be inserted at the end of the Geometry input block. Elastic and piezoelectric constants can be computed at once with the **ELAPIEZO** keyword. Each of these keywords opens a block which must be terminated with an **END** keyword.

13.1 A few theoretical remarks

The elements of the elastic tensor, for 3D systems, are usually defined as:

$$C_{vu} = \frac{1}{V} \left. \frac{\partial^2 E}{\partial \eta_v \partial \eta_u} \right|_0 , \tag{13.1}$$

where η is the rank-2 symmetric tensor of pure strain and Voigt's notation is used according to which $v, u = 1, \ldots, 6$ ($1 = xx$, $2 = yy$, $3 = zz$, $4 = yz$, $5 = xz$, $6 = xy$). Since volume V is not uniquely defined for 1D and 2D systems, it is here omitted (length or surface could be used instead) and all the elements involving non-periodic directions (y, z for 1D and z for 2D systems) are null by definition. As a consequence, for 1D and 2D systems, elastic constants are expressed in energy units (hartree). Second derivatives in equation (13.1) are computed as first numerical derivatives of analytical energy gradients in the present implementation.

In the linear regime, *direct* **e** and *converse* **d** piezoelectric tensors describe the polarization **P** induced by strain η and the strain induced by an external electric field **E** at constant electric field and stress, respectively:

$$\textit{direct} \quad \text{effect} \quad \mathbf{P} = \mathbf{e}\,\eta \quad \text{at constant field} \tag{13.2}$$

$$\textit{converse} \quad \text{effect} \quad \eta = \mathbf{d}^T\mathbf{E} \quad \text{at constant stress} \tag{13.3}$$

Our approach consists in directly computing the intensity of polarization induced by strain (again, since the volume is not defined for 1D and 2D systems, in those cases the polarization reduces to a dipole moment). The Cartesian components of the polarization can then be expressed as follows in terms of the strain tensor components:

$$P_i = \sum_v e_{iv}\eta_v \quad \text{so that} \quad e_{iv} = \left(\frac{\partial P_i}{\partial \eta_v}\right)_E . \tag{13.4}$$

In the above expression, $i = 1, \ldots, 3$; η is the pure strain tensor, the derivative is taken at constant electric field and Voigt's notation is used. In CRYSTAL the polarization can be

computed either via localized Wannier functions or via the Berry phase (BP) approach. The latter scheme is used in the present automated implementation according to which direct piezoelectric constants can be written as follows in terms of numerical first derivatives of the BP φ_l with respect to the strain:

$$e_{iv} = \frac{|e|}{2\pi V} \sum_l a_{li} \frac{\partial \varphi_l}{\partial \eta_v} \,, \tag{13.5}$$

where a_{li} is the i-th Cartesian component of the l-th direct lattice basis vector $\mathbf{a}_l$. Again, for 1D and 2D systems, the volume is omitted, the polarization reduces to a dipole moment and the piezoelectric constants are reported in units of $|e| \times bohr$.
An alternative approach has recently been implemented for the analytical evaluation of the direct piezoelectric tensor of 3D crystals, based on the CPHF/KS scheme. This strategy is adopted when use of the keyword **PIEZOCP** is made.
A simple direct connection exists between *direct* **e** and *converse* **d** piezoelectric tensors:

$$\mathbf{e} = \mathbf{d}\,\mathbb{C} \quad \text{and} \quad \mathbf{d} = \mathbf{e}\,\mathbb{S} \,, \tag{13.6}$$

where $\mathbb{C}$ is the fourth rank elastic tensor of energy second derivatives with respect to pairs of deformations and $\mathbb{S} = \mathbb{C}^{-1}$ is the fourth rank compliance tensor.

Photoelastic constants are the elements of the fourth rank photoelastic (Pockels) tensor and are defined as:

$$p_{ijkl} = \frac{\partial \Delta \epsilon^{-1}_{ij}}{\partial \eta_{kl}} \,, \tag{13.7}$$

where $\Delta\epsilon^{-1}$ is the difference of the inverse dielectric tensor between strained and unstrained configurations. Given the stress-strain relation, the fourth-rank piezo-optic tensor $\boldsymbol{\pi}$ (whose elements are the stress-optical coefficients π_{vu}) can be obtained from the photo-elastic **p** and elastic $\mathbb{C}$ ones as:

$$\boldsymbol{\pi} = \mathbf{p}\,\mathbb{S} \qquad and \qquad \mathbf{p} = \boldsymbol{\pi}\,\mathbb{C} \,. \tag{13.8}$$

At variance with the elastic $\mathbb{C}$ and compliance $\mathbb{S}$ tensors, **p** and $\boldsymbol{\pi}$ are not symmetric (i.e. in general $p_{vu} \neq p_{uv}$ and $\pi_{vu} \neq \pi_{uv}$). It follows that the number of symmetry-independent components to be determined for the stress-optical and strain-optical tensors is generally larger than for the elastic tensors.

The derivatives in the right-hand-sides of equations (13.1), (13.5) and (13.7) are computed numerically by applying finite strains to the crystal lattice. For each strain, N_s configurations are defined according to a strain step δ. By default, $N_s = 3$, corresponding to one "expanded", one unstrained and one "contracted" configuration, and $\delta = 0.01$ for elastic and $\delta = 0.015$ for piezoelectric and photoelastic constants. Parameter N_s can be modified with the sub-keyword **NUMDERIV** while δ can be modified with sub-keyword **STEPSIZE**.
We recall that elastic, piezoelectric and photoelastic constants can be decomposed into purely electronic "clamped-ion" and nuclear "internal-strain" contributions; the latter, computed by default, measures the effect of relaxation of the relative positions of atoms induced by the strain and can be computed by optimizing the atomic positions within the crystal cell. If one wants to compute "clamped-ion" constants can use the sub-keyword **CLAMPION**.
The input geometry is assumed to be optimized; nevertheless, the user can ask this module to perform a pre-optimization of the structure by means of the **PREOPTGEOM** sub-keyword; convergence tolerances on gradient and displacement can be modified by means of the **TOLDEG** and **TOLDEX** sub-keywords.
For elastic, piezoelectric and photoelastic constants a flexible restart option has been prepared which can be activated with the **RESTART** sub-keyword and which uses an external formatted file called ELASINFO.DAT. Detailed printings (recommended) can be activated with the **PRINT** sub-keyword.

13.2 The algorithm

We present here the fully automated procedure for the calculation of the elastic, piezoelectric and photoelastic constants as implemented in CRYSTAL:

1. The starting geometrical structure of the crystal is accurately optimized as concerns both lattice parameters and atomic positions;

2. A single-point self-consistent-field (SCF) calculation is performed on the optimized reference structure. Energy gradients, Berry phase or dielectric tensor (via a CPHF/KS scheme) are also computed;

3. A symmetry analysis is performed in order to find the minimal set of strains, out of a maximum of six, which have to be explicitly applied in order to get all the independent constants. The symmetry features of third and fourth rank tensors is printed in the following way:

```
|  1 A    1 F    1 C     .     1 B     .     |
|         1 A    1 C     .    -1 B     .     |
|                1 E     .      .      .     |
|                       1 D     .    -1 B    |
|                               1 D     .     |
|                                       &     |
```

4. For each irreducible strain, the deformation is applied to the structure and the residual symmetry determined. N_s strained configurations are defined according to a strain step δ;

5. For each strained configuration, the atomic positions are relaxed with an optimization (default option) or not depending on whether one wants to go beyond the "clamped-ion" approximation or not. An SCF calculation is then performed with gradient, Berry phase or dielectric tensor calculation;

6. After the loop over the N_s strained configurations, the energy gradients, Berry phases or dielectric tensors are fitted with singular-value-decomposition routines and their first derivatives determined numerically; Elastic, piezoelectric or photoelastic constants are finally computed and printed.

13.3 Second-order Elastic Constants

A fully-automated procedure for calculating the second-order elastic constants for an arbitrary crystal is activated by specifying the keyword **ELASTCON** in input block 1 (geometry). **ELASTCON** must be the last keyword in geometry input:

```
. . . geometry input . . . .
ELASTCON
END        close ELASTCON input block
END        close geometry input block
```

Note that the user must be confident that the input structure is already well optimized. At the beginning of the run, the forces at the central point are calculated and if they exceed 10^{-4}, then a warning is printed suggesting re-optimization.

Users of this option are kindly requested to cite the following papers[211, 93]:

W.F. Perger, J. Criswell, B. Civalleri and R. Dovesi, *Comp. Phys. Comm.*, **180**, 1753-1759 (2009). *Ab-initio calculation of elastic constants of crystalline systems with the CRYSTAL code*

A. Erba, A. Mahmoud, R. Orlando and R. Dovesi, *Phys. Chem. Minerals*, **41**, 151-160 (2014). *Elastic properties of six silicate garnet end-members from accurate ab initio simulations.*

Users of this option for 1D and 2D systems are kindly requested to cite the following paper[90]:

A. Erba, M. Ferrabone, J. Baima, R. Orlando, M. Rérat and R. Dovesi, *J. Chem. Phys.*, **138**, 054906 (2013). *The vibration properties of the (n,0) Boron Nitride nanotubes from ab initio quantum chemical simulations.*

Keywords, options, and defaults

A default value is chosen for all computational parameters. SCF energy convergence threshold is set to 10^{-8}. To modify it, see keyword **TOLDEE** in input block 3, page 125.

Default choices can be modified by optional keywords (in any order):

rec	variable	meaning
• *A*	**NUMDERIV**	This sets the number of points for the numerical second derivative
∗	INUM	number of points including the central (zero displacement) one [$default = 3$]
• *A*	**STEPSIZE**	This gives the size of the displacement to be used for the calculation of the numerical second derivative.
∗	STEP	size of the strain step along a given deformation [$default = 0.01$]
• *A*	**DEFORM**	Specific deformations are asked for
∗	DEF(I), I=1,6	Six integers have to be provided, each one associated to a given deformation (xx, yy, zz, yz, xz, xy) that tell the program which deformations have to be considered as active. Put 1 for active and 0 for inactive. By default, the program performs a symmetry analysis and finds which deformations are necessary.

rec	variable	meaning
• *A*	**CLAMPION**	This option activates the computation of "clamped-ion" constants.
• *A*	**PRINT**	This option turns on a higher level of diagnostic printing [$default$ minimal printing].
• *A*	**RESTART**	Allows restart using file $ELASINFO.DAT$ from a previous incomplete run.
• *A*	**RESTART2**	Allows a full restart using file $ELASINFO.DAT$ from a previous complete run.
• *A*	**PREOPTGEOM**	A preliminary geometry optimization of cell and atomic positions is performed before starting elastic constant calculation.
• *A*	**SEISMDIR**	Defines additional directions along which seismic wave velocities are computed.
∗	NDIR	Number of additional directions
		Insert NDIR records II
∗	DX,DY,DZ	Cartesian components of each additional direction
		II

Geometry optimization is performed at convergence criteria tighter than the ones given in **OPTGEOM**. Some values can be modified by inserting the following keywords:

• *A*	**TOLDEG**	ELASTCON default [0.0003] - see **OPTGEOM**, page 188
• *A*	**TOLDEX**	ELASTCON default [0.00062] - see **OPTGEOM**, page 188
• *A*	**TOLDEE**	ELASTCON default [8] - see **OPTGEOM**, page 188
• *A*	**MAXCYCLE**	see **OPTGEOM**, page 190
• *A*	**NOTRUSTR**	see **OPTGEOM**, page 188
• *A*	**TRUSTRADIUS**	see **OPTGEOM**, page 188
• *A*	**MAXTRADIUS**	see **OPTGEOM**, page 188

Input Example:

```
. . . geometry input . . . .
ELASTCON
NUMDERIV   modify default choice of n. points [3]
5
STEPSIZE   modify step size for numerical gradient [0.01]
0.02
PRINT      extended printing
END        end of ELASTCON input
END        geometry input
```

The information on the computed elastic constants is printed at the end of the output with the following format:

```
         SYMMETRIZED ELASTIC CONSTANTS FOR HEXAGONAL CASE, IN GPa

        |    89.130   13.792   20.473    0.000   12.584    0.000 |
        |             89.130   20.473    0.000  -12.584    0.000 |
        |                     113.498    0.000    0.000    0.000 |
        |                               58.682    0.000  -12.584 |
        |                                        58.682    0.000 |
        |                                                 37.669 |
```

Compliance tensor is also printed as:

```
              ELASTIC MODULI (COMPLIANCE TENSOR), IN TPa^-1

     |       12.382521 -1.915287 -1.888081 0.000000 -3.066002 0.000000 |
     |                 12.382521 -1.888081 0.000000  3.066002 0.000000 |
     |                           9.491835 0.000000  0.000000 0.000000 |
     |                                   18.355822  0.000000 6.132004 |
     |                                              18.355822 0.000000 |
     |                                                        28.595616 |
```

According to the elastic continuum theory, the three acoustic wave velocities of a crystal, along any general direction represented by unit wave-vector $\hat{\mathbf{q}}$, are related to the elastic constants by Christoffel's equation which can be given an eigenvalues/eigenvectors form, as follows:

$$\mathbf{A}^{\hat{\mathbf{q}}}\mathbf{U} = \mathbf{V}^2\mathbf{U} \qquad \text{with} \qquad A^{\hat{\mathbf{q}}}_{kl} = \frac{1}{\rho}\hat{q}_i C_{iklj}\hat{q}_j \ , \tag{13.9}$$

where ρ is the crystal density, $i,j,k,l = x,y,z$ represent Cartesian directions, $\hat{q}_i$ is the i-th element of the unit vector $\hat{\mathbf{q}}$, $\mathbf{V}$ is a 3×3 diagonal matrix whose three elements give the acoustic velocities and $\mathbf{U} = (\hat{\mathbf{u}}_1,\hat{\mathbf{u}}_2,\hat{\mathbf{u}}_3)$ is the eigenvectors 3×3 matrix where each column represents the polarization $\hat{\mathbf{u}}$ of the corresponding eigenvalue. The three acoustic wave velocities, also referred to as seismic velocities, can be labeled as longitudinal v_p, slow transverse v_{s1} and fast transverse v_{s2}, depending on the polarization direction $\hat{\mathbf{u}}$ with respect to wave-vector $\hat{\mathbf{q}}$.
The seismic wave velocities are computed by default along some Cartesian directions and printed as follows (note that slow and fast transverse wave velocities are sorted according to their values; crossing are possible which should be carefully checked):

```
 SEISMIC VELOCITIES BY CHRISTOFFEL EQUATION (km/s)

       WAVE VECTOR                Vp        Vs1       Vs2

 [ 0.000  0.000  1.000]          9.237     5.068     5.068
 [ 0.000  1.000  0.000]          9.237     5.068     5.068
 [ 1.000  0.000  0.000]          9.237     5.068     5.068
```

```
[ 1.000  1.000  0.000]          9.166     5.195     5.068
[ 1.000  0.000  1.000]          9.166     5.195     5.068
[ 0.000  1.000  1.000]          9.166     5.195     5.068
[ 1.000  1.000  1.000]          9.142     5.153     5.153
```

User-defined Cartesian directions can be added by means of the **SEISMDIR** sub-keyword, along which the seismic wave velocities are computed and printed (typically, this can be done *a posteriori* with a complete restart using the **RESTART2** sub-keyword).

A more complete analysis of directional elastic wave velocities can be performed via the AWE-SoMe program (by D. Munoz-Santiburcio, A. Hernandez-Laguna and J.I. Soto, *Comp. Phys. Commun.*, **192**, 272-277 (2015)), which has been merged with CRYSTAL17. By inserting the **AWESOME** keyword, the full analysis of phase and group velocities is enabled, which generates a series of output files specifically designed to be read by Gnuplot scripts to get 2D and 3D representations of directional seismic wave velocities and related quantities (Gnuplot scripts are available on our on-line tutorial):

```
. . . geometry input . . . .
ELASTCON
AWESOME
END         end of ELASTCON input
END         geometry input
```

13.3.1 Elastic Tensor under Pressure

Reference to be cited when using this module:
A. Erba, A. Mahmoud, D. Belmonte and R. Dovesi, *J. Chem. Phys.*, **140**, 124703 (2014).

The calculation of elastic constants under pressure is a task of great geophysical interest. In the absence of any finite pre-stress, elastic constants can be defined as second energy density derivatives with respect to pairs of infinitesimal Eulerian strains:

$$C_{ijkl} = \frac{1}{V_0}\left(\frac{\partial^2 E}{\partial\epsilon_{ij}\partial\epsilon_{kl}}\right)_{\epsilon=0} . \tag{13.10}$$

The constants above do represent the link between stress and strain via the Hooke's law. In the limit of zero temperature, typical of ab initio simulations, they are also referred to as athermal elastic constants.
If a finite pre-stress $\boldsymbol{\sigma}^{pre}$ is applied in the form of a hydrostatic pressure P ($\sigma_{ij}^{pre} = P\delta_{ij}$), within the frame of finite Eulerian strain, the relevant elastic stiffness constants are transformed as:

$$B_{ijkl} = C_{ijkl} + \frac{P}{2}(2\delta_{ij}\delta_{kl} - \delta_{il}\delta_{jk} - \delta_{ik}\delta_{jl}) , \tag{13.11}$$

provided that V_0 in equation 13.10 becomes the equilibrium volume $V(P)$ at pressure P. In the present fully automated implementation of the calculation of the stiffness tensor $\mathbf{B}$ (and of $\mathbf{S} = \mathbf{B}^{-1}$, the compliance tensor) under pressure, $V(P)$ is obtained from the analytical stress tensor. An option also exists for using the $V(P)$ relation obtained from a previous EOS calculation.
Since both $\boldsymbol{\epsilon}$ and $\boldsymbol{\delta}$ are symmetric tensors, we can rewrite the previous equality as:

$$B_{vu} = C_{vu} + \begin{pmatrix} 0 & P & P & 0 & 0 & 0 \\ P & 0 & P & 0 & 0 & 0 \\ P & P & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{-P}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-P}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{-P}{2} \end{pmatrix} , \tag{13.12}$$

where Voigt's notation has been used.

So, in order to obtain the elastic constants at a given pressure, one basically needs to perform an elastic calculation at the volume corresponding to the desired pressure, $V(P)$, and then correct the computed constants according to equation 13.12. This is automatically done by the program if the following keywords are used:

```
. . . geometry input . . . .
ELASTCON
PREOPTGEOM  A constant-pressure pre-optimization is performed
PRESSURE    Activates the calculation under pressure
5           Set the value (in GPa) of the pressure
END         end of ELASTCON input
END         geometry input
```

In this way, a pressure-constrained pre-geometry optimization is performed. The elastic constants are evaluated starting from the optimized structure.

If one already knows which is the equilibrium structure corresponding to a given pressure (from a previous EOS run, for instance), there is no need to perform a geometry optimization before the elastic constants evaluation. In this case, one should give the optimized structure in input, which refers to a given pressure, and then use the following keywords:

```
. . . geometry input . . . .
ELASTCON
PRESSEOS   Activates the calculation under pressure
5          Set the value (in GPa) of the pressure corresponding to the given structure
END         end of ELASTCON input
END         geometry input
```

13.3.2 Thermo-Elastic Constants

Via a quasi-harmonic approach, the thermal dependence of the elastic constants can be computed.

Users of this option are kindly reminded to cite the following papers:
M. Destefanis, C. Ravoux, A. Cossard and A. Erba, *Minerals*, **9**, 16 (2019) *Thermo-Elasticity of Materials from Quasi-Harmonic Calculations*

J. Maul, D. Ongari, S. M. Moosavi, B. Smit and A. Erba, *J. Phys. Chem. Lett.*, **11**, 8543–8548 (2020) *Thermoelasticity of Flexible Organic Crystals from Quasi-harmonic Lattice Dynamics: The Case of Copper(II) Acetylacetonate*

The implemented options for thermo-elasticity are described in Section 9.4 at page 259.

13.3.3 Nuclear-relaxation Term from Internal-strain Tensor

References to be cited for this option:
A. Erba, *Phys. Chem. Chem. Phys.*, **18**, 13984-13992 (2016)
A. Erba, D. Caglioti, C. M. Zicovich-Wilson and R. Dovesi, *J. Comput. Chem.*, **38**, 257-264 (2017)

Strain-induced response properties of solids can be formally decomposed into a purely electronic "clamped-nuclei" term and a nuclear-relaxation term due to the rearrangement of atomic positions upon strain. The evaluation of the latter is generally much more computationally expensive than that of the former. In principle, two alternative approaches can be used to account for nuclear-relaxation effects: i) performing numerical geometry optimizations to relax atomic positions at actual strained lattice configurations, or ii) evaluating in a more analytical fashion the "internal-strain" tensor of energy second-derivatives with respect to atomic displacements and lattice deformations, as combined with the interatomic force constant Hessian matrix. Given that geometry optimizations at strained configurations are rather slowly-converging numerical procedures requiring particularly tight convergence criteria, the second approach is to

be preferred as it ensures higher accuracy and requires less severe computational parameters to be used.
As a matter of fact, in previous versions of the CRYSTAL program, only the first numerical strategy was available. In CRYSTAL17, the second approach has been implemented for the elastic and piezoelectric tensors, which relies on the calculation of the "internal-strain" tensor by fully-exploiting its point-symmetry features. Beside being more robust, the new strategy has also been documented to be more computationally efficient for most crystalline systems.
The elements of the force-response internal-strain tensor are second-energy derivatives with respect to an atomic displacement and to a lattice distortion:

$$\Lambda_{ai,v} = \left. \frac{\partial^2 E}{\partial u_{ai} \partial \eta_v} \right|_{\mathcal{E}} , \tag{13.13}$$

where u_{ai} are Cartesian components of the displacement vector $\mathbf{u}_a$ of atom a (i=x, y, z). A displacement-response internal-strain tensor $\mathbf{\Gamma}$, which describes first-order atomic displacements as induced by a first-order strain, can be defined as:

$$\Gamma_{ai,v} = - \left. \frac{\partial u_{ai}}{\partial \eta_v} \right|_{\mathcal{E}} = \sum_{bj} (H^{-1})_{ai,bj} \Lambda_{bj,v} , \tag{13.14}$$

where $\mathbf{H}$ is the interatomic force-constant Hessian matrix of energy second derivatives with respect to pairs of periodicity-preserving atomic displacements:

$$H_{ai,bj} = \left. \frac{\partial^2 E}{\partial u_{ai} \partial u_{bj}} \right|_{\mathcal{E},\eta} . \tag{13.15}$$

When mass-weighted and diagonalized, the force-constant matrix of Eq. (13.15) provides vibration frequencies of Brillouin zone-center phonon modes. The $\mathbf{H}^{-1}$ matrix in Eq. (13.14) has to be considered a pseudoinverse of $\mathbf{H}$ where translational degrees of freedom are projected out, as discussed in detail elsewhere.
The nuclear-relaxation contribution to elastic constants can be expressed in terms of the internal-strain tensor $\mathbf{\Lambda}$ (or $\mathbf{\Gamma}$):

$$C_{vw}^{\text{nuc}} = -\frac{1}{V_0} \sum_{ai} \Lambda_{ai,v} \Gamma_{ai,w} , \tag{13.16}$$

In the current fully-automated implementation into the CRYSTAL program, the elements $\Lambda_{ai,v}$ of the force-response internal-strain tensor are here computed as finite differences of analytical lattice gradients with respect to atomic Cartesian displacements, by means of a generalized "Pulay's force method" originally proposed for interatomic force constants.

To activate this option (recommended) use the following keywords:

```
. . . geometry input . . . .
ELASTCON
NUCHESS   Activates use of Hessian matrices for nuclear term
HESSNUM2   Use of a 2-points formula in the evaluation of the Hessian matrices
END        end of ELASTCON input
END        geometry input
```

The keyword **MODEPART** activates the partition of the nuclear-relaxation term into normal modes. The keyword **HESSSTEP** allows to redefine the step (in Å) for the numerical evaluation of the Hessian matrices:

```
. . . geometry input . . . .
ELASTCON
NUCHESS   Activates use of Hessian matrices for nuclear term
HESSNUM2   Use of a 2-points formula in the evaluation of the Hessian matrices
MODEPART
HESSSTEP
0.005
END        end of ELASTCON input
END        geometry input
```

Internal Diagnostics

In order to monitor the quality of the calculation as it proceeds, the total energy after optimization is stored. The recommended use of the **ELASTCON** option assumes that the user supplies an input file from a previously optimized geometry (and not experimental lattice constants and atomic positions, for example). Therefore, in principle, as the various deformations are made, the optimized total energy for each of the deformed geometries should be higher than the energy at the undeformed, equilibrium, geometry. The code monitors each optimized total energy for each deformed geometry and if any deformation lowers the total energy from the equilibrium value, a warning is printed for the user to verify that the input file was really from a previously optimized geometry.

13.4 First-order Piezoelectric Constants

A fully-automated procedure for calculating the first-order piezoelectric constants for an arbitrary crystal is activated by specifying the keyword **PIEZOCON** in input block 1 (geometry). **PIEZOCON** must be the last keyword in geometry input:

```
. . . geometry input . . . .
PIEZOCON
END        close PIEZOCON input block
END        close geometry input block
```

Note that the user must be confident that the input structure is already well optimized. At the beginning of the run, the forces at the central point are calculated and if they exceed 10^{-4}, then a warning is printed suggesting re-optimization.
Due to techinical reasons, the use of PIEZOCON keyword is supported, in CRYSTAL22 by *Pcrystal* only. The program will issue an error and stop if *MPPcrystal* executable is selected and PIEZOCON keyword is active.

Users of this option are kindly requested to cite the following papers[196, 89]:

Y. Noel and C. M. Zicovich-Wilson and B. Civalleri and Ph. D'Arco and R. Dovesi, *Phys. Rev. B*, **65**, 014111 (2001). *Polarization properties of ZnO and BeO: An ab initio study through the Berry phase and Wannier functions approaches.*

A. Erba, Kh. E. El-Kelany, M. Ferrero, I. Baraille and M. Rérat, *Phys. Rev. B*, **88**, 035102 (2013). *Piezoelectricity of $SrTiO_3$: An ab initio description.*

Users of this option for 1D and 2D systems are kindly requested to cite the following paper[90]:

A. Erba, M. Ferrabone, J. Baima, R. Orlando, M. Rérat and R. Dovesi, *J. Chem. Phys.*, **138**, 054906 (2013). *The vibration properties of the (n,0) Boron Nitride nanotubes from ab initio quantum chemical simulations.*

Keywords, options, and defaults

A default value is chosen for all computational parameters. SCF energy convergence threshold is set to 10^{-8}. To modify it, see keyword **TOLDEE** in input block 3, page 125.

Default choices can be modified by optional keywords (in any order):

rec	variable	meaning
•	*A* **NUMDERIV**	This sets the number of points for the numerical second derivative
	∗ INUM	number of points including the central (zero displacement) one [$default = 3$]
•	*A* **STEPSIZE**	This gives the size of the displacement to be used for the calculation of the numerical second derivative.
	∗ STEP	size of the strain step along a given deformation [$default = 0.015$]
•	*A* **DEFORM**	Specific deformations are asked for
	∗ DEF(I), I=1,6	Six integers have to be provided, each one associated to a given deformation (xx, yy, zz, yz, xz, xy) that tell the program which deformations have to be considered as active. Put 1 for active and 0 for inactive. By default, the program performs a symmetry analysis and finds which deformations are necessary.

rec	variable	meaning
•	*A* **CLAMPION**	This option activates the computation of "clamped-ion" constants.
•	*A* **PRINT**	This option turns on a higher level of diagnostic printing [$default$ minimal printing]
•	*A* **RESTART**	Allows restart using file $ELASINFO.DAT$ from a previous incomplete run.
•	*A* **RESTART2**	Allows a full restart using file $ELASINFO.DAT$ from a previous complete run.
•	*A* **PREOPTGEOM**	A preliminary geometry optimization of cell and atomic positions is performed before starting elastic constant calculation.

Geometry optimization is performed at convergence criteria tighter than the ones given in **OPTGEOM**. Some values can be modified by inserting the following keywords:

• *A* **TOLDEG**	PIEZOCON default [0.0003] - see **OPTGEOM**, page 188
• *A* **TOLDEX**	PIEZOCON default [0.00062] - see **OPTGEOM**, page 188
• *A* **TOLDEE**	PIEZOCON default [8] - see **OPTGEOM**, page 188
• *A* **MAXCYCLE**	see **OPTGEOM**, page 190
• *A* **NOTRUSTR**	see **OPTGEOM**, page 188
• *A* **TRUSTRADIUS**	see **OPTGEOM**, page 188
• *A* **MAXTRADIUS**	see **OPTGEOM**, page 188

Input Example:

```
. . . geometry input . . . .
PIEZOCON
NUMDERIV   modify default choice of n. points [3]
5
STEPSIZE   modify step size for numerical gradient [0.015]
0.02
CLAMPION   activates "clamped-ion" approximation
PRINT      extended printing
END        end of PIEZOCON input
END        geometry input
```

The computed piezoelectric tensor is printed as follows at the end of the output:

```
          PIEZOELECTRIC CONSTANTS FOR HEXAGONAL CASE, IN C/m^2

       |     0.000    0.000    0.000    0.000   -0.090    0.000 |
       |     0.000    0.000    0.000   -0.090    0.000    0.000 |
       |    -0.187   -0.187    0.282    0.000    0.000    0.000 |
```

13.5 Piezoelectricity through CPHF/KS Approach

References to be cited when using this module:
J. Baima, A. Erba, L. Maschio, C.M. Zicovich-Wilson, R. Dovesi and B. Kirtman, *Z. Phys. Chem.*, **230**, 719-736 (2016)
A. Erba, *Phys. Chem. Chem. Phys.*, **18**, 13984-13992 (2016)

While the **PIEZOCON** keyword allows to compute the direct piezoelectric tensor through the numerical Berry phase approach, the **PIEZOCP** keyword allows to compute it via an analytical approach based on the CPHF/KS scheme. In this way, not just the electronic term is computed analytically, but also the nuclear-relaxation contribution is evaluated from the internal-strain tensor rather than from numerical geometry optimizations of atomic coordinates at strained configurations:

$$e_{kv}^{\rm nuc} \quad = \quad -\frac{1}{V_0}\sum_{ai} Z^*_{k,ai}\Gamma_{ai,v} \ , \tag{13.17}$$

where the $\mathbf{Z}^*$ tensor in Eq. (13.17) contains the Born dynamical effective charges:

$$Z^*_{k,ai} = \left.\frac{\partial^2 E}{\partial\mathcal{E}_k\partial u_{ai}}\right|_\eta \ . \tag{13.18}$$

```
. . . geometry input . . . .
PIEZOCP  Activates the analytical calculation of the piezoelectric tensor
HESSNUM2 Use of a 2-point formula for the evaluation of the Hessian matrices
TOLALPHA  Set the tolerance on the convergence of the CPHF process
2
END        end of PIEZOCP input
END        geometry input
```

13.6 Elastic and Piezoelectric Constants

A fully-automated procedure for calculating the second-order elastic constants and the first-order piezoelectric (direct and converse) constants for an arbitrary crystal is activated by specifying the keyword **ELAPIEZO** in input block 1 (geometry).
ELAPIEZO must be the last keyword in geometry input:

```
. . . geometry input . . . .
ELAPIEZO
END        close ELAPIEZO input block
END        close geometry input block
```

Note that the user must be confident that the input structure is already well optimized. At the beginning of the run, the forces at the central point are calculated and if they exceed 10^{-4}, then a warning is printed suggesting re-optimization.
Due to techinical reasons, the use of ELAPIEZO keyword is supported, in CRYSTAL22 by *Pcrystal* only. The program will issue an error and stop if *MPPcrystal* executable is selected and ELAPIEZO keyword is active.

Users of this option are kindly requested to cite the following papers[196, 89]:

Y. Noel and C. M. Zicovich-Wilson and B. Civalleri and Ph. D'Arco and R. Dovesi, *Phys. Rev. B*, **65**, 014111 (2001). *Polarization properties of ZnO and BeO: An ab initio study through the Berry phase and Wannier functions approaches.*

A. Erba, Kh. E. El-Kelany, M. Ferrero, I. Baraille and M. Rérat, *Phys. Rev. B*, **88**, 035102 (2013). *Piezoelectricity of $SrTiO_3$: An ab initio description.*

Users of this option for 1D and 2D systems are kindly requested to cite the following paper[90]:

A. Erba, M. Ferrabone, J. Baima, R. Orlando, M. Rérat and R. Dovesi, *J. Chem. Phys.*, **138**, 054906 (2013). *The vibration properties of the (n,0) Boron Nitride nanotubes from ab initio quantum chemical simulations.*

Keywords, options, and defaults

A default value is chosen for all computational parameters. SCF energy convergence threshold is set to 10^{-8}. To modify it, see keyword **TOLDEE** in input block 3, page 125.

Default choices can be modified by optional keywords (in any order):

rec	variable	meaning
•	*A* **NUMDERIV**	This sets the number of points for the numerical second derivative
	∗ INUM	number of points including the central (zero displacement) one [$default = 3$]
•	*A* **STEPSIZE**	This gives the size of the displacement to be used for the calculation of the numerical second derivative.
	∗ STEP	size of the strain step along a given deformation [$default = 0.015$]
•	*A* **DEFORM**	Specific deformations are asked for
	∗ DEF(I), I=1,6	Six integers have to be provided, each one associated to a given deformation (xx, yy, zz, yz, xz, xy) that tell the program which deformations have to be considered as active. Put 1 for active and 0 for inactive. By default, the program performs a symmetry analysis and finds which deformations are necessary.

rec	variable	meaning
•	*A* **CLAMPION**	This option activates the computation of "clamped-ion" constants.
•	*A* **PRINT**	This option turns on a higher level of diagnostic printing [$default$ minimal printing]
•	*A* **RESTART**	Allows restart using file $ELASINFO.DAT$ from a previous incomplete run.
•	*A* **RESTART2**	Allows a full restart using file $ELASINFO.DAT$ from a previous complete run.
•	*A* **PREOPTGEOM**	A preliminary geometry optimization of cell and atomic positions is performed before starting elastic constant calculation.

Geometry optimization is performed at convergence criteria tighter than the ones given in **OPTGEOM**. Some values can be modified by inserting the following keywords:

• *A* **TOLDEG**	ELAPIEZO default [0.0003] - see **OPTGEOM**, page 188
• *A* **TOLDEX**	ELAPIEZO default [0.00062] - see **OPTGEOM**, page 188
• *A* **TOLDEE**	ELAPIEZO default [8] - see **OPTGEOM**, page 188
• *A* **MAXCYCLE**	see **OPTGEOM**, page 190
• *A* **NOTRUSTR**	see **OPTGEOM**, page 188
• *A* **TRUSTRADIUS**	see **OPTGEOM**, page 188
• *A* **MAXTRADIUS**	see **OPTGEOM**, page 188

Input Example:

```
. . . geometry input . . . .
ELAPIEZO
NUMDERIV   modify default choice of n. points [3]
5
STEPSIZE   modify step size for numerical gradient [0.015]
0.02
```

```
CLAMPION   activates "clamped-ion" approximation
PRINT      extended printing
END        end of ELAPIEZO input
END        geometry input
```

Elastic and piezoelectric tensors are printed at the end of the output (see keywords **ELASTCON** and **PIEZOCON** above). Moreover, piezoelectric strain tensor (converse piezoelectric tensor) is printed:

```
 PIEZOELECTRIC STRAIN (CONVERSE) CONSTANTS (pC/N = pm/V)

|     0.000    0.000    0.000    0.000  -14.351    0.000 |
|     0.000    0.000    0.000  -14.351    0.000    0.000 |
|    -0.256   -0.256    2.859    0.000    0.000    0.000 |
```

13.7 Photoelastic Constants

A fully-automated procedure for calculating the photoelastic and piezo-optic constants for an arbitrary crystal is activated by specifying the keyword **PHOTOELA** in input block 1 (geometry). Dielectric tensor is computed via a CPHF/KS procedure. The electronic contribution is evaluated in the limit of infinite frequency $\omega \to \infty$. Dielectric tensor at finite frequency can be computed with sub-keyword **DYNAMIC**.
PHOTOELA must be the last keyword in geometry input:

```
. . . geometry input . . . .
PHOTOELA
END        close PHOTOELA input block
END        close geometry input block
```

Note that the user must be confident that the input structure is already well optimized. At the beginning of the run, the forces at the central point are calculated and if they exceed 10^{-4}, then a warning is printed suggesting re-optimization.

Users of this option are kindly requested to cite the following papers[88]:

A. Erba and R. Dovesi, *Phys. Rev. B*, **88**, 045121 (2013). *Photoelasticity of crystals from theoretical simulations.*
A. Erba, M.T. Ruggiero, T. M. Korter and R. Dovesi, *J. Chem. Phys.*, **143**, 144504 (2015). *Piezo-Optic Tensor of Crystals from Quantum-Mechanical Calculations*

Keywords, options, and defaults

A default value is chosen for all computational parameters. SCF energy convergence threshold is set to 10^{-8}. To modify it, see keyword **TOLDEE** in input block 3, page 125.

Default choices can be modified by optional keywords (in any order). There are some specific sub-keywords controlling CPHF/KS parameters:

rec	variable	meaning
•	*A* **TOLALPHA**	threshold on energy first derivative change between CPHF/KS cycles
	∗ ITOL	$\|\Delta E\| < 10^{-ITOL}$ [default: 4]
•	*A* **FMIXING**	permits to mix the Fock/KS matrix derivatives between CPHF/KS cycles n and $n-1$
	∗ IPMIX	percentage of cycle $n-1$ [IPMIX=0, no mixing; default IPMIX=60]
•	*A* **DYNAMIC**	Activates the computation of frequency-dependent polarizabilities and dielectric constants.
	ILAMBD	wave-length of radiation (expressed in nm).
•	*A* **ANDERSON**	Convergence accelerator. Mixing of Fock/KS matrix derivatives (page 93)
•	*A* **BROYDEN**	Convergence accelerator. Mixing of Fock/KS matrix derivatives (page 95)
	∗ W0	*W0* parameter in Anderson's paper [154]
	∗ IMIX	Percentage of Fock/KS derivative matrices mixing
	∗ ISTART	Iteration after which Broyden method is activated (minimum 2)
•	*A* **NUMDERIV**	This sets the number of points for the numerical second derivative
	∗ INUM	number of points including the central (zero displacement) one [$default = 3$]
•	*A* **STEPSIZE**	This gives the size of the displacement to be used for the calculation of the numerical second derivative.
	∗ STEP	size of the strain step along a given deformation [$default = 0.015$]
•	*A* **DEFORM**	Specific deformations are asked for
	∗ DEF(I), I=1,6	Six integers have to be provided, each one associated to a given deformation (xx, yy, zz, yz, xz, xy) that tell the program which deformations have to be considered as active. Put 1 for active and 0 for inactive. By default, the program performs a symmetry analysis and finds which deformations are necessary.

rec	variable	meaning
•	*A* **CLAMPION**	This option activates the computation of "clamped-ion" constants.
•	*A* **PRINT**	This option turns on a higher level of diagnostic printing [$default$ minimal printing]
•	*A* **PREOPTGEOM**	A preliminary geometry optimization of cell and atomic positions is performed before starting elastic constant calculation.
•	*A* **RESTART**	Allows restart using file $ELASINFO.DAT$ from a previous incomplete run.
•	*A* **RESTART2**	Allows a full restart using file $ELASINFO.DAT$ from a previous complete run.

Geometry optimization is performed at convergence criteria tighter than the ones given in **OPTGEOM**. Some values can be modified by inserting the following keywords:

• *A* **TOLDEG**	PHOTOELA default [0.0003] - see **OPTGEOM**, page 188
• *A* **TOLDEX**	PHOTOELA default [0.00062] - see **OPTGEOM**, page 188
• *A* **TOLDEE**	PHOTOELA default [8] - see **OPTGEOM**, page 188
• *A* **MAXCYCLE**	see **OPTGEOM**, page 190
• *A* **NOTRUSTR**	see **OPTGEOM**, page 188
• *A* **TRUSTRADIUS**	see **OPTGEOM**, page 188
• *A* **MAXTRADIUS**	see **OPTGEOM**, page 188

Input Example:

```
. . . geometry input . . . .
PHOTOELA
NUMDERIV   modify default choice of n. points [3]
5
STEPSIZE   modify step size for numerical gradient [0.015]
0.02
CLAMPION   activates "clamped-ion" approximation
PRINT      extended printing
END        end of PHOTOELA input
END        geometry input
```

The elasto-optic (Pockels) tensor is printed at the end of the output:

```
 ELASTO-OPTIC (PHOTOELASTIC) POCKELS TENSOR

 |    -0.242   -0.010   -0.010    0.000    0.000    0.000 |
 |    -0.010   -0.242   -0.010    0.000    0.000    0.000 |
 |    -0.010   -0.010   -0.242    0.000    0.000    0.000 |
 |     0.000    0.000    0.000   -0.025    0.000    0.000 |
 |     0.000    0.000    0.000    0.000   -0.025    0.000 |
 |     0.000    0.000    0.000    0.000    0.000   -0.025 |
```

The piezo-optic tensor is then printed (in units of Brewsters):

```
 PIEZO-OPTIC TENSOR (IN BREWSTERS; 1B = 10^-12 Pa^-1 = 1 TPa^-1)

 |    -0.958    0.204    0.204    0.000    0.000    0.000 |
 |     0.204   -0.958    0.204    0.000    0.000    0.000 |
 |     0.204    0.204   -0.958    0.000    0.000    0.000 |
 |     0.000    0.000    0.000   -0.160    0.000    0.000 |
 |     0.000    0.000    0.000    0.000   -0.160    0.000 |
 |     0.000    0.000    0.000    0.000    0.000   -0.160 |
```

Chapter 14

One-electron Properties and Wave-function Analysis

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 324).
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 C, page 437, for information on printing.

14.1 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.

   The keyword **PMP2** (see page 344) reads the MP2 correction to the valence density matrix. Properties can then be computed from a MP2 corrected density matrix.

14.2 Properties keywords

RDFMWF wave function data conversion formatted-binary (fort.98 → fort.9)

Preliminary calculations			
NEWK	Eigenvectors calculation	341	I
COMMENS	Density Matrix commensurate to the Monchorst net	311	I
NOSYMADA	No symmetry Adapted Bloch Functions	117	–
PATO	Density matrix as superposition of atomic (ionic) densities	343	I
PBAN	Band(s) projected density matrix (preliminary **NEWK**)	343	I
PGEOMW	Density matrix from geometrical weights (preliminary **NEWK**)	344	I
PDIDE	Energy range projected density matrix (preliminary **NEWK**)	344	I
PSCF	Restore SCF density matrix	350	–

Properties computed from the density matrix			
ADFT	Atomic density functional correlation energy	302	I
BAND	Band structure	304	I
BIDIERD	Reciprocal form factors	305	I
CLAS	Electrostatic potential maps (point multipoles approximation)	311	I
ECHG	Charge density and charge density gradient - 2D grid	319	I
ECH3	Charge density - 3D grid	317	I
EDFT	Density functional correlation energy (HF wave function only)	320	I
EMDLDM	Electron momentum distribution (along a line)	321	I
EMDPDM	Electron momentum distribution (in a plane)	322	I
HIRSHCHG	Hirshfeld population analysis	111	I
KINETEMD	Kinetic tensor from electron momentum density	326	I
PMP2	MP2 correction to the Valence Density Matrix	344	
POLI	Atom and shell multipoles evaluation	345	I
POTM	Electrostatic potential - 2D grid	348	I
POT3	Electrostatic potential - 3D grid	346	I
POTC	Electrostatic properties	347	I
PPAN	Mulliken population analysis	117	
XFAC	X-ray structure factors	352	I

Properties computed from the density matrix (spin-polarized systems)

ANISOTRO	Hyperfine electron-nuclear spin tensor	303	I
HIRSHCHG	Hirshfeld spin population analysis	111	I
ISOTROPIC	Hyperfine electron-nuclear spin interaction - Fermi contact	325	I
POLSPIN	Atomic spin density multipoles	346	I

Properties computed from eigenvectors (after keyword **NEWK**)

ANBD	Printing of principal AO component of selected CO	302	I
BWIDTH	Printing of bandwidth	311	I
DOSS	Density of states	315	I
EMDL	Electron momentum distribution - line	321	I
EMDP	Electron momentum distribution - plane maps	324	I
PROF	Compton profiles and related quantities	349	I
BOLTZTRA	Transport Properties (electron conductivity, Seebeck) within the semiclassical Boltzmann theory	309	I
SPINCNTM	Spin Contamination	351	

New properties

SPOLBP	Spontaneous polarization (Berry phase approach)	357	–
SPOLWF	Spontaneous polarization (localized CO approach)	358	–
LOCALWF	Localization of Wannier functions	327	I
DIEL	Optical dielectric constant	313	I
ISO+POTC	Mössbauer isomer shift and quadrupolar effects	358	I
TOPO	Topological analysis of the electron density	361	I

Auxiliary and control keywords

ANGSTROM	Set input unit of measure to Ångstrom	35	–
BASISSET	Printing of basis set, Fock/KS, overlap and density matrices	305	–
BOHR	Set input unit of measure to bohr	38	–
CHARGED	Non-neutral cell allowed (**PATO**)	75	–
END	Terminate processing of properties input keywords		–
FRACTION	Set input unit of measure to fractional	47	–
MAPNET	Generation of coordinates of grid points on a plane	339	I
MAXNEIGHB	maximum number of equidistant neighbours from an atom	49	I
NEIGHBOR	Number of neighbours to analyse in **PPAN**	59	I
PRINTOUT	Setting of printing options	64	I
RAYCOV	Modification of atomic covalent radii	64	I
SETINF	Setting of inf array options	66	I
SETPRINT	Setting of printing options	66	I
STOP	Execution stops immediately	68	–
SYMMOPS	Printing of point symmetry operators	73	–

Info - Output of data on external units

ATOMIRR	Coordinates of the irreducible atoms in the cell	303	–
ATOMSYMM	Printing of point symmetry at the atomic positions	38	–
COORPRT	Coordinates of all the atoms in the cell	42	–
CRYAPI_OUT	geometry, BS, direct lattice information	312	–
CRYAPI_OUT	Reciprocal lattice information + eigenvalues	312	–
EXTPRT	Explicit structural/symmetry information	44	–
FMWF	Wave function formatted outputi in file fort.98. Section 14.10	324	–
INFOGUI	Generation of file with wf information for visualization	325	–
XML	generation of XML file for electron transport with WanT	**??**	–

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] [17]
	or
SAVIN	Savin weights [251]

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 64)
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 2.2

ANGSTROM - unit of measure

Unit of measure of coordinates (**ECHG**, **POTM**, **CLAS**) See input block 1, page 35.

ANISOTRO - anisotropic spin-dipole 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 [289]

$$\mathbf{T}[\mathrm{mT}] = \frac{1000}{(0.529177 \cdot 10^{-10})^3} \frac{1}{4\pi} \mu_0 \beta_\mathrm{N} \mathrm{g_N} \mathbf{T} = 3.4066697 \mathrm{g_N} \mathbf{T}$$

(see **ISOTROPIC** for the units and conversion factors). The elements of the **T** tensor at nucleus A are defined as follows:

$$\mathrm{T}^\mathrm{A}_{ij} = \frac{1}{n_\alpha - n_\beta} \sum_{\mu\nu} \sum_\mathrm{g} \mathrm{P}^\mathrm{spin}_{\mu\nu\mathrm{g}} \int \varphi_\mu(\boldsymbol{r}) \left(\frac{r^2_\mathrm{A} \delta_{ij} - 3 r_{\mathrm{A}i} r_{\mathrm{A}j}}{r^5_\mathrm{A}} \right) \varphi^\mathrm{g}_\nu(\boldsymbol{r}) \mathrm{d}\boldsymbol{r}$$

where $\frac{1}{n_\alpha - n_\beta}$ is the number of unpaired electrons, $\boldsymbol{r}_\mathrm{A} = |\boldsymbol{r} - \mathrm{A}|$ and $r_{\mathrm{A}i} = (\boldsymbol{r} - \mathrm{A})_i$ (i$^\mathrm{th}$ 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.

ATOMIRR - coordinates of irreducible atoms

Cartesian and fractional coordinates of the irreducible atoms are printed. No input data required.

ATOMSYMM

See input block 1, page 38

14.3 Electronic Band Structure

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 D, page 442
	LPR66	$\neq 0$	printing of eigenvalues
			if ISS > 0 then
			add NLINE records
• ∗	I1,I2,I3		integers that define the starting point of the line (I1/ISS $\mathbf{b}_1$+I2/ISS $\mathbf{b}_2$+I3/ISS $\mathbf{b}_3$), with $\mathbf{b}_1$,$\mathbf{b}_2$,$\mathbf{b}_3$ reciprocal lattice vectors.
	J1,J2,J3		integers that define the final point of the line (J1/ISS $\mathbf{b}_1$+J2/ISS $\mathbf{b}_2$+J3/ISS $\mathbf{b}_3$) with $\mathbf{b}_1$,$\mathbf{b}_2$,$\mathbf{b}_3$ reciprocal lattice vectors.
			if ISS $= 0$ then
			add NLINE records
• ∗	LABELA		label of the the starting point of the line (see tables 14.1 and 14.2 below for a legend).
	LABELB		label of the the final point of the line (see tables 14.1 and 14.2 below for a legend).

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 D, page 442).
When all the starting and terminal points are chosen to coincide with special (high symmetry) ones, it is possible to write the conventional label of each point instead of its coordinates (this option is activated by putting ISS=0). These labels have to be expressed as letters in the latin alphabet: the Γ point is identified by letter G. Apart from Γ, the labels of the special points are different for each Bravais lattice: the convention adopted for the special points and their position in the Brillouin zone can be found in tables 14.1 and 14.2 below. For instance, in the MgO case (fcc lattice), two equivalent inputs would read:

```
BAND
MGO
2 0 30 1 18 1 0
G X
X W
```

and:

```
BAND
MGO
2 12 30 1 18 1 0
0 0 0  6 0 6
6 0 6  6 3 9
```

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 [224] and in Chapter 2 of reference [220], 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 Monkhorst net. Eigenvectors computed by NEWK in k points corresponding to the Monkhorst net are not readable any more. To compute density of states and bands, the sequence must be: BAND - NEWK - DOSS.

6. The ISS=0 option does not recognize the labels of every special points; the ones recognized are only those reported in tables 14.1 and 14.2.

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 $\neq 0$ *insert* **prtrec** *(see page 66)* 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 E.
Printing options:
59 (Density matrix); 60 (Overlap matrix); 64 (Fock/KS matrix).

14.4 Compton Profiles - from $B(r)$ Function

BIDIERD - Reciprocal form factors

This option evaluates the reciprocal form factors (RFF) (also called auto-correlation function) for any direction directly from the direct space density matrix.
Compton Profiles (CPs) can be computed by Fourier transforming the RFF with the **PROF** sub-keyword below. The starting auto-correlation function must be of good quality in order to get good CPs. Both RFFs and CPs can be convoluted (see **CONV** sub-keyword) in order to be compared with the experiments (affected by the finite resolution of the spectrometer); this procedure is performed by multiplying the RFF by the gaussian function $g(r)_{\sigma_r}$:

$$g(r)_{\sigma_r} = e^{-\frac{r^2}{2\sigma_r^2}} \qquad \text{where} \quad \sigma_r = \frac{1}{\sigma_p} = \frac{2\sqrt{2log2}}{FWHM_p}$$

where $FWHM_p$, the convolution parameter, has to be defined in input by the user (in atomic units); the r and p subscripts identify quantities in coordinates and momentum space, respectively. The anisotropies of the RFFs and the CPs can be evaluated via the **DIFF** sub-keyword. This block must be ended by **ENDB**. It works also for open-shell systems.

Table 14.1: Labels and fractional coordinates (referred to reciprocal space lattice parameters of the primitive cell) of the special points recognized in input for each Bravais lattice. FC = face centered, BC = body centered, AC = base centered.

Lattice	Point	Coordinates
P Cubic	M	$\frac{1}{2}\ \frac{1}{2}\ 0$
	R	$\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}$
	X	$0\ \frac{1}{2}\ 0$
FC Cubic	X	$\frac{1}{2}\ 0\ \frac{1}{2}$
	L	$\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}$
	W	$\frac{1}{2}\ \frac{1}{4}\ \frac{3}{4}$
BC Cubic	H	$\frac{1}{2}\ \frac{-1}{2}\ \frac{1}{2}$
	P	$\frac{1}{4}\ \frac{1}{4}\ \frac{1}{4}$
	N	$0\ 0\ \frac{1}{2}$
Hexagonal *or* P Trigonal	M	$\frac{1}{2}\ 0\ 0$
	K	$\frac{1}{3}\ \frac{1}{3}\ 0$
	A	$0\ 0\ \frac{1}{2}$
	L	$\frac{1}{2}\ 0\ \frac{1}{2}$
	H	$\frac{1}{3}\ \frac{1}{3}\ \frac{1}{2}$
Rhombohedral (R Trigonal)	T	$\frac{1}{2}\ \frac{1}{2}\ \frac{-1}{2}$
	F	$0\ \frac{1}{2}\ \frac{1}{2}$
	L	$0\ 0\ \frac{1}{2}$
P Monoclinic	A	$\frac{1}{2}\ \frac{-1}{2}\ 0$
	B	$\frac{1}{2}\ 0\ 0$
	C	$0\ \frac{1}{2}\ \frac{1}{2}$
	D	$\frac{1}{2}\ 0\ \frac{1}{2}$
	E	$\frac{1}{2}\ \frac{-1}{2}\ \frac{1}{2}$
	Y	$0\ \frac{1}{2}\ 0$
	Z	$0\ 0\ \frac{1}{2}$
AC Monoclinic	A	$\frac{1}{2}\ 0\ 0$
	Y	$0\ \frac{1}{2}\ \frac{1}{2}$
	M	$\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}$

Table 14.2: Labels and fractional coordinates (referred to reciprocal space lattice parameters of the primitive cell) of the special points recognized in input for each Bravais lattice. FC = face centered, BC = body centered, AC = base centered.

Lattice	Point	Coordinates
P Orthorombic	S	$\frac{1}{2}\ \frac{1}{2}\ 0$
	T	$0\ \frac{1}{2}\ \frac{1}{2}$
	U	$\frac{1}{2}\ 0\ \frac{1}{2}$
	R	$\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}$
	X	$\frac{1}{2}\ 0\ 0$
	Y	$0\ \frac{1}{2}\ 0$
	Z	$0\ 0\ \frac{1}{2}$
FC Orthorombic	Z	$\frac{1}{2}\ \frac{1}{2}\ 0$
	Y	$\frac{1}{2}\ 0\ \frac{1}{2}$
	T	$1\ \frac{1}{2}\ \frac{1}{2}$
AC Orthorombic	S	$0\ \frac{1}{2}\ 0$
	T	$\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}$
	R	$0\ \frac{1}{2}\ \frac{1}{2}$
	Y	$\frac{1}{2}\ \frac{1}{2}\ 0$
	Z	$0\ 0\ \frac{1}{2}$
BC Orthorombic	S	$\frac{1}{2}\ 0\ 0$
	T	$0\ 0\ \frac{1}{2}$
	R	$0\ \frac{1}{2}\ 0$
	X	$\frac{1}{2}\ \frac{-1}{2}\ \frac{1}{2}$
	W	$\frac{1}{4}\ \frac{1}{4}\ \frac{1}{4}$
P Tetragonal	M	$\frac{1}{2}\ \frac{1}{2}\ 0$
	R	$0\ \frac{1}{2}\ \frac{1}{2}$
	A	$\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}$
	X	$0\ \frac{1}{2}\ 0$
	Z	$0\ 0\ \frac{1}{2}$
BC Tetragonal	M	$\frac{1}{2}\ \frac{1}{2}\ \frac{-1}{2}$
	P	$\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}$
	X	$0\ 0\ \frac{1}{2}$

By using the alternative **PROF** keyword (see page 349), followed by the **BR** sub-keyword, it is possible to obtain the RFF by Fourier transforming the Compton profiles. As the latter implies numerical integration, the **BIDIERD** keyword is expected to provide more accurate results.
Both auto-correlation functions and Compton profiles are saved in two-column format in the external unit CP.DAT

Users of this option for the calculation of CPs are kindly requested to cite the following paper[96]:

A. Erba, C. Pisani, S. Casassa, L. Maschio, M. Schütz and D. Usvyat, *Phys. Rev. B.*, **81**,165108 (2010). *A MP2 versus DFT theoretical investigation of the Compton profiles of crystalline urea.*

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
• A4	**CONV**		Convolution of the $B(r)$ previously computed
• ∗	$FWHM_p$		convolution parameter (a.u.)
• A4	**PROF**		Compton Profiles computed as Fourier Transforms of the $B(r)$
• ∗	NPOIP		number of points along each direction
	STEPC		step along each direction (a.u.)
• A4	**DIFF**		$B(r)$ and CPs anisotropies are computed
• A3	**DIR**		The directions are specified
			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		label of the atom and indices of the cell where the atom is located. The
	ZG		explored direction is defined by the straight line going from the origin to the atom position
• A4	**END**		End block autocorrelation functions

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.

14.5 Transport Properties

BOLTZTRA - Boltzmann Transport Properties

mandatory keywords II

rec	variable	meaning
• *A*	**TRANGE**	Temperature range for transport properties calculation
∗	T1,T2,TS	3 reals, temperature range (K) from T1 to T2, TS is the temperature step
• *A*	**MURANGE**	Chemical potential range for transport properties calculation
∗	E1,E2,ES	3 reals, chemical potential range (eV) from E1 to E2, ES is the chemical potential step
• *A*	**TDFRANGE**	Energy range of distribution function for transport properties calculation
∗	EDF1,EDF2,EDFS	3 reals, energy range of distribution function (eV) from EDF1 to EDF2, EDFS is the energy step

optional keywords II

rec	variable	value	meaning
• *A*	**RELAXTIM**		τ constant within the relation time approximation for transport properties calculation
∗	T		1 real, T is the electron lifetime (fs), τ (default 10.fs)
• *A*	**SMEAR**		smearing of distribution function for transport properties calculation
∗	S		1 real, smearing coefficient (default 0.)
• *A*	**SMEARTYP**		smearing type for transport properties calculation
∗	ST	-2	cold smearing (Marzari-Vanderbilt)
		-1	Fermi-Dirac
		0≤ST≥10	Methfessel-Paxton (default 0)

The electrical conductivity, the Seebeck coefficient and the thermal conductivity can be estimated by calculating the three so-called transport coefficients. Their expression can be obtained solving, in the relaxation time approximation, the Boltzmann's transport equation. A detailed description of the semiclassical Boltzmann transport theory is beyond our scope. For a thorough explanation see Refs. [9, 310, 131].
Here, we limit to the main results of this theory i.e. operative equations: σ, $\sigma\mathbf{S}$ and κ_{el},

$$[\sigma]_{qr}(\mu,T) = e^2 \int \mathrm{d}E\Big(-\frac{\partial f_0}{\partial E}\Big)\Xi_{qr}(E) \tag{14.1}$$

$$[\sigma\mathbf{S}]_{qr}(\mu,T) = \frac{e}{T} \int \mathrm{d}E\Big(-\frac{\partial f_0}{\partial E}\Big)(E-\mu)\Xi_{qr}(E) \tag{14.2}$$

$$[\kappa_e]_{qr}(\mu,T) = \frac{1}{T} \int \mathrm{d}E\Big(-\frac{\partial f_0}{\partial E}\Big)(E-\mu)^2\Xi_{qr}(E) \tag{14.3}$$

where μ is the chemical potential or the Fermi level, E is the energy, f_0 is the Fermi-Dirac distribution and Ξ is the transport distribution function (TDF). In the Equations above, for the sake of simplicity we define core of the transport coefficients, Ξ, as

$$\Xi_{qr}(E) = \tau \sum_{\mathbf{k}} \frac{1}{N_{\mathbf{k}}} \frac{1}{V} \sum_{i,j} v_{i,q}(\mathbf{k}) v_{j,r}(\mathbf{k}) \delta\big(E - E_i(\mathbf{k})\big) \tag{14.4}$$

where $v_{i,q}(\mathbf{k})$ is the velocity of the i^{th} band calculated along the cartesian direction q, τ is the lifetime which is assumed to be constant according to the relax time approximation. The δ function is approximated by a suitable smearing function, that can be set using the **SMEARTYP** keyword.
From the theoretical point of view, the band velocity is the critical quantity that has to be

calculated. In atomic unit (a.u.) its expression is the derivative of the band energies $E(i, \mathbf{k})$, with respect to a reciprocal space vector k_q

$$v_{i,q}(\mathbf{k}) = \frac{\partial E_i(\mathbf{k})}{\partial k_q} \quad . \tag{14.5}$$

At the end of the calculation, the code saves five formatted DAT files: SIGMA.DAT which contains Eq. 14.1, SEEBECK.DAT where the Seebeck coefficient is stored, SIGMAS.DAT for Eq. 14.2, KAPPA.DAT for Eq. 14.3 and TDF.DAT for Eq. 14.4.
The structure of the files for a closed shell calculation (CS) is common for all the different kinds and it is as follow:

```
1 # 0 [Electrical conductivity in SI units, i.e. in 1/Ohm/m]
2 # Mu(eV) T(K) N(#carriers) sigma_xx sigma_xy sigma_yy
3 #T(K) = 300.000 Npoints =    1001 V(cm^3) =  7.676E-23
      ...      ...         ...            ...             ...              ...
1005 #T(K) = 500.000 Npoints =    1001 V(cm^3) =  7.676E-23
      ...      ...         ...            ...             ...              ...
```

The first line of the header specifies the kind of calculation, 0 or 1 for CS and OS respectively, and the transport phenomenon. In line 2 we explain the meaning of the numbers of each column, from left they stand for chemical potential,temperature, number of carriers and tensor components. In line 3, further information are provided for plotting with the CRYSPLOT package (see boh). The structure for an open shell calculation (OS) is:

```
1 # 1 [Alpha Electrical conductivity in SI units, i.e. in 1/Ohm/m]
2 # Mu(eV) T(K) N(#carriers) sigma_xx sigma_xy sigma_yy
3 #T(K) = 300.000 Npoints =    1001 V(cm^3) =  7.676E-23
    ...      ...         ...            ...             ...              ...              ...              ...
3 #T(K) = 500.000 Npoints =    1001 V(cm^3) =  7.676E-23
    ...      ...         ...            ...             ...              ...              ...              ...
1 # 1 [Beta Electrical conductivity in SI units, i.e. in 1/Ohm/m]
2 # Mu(eV) T(K) N(#carriers) sigma_xx sigma_xy sigma_yy
3 #T(K) = 300.000 Npoints =    1001 V(cm^3) =  7.676E-23
    ...      ...         ...            ...             ...              ...              ...              ...
3 #T(K) = 500.000 Npoints =    1001 V(cm^3) =  7.676E-23
    ...      ...         ...            ...             ...              ...              ...              ...
```

For these calculations, in the first line of the header wil appear 1, see above, and the electron spin specification, alpha or beta.
Notes:

- the Seebeck coefficient, SEEBECK.DAT file, is not spin dependent;
- the SIGMAS.DAT, in OS calculations, will have an alpha and beta headers;
- the numbers of carriers (N) has been computed by means of:

$$N_{\mu,T}(E) = \frac{n}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_{i} f_0\Big(E, \mu, T\Big)\Theta\Big(E - E_i(\mathbf{k})\Big) = \frac{n}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_{i} \frac{1}{exp\Big(\frac{E-\mu}{k_B T}\Big) + 1}\Theta\Big(E - E_i(\mathbf{k})\Big) \tag{14.6}$$

  where f_0 is the Fermi-Dirac distribution, $\Theta\Big(E - E_i(\mathbf{k})\Big)$ is the theta function, n is the number of electron per state and $N_{\mathbf{k}}$ is the number of k-points in the IBZ.

BOHR - unit of measure

Unit of measure of coordinates (**ECHG**, **POTM**, **CLAS**) See input block 1, page 38.

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 75.
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 339)

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 339). The potential values are written formatted in file fort.25. (see Appendix D, page 441).

2. When IDER $\neq$ 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 42.

COMMENS - Density Matrix commensurate to Monkhorst grid

rec		variable	value	meaning
•	∗	ICASO	0	total density matrix
			1	core density matrix
			2	valence density matrix

The **COMMENS** keyword has to be inserted as first record in the PROPERTIES input before the **NEWK** keyword. It activates the construction of the Density Matrix (DM) $\mathbf{P} \equiv P^{\mathbf{g}}_{\mu\nu}$ (where μ and ν label two atomic orbitals $\chi_\mu(\mathbf{r})$ and $\chi_\nu(\mathbf{r})$ and $\mathbf{g}$ represents a lattice vector) with the index $\mathbf{g}$ running over a number of cells L commensurate to the number L of **k**-points of the Monkhorst-Pack grid where the one-electron Hamiltonian is diagonalized and the crystalline orbitals computed.
If this option is not activated, the truncation over $\mathbf{g}$ in the definition of the DM is imposed by the first tolerance `TOL1` of the integrals: those elements $P^{\mathbf{g}}_{\mu\nu}$ are disregarded that correspond to product distributions $\Pi(\mathbf{r}) = \chi_{\mu\mathbf{0}}(\mathbf{r})\chi_{\nu\mathbf{g}}(\mathbf{r})$ for which the pseudo-overlap between the two AOs is less than 10^{-TOL1}.
The DM obtained in this way can be used for computing electron momentum densities, auto-correlation functions and Compton profiles (see keywords **BIDIERD**, **EMDLDM**, **EMDPDM**). See the discussion in section 18.8 (page 401) for further details.

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 327) are appended to file GRED.DAT

k points coordinates (Monkhorst sampling net) and eigenvectors (if computed by **NEWK** page 341) 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 D, page 450.

DIEL/DIELECT - Optical dielectric constant

Computes the dielectric constant of the system along the periodic direction where an external electric field has been applied during the SCF calculation (using keyword **FIELD**, page 44). The dielectric constant is calculated by using the concept of macroscopic average of the total charge density (see for example Fu *et al.* [113]) and Poisson's equation. The charge density is first averaged with respect to the (infinite) plane orthogonal to the field

$$\overline{\rho}(z) = \frac{1}{A} \int_A \rho(z)\, dA \tag{14.7}$$

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:

$$\overline{\rho}(z) = \frac{1}{V} \sum_{\ell=-\infty}^{+\infty} F_{00\ell}\, e^{-\imath \frac{2\pi \ell z}{C}} \tag{14.8}$$

$F_{00\ell}$ 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

$$\overline{\overline{\rho}}(z) = \frac{1}{\Delta z} \int_{z-\Delta z/2}^{z+\Delta z/2} \overline{\rho}(z')\, dz' \tag{14.9}$$

that is

$$\overline{\overline{\rho}}(z) = \frac{1}{V} \sum_{\ell=-\infty}^{+\infty} F_{00\ell}\, sinc\left(\ell\pi \frac{\Delta z}{C}\right) e^{-\imath \frac{2\pi \ell z}{C}} \tag{14.10}$$

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 $\overline{\overline{\rho}}(z)$:

$$\frac{\partial^2 \overline{\overline{V}}(z)}{\partial z^2} = -4\pi \overline{\overline{\rho}}(z) \tag{14.11}$$

or

$$\frac{\partial \overline{\overline{E}}(z)}{\partial z} = 4\pi \overline{\overline{\rho}}(z) \tag{14.12}$$

because

$$\frac{\partial \overline{\overline{V}}(z)}{\partial z} = -\overline{\overline{E}}(z) \tag{14.13}$$

$\overline{\overline{V}}(z)$, $\overline{\overline{F}}(z)$ 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:

$$F_{00\ell} = F_{00\ell}^{\Re} + \imath F_{00\ell}^{\Im} \tag{14.14}$$

Exploiting the following properties of the structure factors:

$$\begin{aligned} F_{000} &= N_e \quad (\textbf{number of electrons in the supercell}) \\ F^{\Re}_{00\ell} &= F^{\Re}_{00-\ell} \\ F^{\Im}_{00\ell} &= -F^{\Im}_{00-\ell} \end{aligned} \tag{14.15}$$

the real and imaginary parts of $\overline{\overline{\rho}}$ take the following form:

$$\Re\left[\overline{\overline{\rho}}(z)\right] = \frac{N_e}{V} + \frac{2}{V}\sum_{\ell=1}^{+\infty}\left[F^{\Re}_{00\ell}\cos\left(\frac{2\pi\ell z}{C}\right) + F^{\Im}_{00\ell}\sin\left(\frac{2\pi\ell z}{C}\right)\right] sinc\left(\ell\pi\frac{\Delta z}{C}\right) \tag{14.16}$$

$$\Im\left[\overline{\overline{\rho}}(z)\right] = 0 \tag{14.17}$$

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 14.13, 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. 14.12):

$$\overline{\overline{E}}(z) = \frac{8\pi}{V}\sum_{\ell=1}^{+\infty}\left[F^{\Re}_{00\ell}\frac{\sin\left(\frac{2\pi\ell z}{C}\right)}{\left(\frac{2\pi\ell}{C}\right)} - F^{\Im}_{00\ell}\frac{\cos\left(\frac{2\pi\ell z}{C}\right)}{\left(\frac{2\pi\ell}{C}\right)}\right] sinc\left(\ell\pi\frac{\Delta z}{C}\right) \tag{14.18}$$

The corresponding macroscopic electric potential can be written as follows:

$$\overline{\overline{V}}(z) = \frac{-8\pi}{V}\sum_{\ell=1}^{+\infty}\left[F^{\Re}_{00\ell}\frac{\cos\left(\frac{2\pi\ell z}{C}\right)}{\left(\frac{2\pi\ell}{C}\right)^2} + F^{\Im}_{00\ell}\frac{\sin\left(\frac{2\pi\ell z}{C}\right)}{\left(\frac{2\pi\ell}{C}\right)^2}\right] sinc\left(\ell\pi\frac{\Delta z}{C}\right) \tag{14.19}$$

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:

$$\epsilon = \frac{E_0}{E_0 + \overline{\overline{E}}} \tag{14.20}$$

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 D, page 442.
Available keywords are:

rec	variable	meaning
•	*A* **END**	end of **DIEL** input block
		optional keywords II
•	*A* **PRINT**	extended output

14.6 Electronic Density-of-State

DOSS - Density of states

rec		variable	value	meaning
•	∗	NPRO	0	only total DOS is calculated
			> 0	total DOS and NPRO projected densities are calculated.
		NPT		number of uniformly spaced energy values 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 to file fort.25 for plotting (Appendix D, page 442).
			2	formatted output to file DOSS.DAT for plotting (Appendix D, page 443).
		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 ≠ 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 ≠ 0, insert **prtrec** *(see page 66)* 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:

$$\rho_\mu(\epsilon) = 2/V_B \sum_j \sum_\nu \int_{BZ} d\underline{k} S_{\mu\nu}(\underline{k}) a_{\mu j}(\underline{k}) a^*_{\nu j}(\underline{k}) \;\; \delta[\epsilon - \epsilon_j(\underline{k})] \tag{14.21}$$

$$\rho_A(\epsilon) = \sum_{\mu \in A} \rho_\mu(\epsilon) \tag{14.22}$$

$$\rho_{tot}(\epsilon) = \sum_A \rho_A(\epsilon) \tag{14.23}$$

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 18.7, page 401). 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 **DOSS**, 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, ([221, 222]). 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 second 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 124) or level shifter (**LEVSHIFT**, page 113) 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.

14.7 Crystal Orbital Overlap/Hamiltonian Populations

Similar to the density of states calculation, a crystal orbital overlap/Hamiltonian population (COOP/COHP) analysis allows for the investigation of electronic states within a material. But in contrast to the DOS, a COOP/COHP calculation is specifically aimed at the interaction between two selected groups of orbitals/atoms. Thus, information regarding bonding and antibonding states, bond orders, and interaction strengths can be obtained in a straightforward fashion.
The difference between the COOP and COHP analyses lie in the usage of either the overlap matrix (**S**) or the Hamiltonian matrix (**H**) elements for COOP and COHP, respectively. The former provides a measure of the bond order, while the latter a measure of bond strength. The mathematical form is very similar to that of a projected DOS equation:

$$COOP_{A-B}(\epsilon) = \frac{2}{V_B} \sum_j \sum_{\mu \in A} \sum_{\nu \in B} \int_{BZ} S_{\mu\nu}(\mathbf{k}) a^*_{\mu j}(\mathbf{k}) a_{\nu j}(\mathbf{k}) \delta(\epsilon - \epsilon_j(\mathbf{k})) \mathrm{d}\mathbf{k}$$

$$COHP_{A-B}(\epsilon) = \frac{2}{V_B} \sum_j \sum_{\mu \in A} \sum_{\nu \in B} \int_{BZ} H_{\mu\nu}(\mathbf{k}) a^*_{\mu j}(\mathbf{k}) a_{\nu j}(\mathbf{k}) \delta(\epsilon - \epsilon_j(\mathbf{k})) \mathrm{d}\mathbf{k}$$

Two keywords are available to compute these quantities: **COOP** and **COHP**.
As with the calculation of the density of states, the COOP/COHP calculations require the Fock/KS eigenvectors at each **k** point defined by Pack-Monkhorst net, which need to be recalculated using the keyword **NEWK**.
The INPUT format for the crystal orbital overlap population (**COOP**) and the crystal orbital Hamiltonian population (**COHP**) calculations are essentially identical, and both can be run simultaneously.

Users of this module are kindly asked to cite the following reference:

M.T. Ruggiero, A. Erba, R. Orlando and T. M. Korter, *Phys. Chem. Chem. Phys.*, **17**, 31023-31029 (2015)

rec	variable	value	meaning
• ∗	NINT		number of interactions A-B
	NPT		number of uniformly spaced energy values where COOP/COHP are calculated, from bottom of band INZB to top of band IFNB.
	INZB		first band considered in COOP/COHP calculation
	IFNB		last band considered in COOP/COHP calculation
	IPLO	0	COOP/COHP are not stored on disk
		1	formatted output to file fort.25 for plotting (Appendix D, page 442).
		2	formatted output to file COOP.DAT or COHP.DAT for plotting (Appendix D, page 443).
	NPOL		number of Legendre polynomials used to expand COOP/COHP (≤ 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
			insert NINT records II
• ∗	NA	> 0	Number of AOs of center A (to be listed in the line below)
		< 0	All AOs of NA atoms (to be listed in the line below) included in center A
	NDM(J),J=1,NA		List of AOs (if NA>0) or atoms (if NA<0) by their sequence number (basis set order)
• ∗	NB	> 0	Number of AOs of center B (to be listed in the line below)
		< 0	All AOs of NB atoms (to be listed in the line below) included in center B
	NDM(J),J=1,NB		List of AOs (if NB>0) or atoms (if NB<0) by their sequence number (basis set order)
			if NPR $\neq$ 0, insert **prtrec** *(see page 66)* II

For the calculation of COOP/COHP the following information must be specified:
The number of interactions to explore
The number of points along the energy range at which the COOP/COHP is calculated
The range of energy to explore. It can be defined by two bands (the range is from the bottom of the first band to the top of the second band), or by two energies.
The way in which the COOP/COHP data is written
The degree of the polynomials to be used in the COOP/COHP expansion
Any printing options (see the manual)
Followed by the selection of the orbitals or atoms that are to be included in the two groups
Selection of either orbital (positive) or atomic (negative) indices and the number of orbitals/atoms to include in the group
The label of the orbital(s) or atom(s)

14.8 3D Electron Charge Density

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	**RANGE**
length scales for non-periodic dimensions	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 [and 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 to file fort.31 (function value at the grid points) in the format required by the visualization program DLV.
See Appendix D, page 447, for description of the format.
Function data computed at 3D grid points are written according to GAUSSIAN CUBE format in files (for calculations from a 1c-SCF):

DENS_CUBE.DAT	charge density
SPIN_CUBE.DAT	spin density

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 in the cell.
For calculations from a 2c-SCF, the formats are:

DENS_CUBE.DAT	electron density ρ
MX_CUBE.DAT	m_x component of the magnetization $\boldsymbol{m}$
MY_CUBE.DAT	m_y component of the magnetization $\boldsymbol{m}$
MZ_CUBE.DAT	m_z component of the magnetization $\boldsymbol{m}$
JX_CUBE.DAT	j_x component of the particle current $\boldsymbol{j}$
JY_CUBE.DAT	j_y component of the particle current $\boldsymbol{j}$
JZ_CUBE.DAT	j_z component of the particle current $\boldsymbol{j}$
JSXX_CUBE.DAT	J_x^x component of the spin current $\boldsymbol{J}^x$
JSXY_CUBE.DAT	J_y^x component of the spin current $\boldsymbol{J}^x$
JSXZ_CUBE.DAT	J_z^x component of the spin current $\boldsymbol{J}^x$
JSYX_CUBE.DAT	J_x^y component of the spin current $\boldsymbol{J}^y$
JSYY_CUBE.DAT	J_y^y component of the spin current $\boldsymbol{J}^y$
JSYZ_CUBE.DAT	J_z^y component of the spin current $\boldsymbol{J}^y$
JSZX_CUBE.DAT	J_x^z component of the spin current $\boldsymbol{J}^z$
JSZY_CUBE.DAT	J_y^z component of the spin current $\boldsymbol{J}^z$
JSZZ_CUBE.DAT	J_z^z component of the spin current $\boldsymbol{J}^z$

In the case of a calculation from a 2c-SCF wavefunction, it is necessary to specify which variables (i.e. either ρ, $\boldsymbol{m}$, $\boldsymbol{j}$, and/or the three $\boldsymbol{J}^x$, $\boldsymbol{J}^y$, $\boldsymbol{J}^z$) are to be calculated using a prior call to the keyword **PROPS2COMP**, see p. 179.

14.9 2D Electron Charge Density

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 14.11, page 339)	

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 339). The electron density values (electron bohr^{-3}) are written formatted in file fort.25 (see Appendix D, page 441).

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 D, page 441)

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 D, page 441).

5. The electron charge density is computed from the density matrix stored in file fort.31. 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**).

For a calculation from a 2c-SCF wavefunction the density variables to be calculated must be specified by a prior call to the keyword **PROPS2COMP**, see p. 179. The various density variables are written to fort.25 (see Appendix D, page 441) in the following order, with each record being separated by a string with the header "MAPN":

1. electron density ρ
2. m_x component of the magnetization $\boldsymbol{m}$
3. m_y component of the magnetization $\boldsymbol{m}$
4. m_z component of the magnetization $\boldsymbol{m}$
5. j_x component of the particle current $\boldsymbol{j}$
6. j_y component of the particle current $\boldsymbol{j}$
7. j_z component of the particle current $\boldsymbol{j}$
8. J_x^x component of the spin current $\boldsymbol{J}^x$

9. J_y^x component of the spin current $\boldsymbol{J}^x$
10. J_z^x component of the spin current $\boldsymbol{J}^x$
11. J_x^y component of the spin current $\boldsymbol{J}^y$
12. J_y^y component of the spin current $\boldsymbol{J}^y$
13. J_z^y component of the spin current $\boldsymbol{J}^y$
14. J_x^z component of the spin current $\boldsymbol{J}^z$
15. J_y^z component of the spin current $\boldsymbol{J}^z$
16. J_z^z component of the spin current $\boldsymbol{J}^z$

Values of IDER > 0 are not supported for calculations including magnetization or current densities. For subsequent reading and processing of data involving magnetization and current densities, it may be useful to also insert the keyword **ADDSPACE**, see p. 179.

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] [17]
or	
SAVIN	Savin weights [251]

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 64)

TOLLDENS		
• ∗	ID	DFT density tolerance [default 9]

TOLLGRID		
• ∗	IG	DFT grid weight tolerance [default 18]

14.10 Electron Momentum Density

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 to file fort.25 for plotting (Appendix D, page 443).
			2	formatted output to file LINEA.DAT for plotting (Appendix D, page443).
		LPR113	$\neq 0$	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 $\neq 0$ *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 $\neq 0$ *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 The calculation of the Fermi energy is necessary for metallic systems (**NEWK** keyword with IFE = 1). The Electron Momentum Density is calculated along given directions (equation 18.22, page 402). It can be computed also for open-shell systems. 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.

EMDLDM - Electron Momentum Density from Density Matrix - line path

Users of this option are kindly requested to cite the following papers[95, 225]:

A. Erba and C. Pisani, *J. Comput. Chem.*, **33**, 822 (2012). *Evaluation of the electron momentum density of crystalline systems from ab initio linear combination of atomic orbitals calculations.*

C. Pisani, A. Erba, S. Casassa, M. Itou and Y. Sakurai, *Phys. Rev. B*, **84**, 245102 (2011). *The anisotropy of the electron momentum distribution in α-quartz investigated by Compton scattering and ab initio simulations.*

rec		variable	value	meaning
•	∗	N		number of directions (≤ 10)
		PMAX		maximum momentum value (a.u.) for which the EMD is computed
		STEP		discretization step for the EMD
		IPLO	0	no data stored on disk
			1	formatted output to file fort.25 for plotting (Appendix D, page 443).
			2	formatted output to file EMDLDM.DAT for plotting (Appendix D, page 443).
				insert N records
•	∗	H K L		three integers defining the direction with respect to the conventional cell
				end
•	∗	ICASO	1	the total density matrix is used
			2	the core density matrix is used
			3	the valence density matrix is used
•	∗	NSA1	0	no spherically averaged EMD
			1	the spherically averaged EMD is computed
		NSA2	0	no EMD-anisotropy
			1	EMD-anisotropies are computed

The Electron Momentum Density (EMD) is calculated along given crystallographic directions (defined in oblique coordinates with respect to the conventional cell) directly from the Density Matrix. The EMD is a non-periodic function; it falls rapidly to zero outside the first Brillouin zone. $\pi(\mathbf{0})$ gives the number of electrons at rest. For Open-Shell systems the $\alpha + \beta$ and the $\alpha - \beta$ EMD are computed.
The **NEWK** keyword must be called with the option activating the Fermi level calculation before **EMDLDM** if the core or valence Density Matrix is desired (`ICASO`=2,3).
The spherically averaged EMD $\pi_{SA}(|\mathbf{p}|)$ is computed according to the procedure described in section 18.8 (page 401). The EMD-anisotropy is $\Delta\pi(\mathbf{p}) = \pi(\mathbf{p}) - \pi_{SA}(|\mathbf{p}|)$.

EMDPDM - Electron Momentum Density from Density Matrix - maps

Users of this option are kindly requested to cite the following papers[95, 225]:

A. Erba and C. Pisani, *J. Comput. Chem.*, **33**, 822 (2012). *Evaluation of the electron momentum density of crystalline systems from ab initio linear combination of atomic orbitals calculations.*

C. Pisani, A. Erba, S. Casassa, M. Itou and Y. Sakurai, *Phys. Rev. B*, **84**, 245102 (2011). *The anisotropy of the electron momentum distribution in α-quartz investigated by Compton scattering and ab initio simulations.*

rec	variable	value	meaning
• ∗	N		number of planes (≤ 10)
	PMAX		maximum momentum value (a.u.) for which the EMD is computed in the two main directions defining the plane
	STEP		discretization step for the EMD
	IPLO	0	no data stored on disk (data only in output)
		1	formatted output to file fort.25 for plotting (Appendix D, page 441)
		2	formatted output to files 3DEMDTOTAL.DAT and 3DEMDANISO.DAT for 3D plotting of total EMD and EMD-anisotropy (for squared map-windows only); see below
	IDIR	0	planes defined via the **MAPNET** dummy keyword
		1	planes defined via two crystallographic directions
• ∗	ICASO	1	the total density matrix is used
		2	the core density matrix is used
		3	the valence density matrix is used
• ∗	NSA1	0	no spherically averaged EMD
		1	the spherically averaged EMD is computed
	NSA2	0	no EMD-anisotropy
		1	EMD-anisotropies are computed
	NSA3	0	no restart of spherically averaged EMD
		$\neq 0$	a previously computed spherically averaged EMD is read from input
			if NSA3 ≠ 0 insert NSA3 records
• ∗	P		value of $\|\mathbf{p}\|$
	SAEMD		value of the spherically averaged EMD $\pi(\|\mathbf{p}\|)$
			end
			insert N records
			if IDIR = 0
			insert **MAPNET** input records (Section 14.11, page 339)
			else if IDIR = 1
• ∗	H K L		three integers defining the first direction with respect to the conventional cell
	H' K' L'		three integers defining the second direction with respect to the conventional cell
			end if
			end

The Electron Momentum Density (EMD) is calculated in given crystallographic planes (defined in oblique coordinates with respect to the conventional cell) directly from the Density Matrix. The EMD is a non-periodic function; it falls rapidly to zero outside the first Brillouin zone. $\pi(\mathbf{0})$ gives the number of electrons at rest. EMD-maps can be computed for closed-shell systems only.

The **NEWK** keyword must be called with the option activating the Fermi level calculation before **EMDPDM** if the core or valence Density Matrix is desired (`ICASO`=2,3).

The spherically averaged EMD $\pi_{SA}(|\mathbf{p}|)$ is computed according to the procedure described in section 18.8 (page 401). The EMD-anisotropy is $\Delta\pi(\mathbf{p}) = \pi(\mathbf{p}) - \pi_{SA}(|\mathbf{p}|)$. If one activates the computation of the spherically averaged EMD (`NSA1`=1), the `STEP` should be small in order to reduce the numeric noise in its fitting procedure.

Formatted external units 25 (fort.25) are generated that can be read by the Crgra2006 graphics software for creating 2D maps. See Appendix D at page 441. The formatted external files 3DEMDTOTAL.DAT and 3DEMDANISO.DAT consist of a series of records PX, PY, EMD(PX,PY) and can easily be used for representing 3D surfaces of the EMD on a plane.

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 the EMD function in output
				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 The Fermi energy must be computed for metallic systems (**NEWK** keyword with IFE = 1). Calculation of electron momentum density on definite planes (equation 18.22, page 402). It works also for open-shell systems. If LPR115 ≠ 0 the EMD function is printed in output (recommended). 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 44

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.98
	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 47.

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, to file fort.31. See Appendix D, page 447, for description of the format.

HIRSHCHG/HIRSHBLK - Iterative Hirshfeld Population Analysis

See input block 3, page 111.

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 D, page 447, 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

As an additional information, the total electron density at the nuclei is computed for all systems.
In the case of open shell systems, the spin density at the nuclei ($\langle \rho^{\mathrm{spin}}(\boldsymbol{r}_{\mathrm{N}}) \rangle$) is evaluated. This

quantity is given in bohr^{-3} and is transformed into the hyperfine coupling constant a_N[mT] through the relationship [289]

$$a_N[\text{mT}] = \frac{1}{n_\alpha - n_\beta} \frac{1000}{(0.529177 \cdot 10^{-10})^3} \frac{2}{3} \mu_0 \beta_N g_N \langle \rho^{\text{spin}}(\boldsymbol{r}_N) \rangle = \frac{1}{n_\alpha - n_\beta} 14.277235\, g_N \langle \rho^{\text{spin}}(\boldsymbol{r}_N) \rangle$$

where

$$n_\alpha - n_\beta \qquad \text{is the number of upaired electrons (summed spin density)}$$

$$mu_0 = 4\pi \cdot 10^{-7} = 12.566370614 \cdot 10^{-7} [\text{T}^2\text{J}^{-1}\text{m}^3] \qquad \text{(permeability of vacuum)}$$

$$\beta_N = 5.0507866 \cdot 10^{-17} [\text{JT}^{-1}] \qquad \text{(nuclear magneton)}$$

the nuclear g_N factors for most of the nuclei of interest are available in the code and are taken from [289]. Conversion factors:

$$a_N[\text{MHz}] = \frac{a_N[\text{mT}] g_e \beta_e}{10^9 h[\text{Js}]} = 14.012470 \cdot a_N[\text{mT}]$$

$$a_N[\text{cm}^{-1}] = \frac{a_N\ [\text{MHz}] 10^8}{c[\text{ms}^{-1}]} = 0.1667820 \cdot 10^{-4} \cdot a_N[\text{MHz}]$$

$$a_N[\text{J}] = g_e\ \beta_e\ 10^{-3} a_N[\text{mT}] = 1.856954 \cdot 10^{-26} a_N[\text{mT}]$$

where:

$$\beta_e = 9.2740154 \cdot 10^{-24}\ [\text{JT}^{-1}] \qquad \text{(bohr magneton)}$$

$$g_e = 2.002319304386 \qquad \text{(free-electron g factor)}$$

$$c = 2.99792458 \cdot 10^8\ [\text{ms}^{-1}] \qquad \text{(speed of light in vacuum)}$$

$$h = 6.6260755 \cdot 10^{-34}\ [\text{Js}] \qquad \text{(Planck constant)}$$

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.

KINETEMD - Kinetic Tensor computed from the Electron Momentum Distribution

rec		variable	value	meaning
•	∗	PMAX		maximum momentum value (a.u.) for which the EMD is computed
		PINT		Maximum momentum value (a.u.) defining an inner sphere where the EMD is computed exactly (PINT < PMAX)
		STEP		discretization step for the computation of the EMD in the inner region of **p**-space ($p <$ PINT)
		STEPDIST		discretization step for the computation of the EMD in the outer region of **p**-space (PINT $< p <$ PMAX)
•	∗	ICASO	1	the total density matrix is used
			2	the core density matrix is used
			3	the valence density matrix is used

The **KINETEMD** keyword activates the computation of the kinetic tensor **T**, whose trace is the total kinetic energy of the system, as the set of second moments of the electron momentum distribution. See the discussion in section 18.8 (page 401).

KNETOUT - Reciprocal lattice information - Fock/KS eigenvalues

Obsolete. See **CRYAPI_OUT**, page 312.

14.11 Localized Crystalline Orbitals - Wannier Functions

LOCALWF - Localization of Wannier Functions (WnF)

Wannier functions are computed fromBloch functions, and then localized, following the method described in [308] and [307] The method 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, recommended to developers only.
The **LOCALWF** block is closed by the **END** keyword.
For UHF calculations two set of blocks must be inserted for the α and β electrons, each one ending with the keyword **END**.

1. The **NEWK** option must be executed before running **LOCALWF**, to compute theBloch functions.

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, an IS slightly larger 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. The IS value chosen determines a Born-von Karman supercell (or cyclic cluster) from which the program *a priori* estimates the memory space that should be enough to contain all WnF coefficients lower than the threshold 10^{-ITDP} (see the meaning of $ITDP$ in what follows) in real space. The size of this crystal domain in terms of unit cells is provided in output before the localization procedure. If the crystal domain is too small usually the localization fails and IS must be increased. On the other hand, if it is too large (very large IS) the memory space reserved for the WnF coefficients becomes overestimated and the calculation may stop because of a lack of memory for array allocation.

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 file fort.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.

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	$10^{-\mathrm{ITPH0}}$ is the initial tolerance on the atomic charge population to attribute the phase to atoms in the wannierization step [2]
	ITPH	> 0	$10^{-\mathrm{ITPH}}$ is the final tolerance used to attribute this phase [3]
	ICHTOL	> 0	DM tolerance of the cycle where ITPH0 changes to ITPH. [ITDP0+1]

Restart Keywords

With this option the WnFs of a previous run are read from file fort.81 (copy of file fort.80, written by the previous run) and projected onto the current occupied subspace. Along with the projection the WnFs are re-orthonormalized within the Born-von Karman cyclic cluster.
Tolerances and active bands must be the same as in the previous run.
Three variants are possible, corresponding to three different keywords.

RESTART: can be used to start a new localization or to improve a previous one. After the projection, cycles of wannierization and localization are performed until convergence is attained.

RESTORTH: similar to the RESTART keyword but here the WnF are orthonormalized in direct space (where the tails feature different topological properties than in the Born-von Karman cyclic cluster) immediately after their reading. Next the localization properties in direct space of the resulting WnFs *may optionally* be improved using the appropriate keywords (**WANDM** and **FULLBOYS**).

FIXWF: this keyword can be used only if file fort.80 of the previous run has been generated with the keyword **SYMMWF**. After projection, no further Boys localization step is performed. Both the original symmetry and the WnF labeling are kept with a negligible loss in the localization indices.

The last option is recommended when a sequence of CRYSCOR calculations are to be performed corresponding each time to small geometrical changes. Accordingly, to ensure a smooth evolution of the energy and wave function, all indices concerning the symmetry and the labeling of the WnFs are kept to be the same along the sequence. This is required in particular in geometry optimizations and scanning along geometrical parameters (for instance: in the calculation of molecular physisorption energy curves – see the *CRYSCOR Users's Manual* and the *CRYSCOR Tutorials* for a detailed description of the procedure). As for the **FIXINDEX** option (page 104), the reference calculation should be the one with the most compact structure.

General Keywords

MAXCYCLE

rec	variable	value	meaning
• ∗	NCYC	> 0	maximal number of short cycles for the iterative process 30

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 [219]) 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		$10^{-\mathrm{IBTOL}}$ is the threshold used for the stability condition on B_{st}. 4
	IBFRZ		If for a pair of WnFs st, $\|A_{st}\| \leq 10^{-\mathrm{IBFRZ}}$, then the corresponding WnFs are not mixed. 4
	MXBCYC		Maximum number of cycles allowed in the Boys localization process 500

EMDWF

The **EMDWF** keyword activates the computation of the contribution to the total Electron Momentum Density (EMD) $\pi(\mathbf{p})$ of a selected Wannier Function, according to the partitioning scheme illustrated in section 18.8 (page 401). Partitioning the EMD in terms of contributions coming from chemically meaningful objects like WFs (that can easily be assigned to bonds, lone pairs, etc.) is an appealing way of extracting information from a function which is still far from being completely understood.

The contribution to the EMD is computed on a given plane. The user can provide, via input, a previously computed spherically averaged EMD (see keywords **EMDLDM** and **EMDPDM**) in order to compute the contribution of a selected WF to the global anisotropy of the EMD. See the discussion of section 18.8 (page 401).

The formatted external file WFEMD.DAT consists of a series of records PX, PY, EMD(PX,PY), [SAEMD(PX,PY)] and can easily be used for representing 3D surfaces or 2D maps of the EMD on the plane.

Users of this option are kindly requested to cite the following papers[95, 225]:

A. Erba and C. Pisani, *J. Comput. Chem.*, **33**, 822 (2012). *Evaluation of the electron momentum density of crystalline systems from ab initio linear combination of atomic orbitals calculations.*

C. Pisani, A. Erba, S. Casassa, M. Itou and Y. Sakurai, *Phys. Rev. B*, **84**, 245102 (2011). *The anisotropy of the electron momentum distribution in α-quartz investigated by Compton scattering and ab initio simulations.*

rec	variable	value	meaning
• ∗	WF		label of the selected WF
	NTOL1	[1]	initial tolerance used to calculate the Dipole Moment matrix elements: $10^{-\texttt{NTOL1}}$
	IDIR	0	planes defined via the **MAPNET** dummy keyword
		1	planes defined via two crystallographic directions
	PMAX		maximum momentum value (a.u.) for which the EMD is computed in the two main directions defining the plane
	STEP		discretization step for the EMD
	IPLO	0	no data stored on disk (data only in output)
		1	formatted output to file WFEMD.DAT for 2-3D plotting of total EMD and EMD-anisotropy (for squared map-windows only)
• ∗	NSA1	0	the spherically averaged EMD is not read
		≠0	the spherically averaged EMD is read from input
	NELVAL		number of valence electron for which the spherically averaged EMD has been computed. If `NSA1`=0 then `NELVAL` becomes a dummy entry
			if NSA1 ≠ 0 insert NSA1 records
• ∗	P		value of $\|\mathbf{p}\|$
	SAEMD		value of the spherically averaged EMD $\pi(\|\mathbf{p}\|)$
			end
			if IDIR = 0
			insert **MAPNET** input records (Section 14.11, page 339)
			else if IDIR = 1
• ∗	H K L		three integers defining the first direction with respect to the conventional cell
	H' K' L'		three integers defining the second direction with respect to the conventional cell
			end if

EMDWFKIN

The **EMDWFKIN** keyword activates the computation of the kinetic tensor $\mathbf{T}^i$ related to a selected Wannier function W_i. The contribution of W_i to the electron momentum density $\pi_i(\mathbf{p})$ is computed and then its second moment is obtained via numerical integration (see the discussion in section 18.8 at page 401).

rec	variable	value	meaning
• ∗	WF		label of the selected WF
	NTOL1	[1]	initial tolerance used to calculate the Dipole Moment matrix elements: $10^{-\texttt{NTOL1}}$
• ∗	PMAX		maximum momentum value (a.u.) for which the EMD is computed
	PINT		Maximum momentum value (a.u.) defining an inner sphere where the EMD is computed exactly (PINT < PMAX)
	STEP		discretization step for the computation of the EMD in the inner region of **p**-space (p < PINT)
	STEPDIST		discretization step for the computation of the EMD in the outer region of **p**-space (PINT < p < PMAX)

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) [177].

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 64). 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		$10^{-\text{ICNVMP}}$ is the convergence threshold for the Magnasco-Perico pre-localization 8
	IOVPOP		Just those pairs of atoms whose overlap population are greater than $10^{-\text{IOVPOP}}$ 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 64). An inter-atomic distance lower than the capture indicates that I and J can be covalently bonded. Default value [2.0].

Symmetry adaptation of Wannier Functions

SYMMWF

The WnFs *a-posteriori* symmetrization procedure [45], activated by the **SYMMWF** keyword, is mandatory in the case of a subsequent CRYSCOR calculations. A brief outline of the procedure can help to orient in the particular nomenclature adopted to define the symmetrized WnFs and their symmetry relations.

1. after the localization step, the WnFs are centered at different Wickoff sites of the reference cell which are invariant with respect to point-symmetry subgroup H of the space group G of the crystal;

2. for each site, a coset decomposition of G by H is performed, thus leading to the definition of $N_F = \frac{|G|}{|H|}$ symmetry operators (coset representatives) which rotate the reference site into equivalent ones;

3. among each set of equivalent sites, a reference site is chosen;

4. the WnFs of the reference site are symmetrized according to the corresponding subgroup H: each of these symmetrized WnFs is basis of an irreducible representations (IRREP) of the subgroup H and hereafter each component of the irreducible basis set will be referred to as *petal*;

5. the collection of *petals* belonging to the same IRREP of H ,the whole irreducible basis set, constitutes a so-called *flower*. It is worth noting that bi/three-dimensional IRREPs gives rise to *flower* made up of two/three *petals*, respectively;

6. the rotation of the reference *flower* performed by means of the corresponding N_F-1 coset representatives (identity excluded) yields the creation of others N_F-1 symmetry related *flowers*; the set of such equivalent *flowers* constitutes a *bunch*;

7. in general, more than one *bunch* could be associated to the same reference site;

8. as a result of this procedure, each WnF is fully classified by four indices $|\mathtt{b,f,p,g}\rangle$ (`b` = bunch, `f`=flower, `p`= petal, `g`= crystal cell) such that a general symmetry operator of the system $\widehat{R} \in$ G, applied to a WnF gives: $\widehat{R}\,|\mathtt{b,f,p,g}\rangle = \sum_{p'} [A(R)]_{pp'}\ |\mathtt{b,f}^R\mathtt{,p',g}^R\rangle$.

In addition to the general keywords of the localization step, a set of optional keywords, recommended to developers only, can be used to modify some default settings.
The **SYMMWF** input block must be closed by the keyword **END**.

rec	variable	value	meaning
• A	**TOLBOND**		redefinition of the tolerance to classify WFs as *bond* or *atomic* - default value [0.2]
• F	TOLB		if $\|[p_1^i - p_2^i]\| < tolb$ then the WnF ω_i is a *bond* one (p_x^i is the atomic population of the two atoms which most contribute to the WnF i)
• F	**TOLSYM**		a WnF ω_i is classified as already symmetric if $\langle\omega_j\|R\|\omega_i\rangle > TOLSYM$ - default value [0.99]
• A	**NOSYMAP**		an alternative algorithm for the unpacking of WnFs (from the reciprocal to the direct lattice) is followed
• A	**PRINT**		default value [0]

WnFs's Quality. Get WnFs of good quality, in terms of norm and symmetry, is a necessary prerequisites for obtaining reliable energy at the MP2 level. The output file contains some useful information that can be used to check the quality of the solution: as a general and safety rule, the final value for "ERR PER WF" should be less than $1*10^{-5}$ and the precision on the scalar products between WnFs (printed by setting the PRINT option equals 2) should not be less than $1*10^{-5}$.
The solution can be eventually improved by tuning some computational parameters; in particular the user can:
• set tighter tolerances for the evaluation of two-electron integrals in the HF reference solution (**TOLINTEG**): despite the increase of computational time, it turns out that the localization procedure is particularly sensitive to the the first threshold;
• increase the number of **k** points (**NEWK**) according to the suggestions reported in the LOCALWF introduction paragraph;
• use more severe values for the short and large cycles tolerances (**CYCTOL**).
Finally, in some case, the activation of the **FULLBOYS** option can be decisive.

Printing Options and Plot of the WnFs

WnFs can be printed in terms of their coefficients or can be plotted as 2D or 3D maps.

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]

If IPLOT $\neq 0$ insert after the LOCALWF keyword block (defining the localization procedure computational parameters, and terminated by END) the following data:

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 14.11, page 339)

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**, 339. The package Crgra2006 (http://www.crystal.unito.it/Crgra2006.html) allows plotting the function as contour lines. The WnFs and the WnFs densities (in this order) within the selected regions are given in file fort.25.

If IPLO < 0 the WnFs are computed in a 3D grid, which is generated considering the coordinates of WnFs centroids. Insert the following data:

rec	variable	value	meaning
• ∗	ICUBE	0	data are saved in fort.31
		> 0	data are saved in fort.31 and in the CUBE format in the external units *inputfilename*_WAN_CUBE.DAT
	RADIUS	R	defines the spatial region of the 3D grid. The value is in Angstrom
	NPOINTS	NP	number of points along the x direction
	NWFS	0	all the WnFs are plotted
		> 0	number of WnFs to plot
			if NWFS > 0 *insert*
• ∗	WnFs=1,NWFS		sequence number (output order) of the WnFs to be plotted

Let us consider the following input for generating the values of WnFs on a 3D grid of points:

```
NEWK
4 4
0 0
LOCALWF
VALENCE
SYMMWF
END
PRINTPLO
0 0 0 -8
1 2 60 2
1 2
END
END
END
```

The previous input allows for the computation of all the valence (see keyword **VALENCE**) WnFs of the system and for the 3D plotting of the first and second valence WnFs. The data are saved in CUBE format in external units *inputfilename*_WAN_CUBE.DAT. It is possible to visualize the structure of the considered system and the WnFs contained in the CUBE files by using the Jmol program, for instance. In that case, it would be necessary to insert these commands in the Jmol console:
load "file: path-to-the-cube-file"
isosurface sign red blue cutoff 0.7 "file: path-to-the-cube-file"
The default cutoff value is 0.2 but it is possible to modify it as shown above, in order to get a better graphical representation. Some examples of graphical representations are given in Figures (14.1 and 14.2).

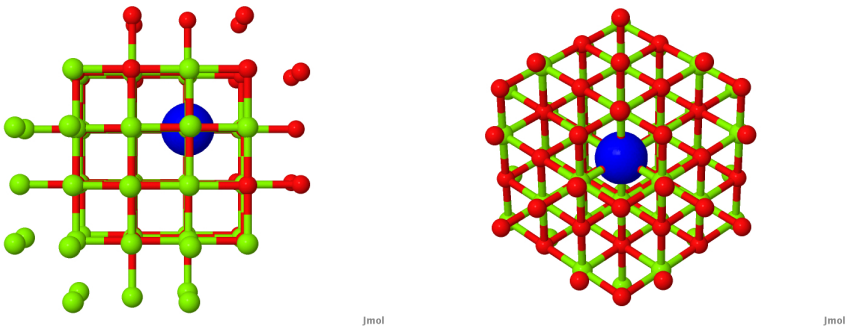


Figure 14.1: First valence WnF of MgO; two different views.

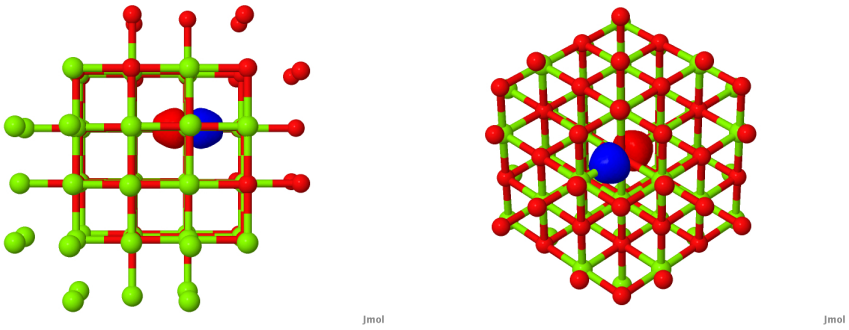


Figure 14.2: Second valence WnF of MgO; two different views.

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 14.3)
		optional keyword II
•	**MARGINS**	definition of a new A",B",C",D" window including ABCD (or A'B'C'D') (see Fig 14.4)
•	∗ 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 fractional 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 D.

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 D, page 441.

5. The position of the three points A, B and C can be specified in two alternative ways:

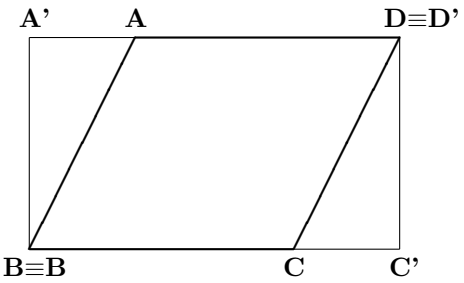


Figure 14.3: Definition of the window where the function F is mapped Effect of optional keyword RECTANGU.

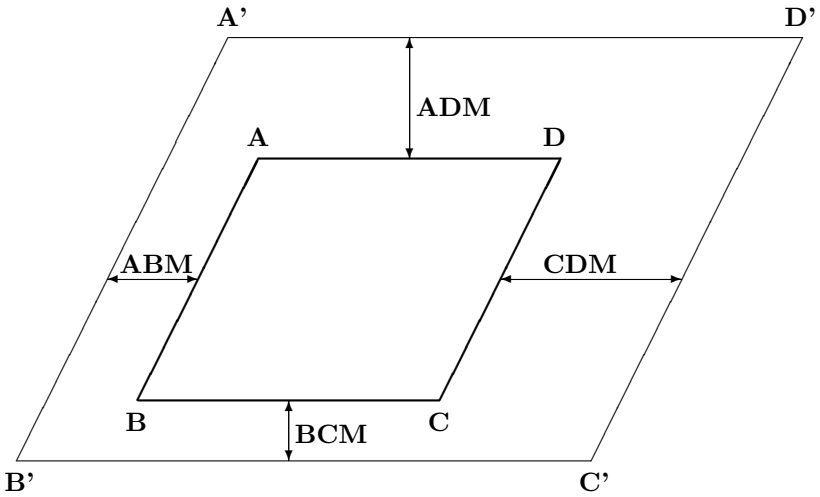


Figure 14.4: Definition of frame around the original window where the function F is mapped. Effect of optional keyword MARGINS.

COORDINA the cartesian coordinates of the three points are given in bohr / Ångstrom / fractional units (default Ångstrom; see Section 3.1, page 35)

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.

NEIGHBOR/NEIGHPRT

See input block 1, page 59

NEWK - Hamiltonian 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 27).
		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 66)* 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 **CRYAPI_OUT** (page 312).

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 18.7, page 401) 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 18, equation 18.5).

3. A Fermi energy calculation must be performed (IFE=1) to run the Compton profiles option **PROF**, the electron momentum density options **EMDL** and **EMDP**, **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 113; **SPINLOCK**, page 124; **BETALOCK**, 94.
   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 117

14.12 3D plotting of crystalline (molecular) orbitals

ORBITALS - Crystalline (Molecular) orbitals 3D plot

rec	variable	value	meaning
• ∗		VAR	filename
• ∗	ICAR	0	Cartesian coordinates (default for molecules)
		1	Fractionary coordinates (default for periodic systems)
	ILOC	0	crystalline orbitals as Bloch functions
		1	crystalline orbitals as Wannier functions

The crystalline (or molecular) orbitals are computed and written in a Molden format to be visualized with graphical tools. Molecular orbitals can be visualized with multiple programs, e.g. Gabedit, Molden and many others supporting the Molden format. Crystalline orbitals are written in a modified Molden format which is currently only supported by a recent version of Jmol.

1. The **NEWK** option *must* be executed (to compute Hartree-Fock/KS eigenvectors and eigenvalues) before running **ORBITALS**.
2. The *filename* corresponds to the row name of the files that contain the information for plotting the orbitals. For periodic systems a file is generated for each real k point while two files are created for complex k points.
3. When ICAR=0, the 3D plot of the orbitals is referred to Cartesian coordinates. This is the default when plotting molecular orbitals.
4. When ICAR=1, orbitals are plotted in fractionary coordinates. This is the default for periodic systems.
5. When ILOC=0, canonical molecular orbitals (or crystalline orbitals - Bloch functions) are calculated
6. When ILOC=1, localized molecular orbitals (or crystalline orbitals - Wannier functions) are computed. Note that the localized orbitals *must* be previously computed by running the option **LOCALI**.

Examples

```
COs - Bloch functions
NEWK
3 3
```

```
Localized-COs - Wannier functions
NEWK
3 3
```

```
1 0
END
ORBITALS
Corundum-COs
1
0
END
END
```

```
1 0
LOCALI
END
ORBITALS
Corundum-LCOs
1
1
END
END
```

14.13 Density Matrix of Non-interacting Atoms

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.
		≠ 0	new basis set and/or new electron configuration is given
	NPR	≠ 0	printing of the density matrix for the first NPR direct lattice vectors
			if IBN ≠ 0 insert basis set input (page 20) 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 reducible density matrix is written to file fort.13.

2. The atomic wave function is computed by the atomic program [240], 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 2.2). 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 93, and **CHARGED**, input block 2, page 75.

PBAN/PDIBAN - Band(s) projected density matrix

rec	variable	meaning
• ∗	NB	number of bands to consider.
• ∗	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.
The **PBAN** keyword must be preceeded by the **NOSYMADA** and **NEWK** (with IFE set to 1) ones (in this order).
To be combined *only* with **ECHG** and **PPAN**.
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).

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 file fort.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 [168] within the Tersoff-Haman approximation [272].

PMP2 - Reads the MP2 correction to the Valence Density Matrix

1. The MP2 correction to the Valence Density Matrix, as computed by the CRYSCOR program, is read from the external unformatted file fort.63, which has to be provided by the user.

2. The **PMP2** keyword must be preceded by **NEWK** (with IFE=1) since the Fermi energy has to be computed in order to split the total SCF DM into core and valence density matrices. The MP2 correction to the DM is summed to the valence SCF DM.

3. By default the PROPERTIES program adopts the SCF density matrix (DM) but if one adds the **PMP2** keyword in the input file then any property whose corresponding keyword is inserted after **PMP2** is computed using the HF+MP2 density matrix.

4. For instance, in order to evaluate the MP2 correction to the X-rays Structure Factors F_{hkl}, the following sequence of keywords has to be used: **NEWK-XFAC-PMP2-XFAC-END**; in the output file will be printed the HF F_{hkl} followed by the HF+MP2 F_{hkl}.

We report in the following the list of the properties for which the MP2 correction can be evaluated, along with some advices one has to take in mind:

- *Directional Compton Profiles* (**PROF**) The ICORE variable must be 1 or 3 and the IVIA variable must be 1
- *Auto-correlation Function and Compton Profiles* (**BIDIERD**) The ICASO variable must be 1 or 3
- *Electron Momentum Density* (**EMDLDM** and **EMDPDM**) The ICASO variable must be 1 or 3
- *Electron Charge Density* (**ECHG**)
- *Mulliken's Populations* (**PPAN**)
- *Structure Factors* (**XFAC**)

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 [224]) 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.

14.14 3D Electrostatic Potential

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 348)

if non-3D system

keyword to choose the type of grid on the non-periodic direction(s):

SCALE	**RANGE**
length scales for non-periodic dimensions	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 D, page 447, for description of the format.

Function data computed at 3D grid points are written according to GAUSSIAN CUBE format in file POT_CUBE.DAT

14.15 2D Electrostatic Potential

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 3.1, 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

```
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 D, page 444.

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 339)

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 339). The electrostatic potential values are written formatted in file fort.25 (see Appendix D, page 441).
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 D, page 441).
3. The electrostatic potential at **r** is evaluated [250] 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 $\rho_\lambda(\underline{r} - \underline{h})$ [249]. 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 **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 [250, 249].
6. The multipoles of shell charges are computed if **POLI** option was not run before **POTM**.

14.16 Mulliken Population Analysis

See input block 3, page 117.

PRINTOUT - Setting of printing environment

See input block 1, page 64.

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 D, page 443)
			2	formatted CP data stored in Fortran unit 24 (Appendix D, page 443)
				if NPR ≠ 0 insert **prtrec** *(see page 66)* II
•	A2		**CP**	calculation of Compton profiles (J(q)) along selected directions (eq. 18.26).
•	∗	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**	auto-correlation function B(r) calculation.
•	∗	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)
•	∗	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. 18.26). 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); the valence contribution is computed numerically (IVIA=0) by integrating the EMD. 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 (auto-correlation 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 64

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 99)

SETINF - Setting of INF values

See input block 1, page 66

SETPRINT - Setting of printing options

See input block 1, page 66.

SPINCNTM - Spin Contamination

This keyword requires a prior call to **NEWK**, p. 341.
In unrestricted calculations (see keywords **UHF** or **SPIN**, p. 126 and 135), the spatial part of the spin-up and down orbitals may be different $\varphi_i^\alpha \neq \varphi_i^\beta$. This results in a larger variational freedom, typically leading to lower energy solutions, but the resulting wavefunction is no longer guaranteed to be an eigenfunction of $\hat{S}^2$, as it may be contaminated by components with higher spin multiplicity.
Let us consider an open-shell system with N_α spin-up electrons and $N_\beta < N_\alpha$ spin-down electrons. The spin-contamination ΔS of the unrestricted wavefunction is defined as follows (a molecular formalism is here adopted for simplicity):

$$\Delta S = \langle \hat{S}^2 \rangle - \langle \hat{S}^2 \rangle_{\text{pure}} = N_\beta - \sum_{i=1}^{N_\alpha} \sum_{j=1}^{N_\beta} \langle \phi_i^\alpha | \phi_j^\beta \rangle \langle \phi_j^\beta | \phi_i^\alpha \rangle \ , \tag{14.24}$$

where $\langle \hat{S}^2 \rangle_{\text{pure}}$ is the ideal expectation value of the $\hat{S}^2$ operator for a pure spin state:

$$\langle \hat{S}^2 \rangle_{\text{pure}} = S_z(S_z+1) \ \text{ with } \ S_z = \left(\frac{N_\alpha - N_\beta}{2} \right). \tag{14.25}$$

In order to evaluate the spin contamination, the last term in Eq. (14.24) has to be computed, which can be rewritten as follows by expressing the spin-orbitals as a linear combination of

atomic orbitals (AOs): $|\phi_i^\sigma\rangle = \sum_\mu C_\mu^{i\sigma}|\chi_\mu\rangle$, with σ either α or β:

$$\sum_{i=1}^{N_\alpha}\sum_{j=1}^{N_\beta}\langle\phi_i^\alpha|\phi_j^\beta\rangle\langle\phi_j^\beta|\phi_i^\alpha\rangle = \sum_{i,j}^{\text{occ.}}\sum_{\mu\nu\lambda\rho} C_\mu^{i_\alpha *}C_\nu^{j_\beta}C_\lambda^{j_\beta *}C_\rho^{i_\alpha}\langle\chi_\mu|\chi_\nu\rangle\langle\chi_\lambda|\chi_\rho\rangle \quad = \sum_{\mu\nu\lambda\rho} P^\alpha_{\rho\mu}P^\beta_{\nu\lambda}S_{\mu\nu}S_{\lambda\rho} = \text{Tr}\big(\mathbf{P^\alpha S P^\beta S}\big) \quad , \tag{14.26}$$

where **P** and **S** are the density and overlap matrices, respectively, in the basis of the AOs.

An example input follows:

```
NEWK      call NEWK to compute eigenvectors
8 8       8*8*8 Monkhort-Pack and Gilat nets
1 0       1 to recalculate Fermi energy
SPINCNTM  now calculate spin contamination
END       end of properties input
```

STOP

Execution stops immediately. Subsequent input records are not processed.

SYMADAPT

See input block 3, page 124

14.17 X-ray Structure Factors and XRD Spectra

XFAC - X-ray structure factors

The Fourier transform of the *static* ground state charge density $\rho(\mathbf{r})$ of a crystalline system provides the set of *static* structure factors $\{F_{hkl}\}$ 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 (*vide infra*):

$$F_{hkl} = \int d\mathbf{r}\rho(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}} \quad \text{with} \quad \mathbf{k} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3 \ ,$$

where $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ are the fundamental reciprocal lattice vectors and h,k,l are Miller's indices. The electron charge density is a function that exhibits translational invariance so that we can consider its expression just in the reference cell $\rho_{cell}(\mathbf{r})$. Let us introduce an, to some extent arbitrary, approximation: the partition of $\rho_{cell}(\mathbf{r})$ into *atomic* contributions, based on some criteria:

$$\rho_{cell}(\mathbf{r}) = \sum_{A\in cell}\rho_A(\mathbf{r}-\mathbf{r}_A)$$

where $\mathbf{r}_A$ is the equilibrium position of atom A in the cell. Within such an approximation the expression of the *static* structure factors simply becomes:

$$F_{hkl} = \sum_{A\in cell} f_A(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}_A} \quad \text{with} \quad f_A(\mathbf{k}) = \int d\mathbf{r}\rho_A(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}$$

A simple way for introducing the effect of thermal nuclear motion in the above expression is represented by the so-called *atomic* Debye-Waller factors $\text{DW}_A(\mathbf{k})$:

$$\tilde{F}_{hkl} = \sum_{A\in cell} f_A(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}_A}\times e^{-\text{DW}_A(\mathbf{k})}$$

where the symbol $\tilde{F}_{hkl}$ represents *dynamic* structure factors and where, if one assumes a harmonic anisotropic potential acting over the atoms and a Gaussian probability of finding an atom displaced by its equilibrium position, the Debye-Waller factors can be expressed as

$$\mathrm{DW}_A(\mathbf{k}) = \frac{1}{2}\langle \mathbf{k}|\mathbf{U}_A|\mathbf{k}\rangle$$

where $\mathbf{U}_A$ is the mean square displacement tensor (that depends upon temperature):

$$\mathbf{U}_A = \begin{bmatrix} \langle x_1^2\rangle & \langle x_1x_2\rangle & \langle x_1x_3\rangle \\ \langle x_2x_1\rangle & \langle x_2^2\rangle & \langle x_2x_3\rangle \\ \langle x_3x_1\rangle & \langle x_3x_2\rangle & \langle x_3^2\rangle \end{bmatrix}$$

with x_i components of the displacement of atom A with respect to its equilibrium position. The above tensor and its eigenvalues are sometimes referred to as *anisotropic displacement parameters* (ADP).

Users of this option for the calculation of dynamical structure factors are kindly requested to cite the following paper [91]:

A. Erba, M. Ferrabone, R. Orlando and R. Dovesi, *J. Comput. Chem.*, **34**, 346 (2013). *Accurate dynamical structure factors from ab initio lattice dynamics: The case of crystalline silicon.*

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
	PC	0	Miller's indices h,k,l refer to the *primitive* cell
		1	Miller's indices h,k,l refer to the *conventional* cell (if any)
			if ISS > 0 *insert ISS records*
• ∗	H,K,L		Miller's indices h,k,l of the reflections
•	A2**DW**		Debye-Waller atomic factors $\mathrm{DW}_A(\mathbf{k})$ are computed to obtain dynamic structure factors. Anisotropic displacement parameters (ADP) for a given temperature are read from the external unit ADP.DAT where they are inserted as tensors
• ∗	UNITS	0	The ADP in the unit are given in atomic units (bohr^2)
		1	The ADP in the unit are given in 10^{-4} Å^2
		2	The ADP in the unit are given in Å^2
	DWPRT	0	Debye-Waller atomic factors are not printed
		1	Debye-Waller atomic factors are printed
•	A2**END**		End of the block of instructions of **XFAC**

Note: if dynamic structure factors are computed for a given temperature T, the corresponding ADPs have to be provided via the external formatted unit ADP.DAT. Such unit has to be put in the scratch directory where the PROPERTIES program is executed. If the standard script `runprop09` is used, then such unit can be renamed *inputfilename.adp* and put in the same directory where the input file *inputfilename.d3* is.
The structure of the ADP.DAT unit is as follows (this example refers to the crystal of silicon):

```
298.1500
1
0.017717   0.000000   0.000000
0.000000   0.017717   0.000000
0.000000   0.000000   0.017717
2
0.017717   0.000000   0.000000
```

```
0.000000    0.017717    0.000000
0.000000    0.000000    0.017717
```

where the first entry is the temperature (in Kelvin) at which the ADPs have been measured/computed. After that, a list of the ADP tensors is given for all the atoms in the cell. In the example above there are just two atoms per cell and their ADP tensors are equal and diagonal.

XRDSPEC - X-ray diffraction Spectra

X-ray static structure factors F_{hkl} can be computed as discussed above, and correspond to a discrete Fourier transform of the electron charge density of the crystal:

$$F_{hkl} = \int_{\text{cell}} \rho(\mathbf{r}) e^{i2\pi\mathbf{k}\cdot\mathbf{r}} \mathrm{d}\mathbf{r} \ , \tag{14.27}$$

where $\mathbf{k} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3$ is a reciprocal lattice vector (being $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ the fundamental reciprocal lattice vectors) and h,k,l Miller's indices. From the CRYSTAL09 version of the program, Debye-Waller thermal factors can be computed to transform static into dynamical structure factors, starting from harmonic atomic anisotropic displacement parameters (ADPs). The intensity of the diffraction peaks is affected by many factors and is typically proportional to:

$$I_{hkl} \propto |F_{hkl}|^2 \times M_{hkl} \times LP(\theta) \times e^{-B} \ , \tag{14.28}$$

where M_{hkl} is the symmetry multiplicity of the structure factor and $LP(\theta)$ is a correction for Lorentz and Polarization effects, with a functional form depending on Bragg's angle θ (linked to Miller's indices through the spacing of the corresponding crystallographic planes and the wavelength of the experiment). In CRYSTAL17, an option has been implemented, which computes the structure factor multiplicities M_{hkl}, converts h,k,l Miller's indices into Bragg's angle θ, and corrects the intensities for the Lorentz and Polarization (LP) effects.

rec		variable	value	meaning
•	∗	NRIF	> 0	a set of non-equivalent reflections with indices h,k,$l < \|NRIF\|$ is generated
		LAMBD	λ	The wavelength λ (in Å)
		B		The isotropic Debye-Waller thermal factor $B = 8\pi^2 \times ADP$ (in Å^2) [typical values in the range 0.5 - 1.5]

14.18 Spontaneous polarization

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 either a Berry phase approach [239, 57] or a localized Wannier functions approach, as the polarization difference between one of the two enantiomorphic structures (λ=+1 or λ=−1) and the intermediate geometric structure (λ=0).
Three subsequent runs are required.

1. First run: preliminary calculation related to λ=0 structure
2. Second run: preliminary calculation related to λ=+1 (or λ=−1) structure
3. Third run: merging of previous data

Some comments:

1. The unit-cell has to contain an even number of electrons.

2. 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.

3. 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.

4. See page 341 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.

5. Data evaluated with the keywords **POLARI** or **LOCALI** in the first two runs do not have any physical meaning if considered independently. Only the output produced choosing the keywords **SPOLBP** or **SPOLWF** in the third run is significant.

6. 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.

7. The spontaneous polarization is obtained through either the Berry phase approach or the localized Wannier functions 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.

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

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 2 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 **SPOLWF** 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 37 as sys0.f37		

2. Second run: preliminary calculation related to system $\lambda = 1$

Program	Keyword	comments
crystal		see deck 2 for input blocks 1 and 1b
properties	NEWK	same input as in first run
	LOCALI	
save Fortran unit 37 as sys1.f37		

3. Third run: merging of previous data.

copy sys0.f37 to Fortran unit 38		
copy sys1.f37 to Fortran unit 39		
Program	Keyword	comments
properties	SPOLWF	

14.19 Mössbauer Spectroscopy

A nuclear spin transition can be promoted when a γ photon, originated from a nuclear spin relaxation, interacts with a chemical equivalent nucleus. The resonant condition are satisfied if the crystal lattice can conveniently absorb the recoil energy. This transition, that in the case of ^{57}Fe occurs between $I_{1/2} \rightarrow I_{3/2}$ and involves an energy of 14.4 keV, can be accompanied by at least two effects resulting in a sensitive shift of the energy levels:

- the isotropic effect and
- the anisotropic or quadrupolar interaction.

The code is capable of calculating the electron density and the eigenvalues of the 3x3 matrix of the electric field gradient, at the nuclei. These particular quantities can be related to experimental observable, as shown in the following subsections, and can provide a deeper insight into the chemical-physical environment surrounding the resonant nucleus. For a better

comprehension of the Mossbauer effect and its exploitation in the characterization of solid state materials, please refer to literature[192, 260, 97, 108].

Isotropic effect (IS)

The energy associated to a nuclear spin transition, ΔE_γ, is directly proportional to the total electron density $\rho(\boldsymbol{r})$ and this means than nuclei subjected to a different field due to their chemical surrounding, absorb at a slightly different frequency. Experimentally what is observed is the shift δ, expressed in terms of the Doppler velocity (mm/s) needed to achieve the resonance absorption between the source S (i.e. the nucleus which emits the γ-ray) and the absorber A (i.e. the nucleus which undergoes the spin transition):

$$\delta = \frac{c}{E_\gamma}(\Delta E_\gamma^A - \Delta E_\gamma^S) \tag{14.29}$$

where c is the light velocity and E_γ the energy of the γ ray. Exploiting the proportionality between ΔE_γ and $\rho(\boldsymbol{r})$, it is possible to state that:

$$\delta = A[\rho_e^A(0) - \rho_e^S(0)] \tag{14.30}$$

where the constant A groups a certain number of terms (relativist effects among the others) for a given isotope. Since all components in equation 14.30 are constant for a given isotope but the electron density of the absorber, it is sufficient to consider the simplified equation:

$$\delta = a[\rho_e^A(0) - b] \tag{14.31}$$

with a and b to be determined in a calibration procedure in which the calculated electron density at a given nucleus is plotted versus the experimentally determined isotopic shift (IS) δ, for a series of compounds containing that nucleus. Once the a and b parameters, for a series of omogeneus compounds (i.e. organic, organo-metallic, inorganic,..) of a given nucleus, have been calculated can be used to infer the chemical shift of such nucleus in systems for which the experimental data are unknown or ambiguous[295, 286].

Anisotropic effect

If the nucleus possess a quadrupolar tensor, Q, the first excited nuclear spin state splits into two double degenerated and equally probable spin states, characterized by the energies $E_{3/2} \pm E_{\mathrm{QI}}$, where the quadrupolar interaction (QI) is of the form:

$$E_{\mathrm{QI}} = -e^2 \frac{1}{6} \sum_{i,j=1,3} V_{ij} Q_{ij} = V \cdot Q \tag{14.32}$$

the Q tensor, representing the deviation of the nuclear charge from the ideal spherical shape, is almost a constant for a given nucleus and can be obtained experimentally. The 3x3 V matrix contains the electric field gradients at the nucleus and can be diagonalize to obtain the eigenvalues $V_{\mathrm{AA}}, V_{\mathrm{BB}}$ and V_{CC} and the asymmetric parameter:

$$\eta = \frac{V_{\mathrm{BB}} - V_{\mathrm{AA}}}{V_{\mathrm{CC}}} \tag{14.33}$$

being always $|V_{\mathrm{CC}}| \geq |V_{\mathrm{BB}}| \geq |V_{\mathrm{AA}}|$. Experimentally a doublet is observed, corresponding to the transitions from the ground state ($I = 1/2$) towards the two excited states and the difference between the two peaks can be directly related to η and V_{CC}:

$$\begin{aligned} \Delta E_{\mathrm{QI}} &= E_{[3/2,\mathrm{QI}]} - E_{[3/2,-\mathrm{QI}]} = E_{3/2} + E_{\mathrm{QI}} - E_{3/2} + E_{\mathrm{QI}} = \\ &= 2E_{\mathrm{QI}} = \frac{1}{2} e\, Q\, V_{\mathrm{CC}} (1 + \frac{\eta^2}{3})^{1/2} \end{aligned} \tag{14.34}$$

where Q is generally expressed in *barn* ($1\,barn = 1\,10^{-28} m$). Equation 14.34 can be used in two ways: (i) if the quadrupolar moment, Q, is known the value of ΔE_{QI} can be estimated; (ii) if ΔE_{QI} is given then Q can be derived.

Input and Output

By mean of the following input (see page 325 for the *ISOTROPIC* keyword and page 347 for the *POTC* one):

```
ISOTROPIC
UNIQUE
POTC
0 0 0
END
```

the following information for metal Iron, at the B3LYP level, can be obtained:

```
..
 ********************************************************************************
 TOTAL DENSITY AT THE NUCLEAR POSITIONS
 ********************************************************************************
 POINT  ATOM     X(AU)     Y(AU)     Z(AU)         BOHR**(-3)
   1    1 FE    0.0000    0.0000    0.0000      11613.823418809
 ********************************************************************************
...
..
 ********************************************************************************

 TOTAL ELECTROSTATIC POTENTIAL

 ********************************************************************************
.....

 TRACELESS ELECTRIC FIELD GRADIENT TENSOR


 POINT   1 POSITION    2.7483  -4.7602   0.5717

 TENSOR IN PRINCIPAL AXIS SYSTEM
 AA -1.316699E-14 BB -3.355213E-12 CC  3.368380E-12

....
 CONVERSION FACTORS (CODATA RECOMMENDED VALUES 1998)
...
 THE PRODUCT OF THE EFG COMPONENT [AU] AND THE NUCLEAR
 QUADRUPOLE MOMENT COMPONENT [BARN] MUST BE MULTIPLIED BY 6.073349
 TO OBTAIN THE QUADRUPOLAR COUPLING CONSTANT [mm/s]

 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT POTC        TELAPSE       95.57 TCPU       95.37
```

The tensor "in principal axis system" represent the electric field gradients (EFG). To get the multiplicative factor, let's rewrite equation 14.34 in terms of units (written in bold for sake of clarity):

$$\begin{aligned}
\Delta E_{\mathrm{QI}}[\boldsymbol{mm/s}] &= \frac{1}{2} e\, Q\, V_{\mathrm{CC}} (1 + \frac{\eta^2}{3})^{1/2} \\
&= \frac{1}{2}\ 1.6021761 \cdot 10^{-19} [\boldsymbol{JV^{-1}}]\ Q \cdot 10^{-28} [\boldsymbol{m^2}] \cdot \\
&\quad \cdot 9.71736153 \cdot 10^{21}\, V_{\mathrm{CC}} (1 + \frac{\eta^2}{3})^{1/2}\, [\boldsymbol{Vm^{-2}}] \\
&= \frac{1}{2} 15.5689 \cdot 10^{-26}\ Q V_{\mathrm{CC}} (1 + \frac{\eta^2}{3})^{1/2}\, \boldsymbol{J}
\end{aligned}$$

using the following conversion factors:

$$\begin{aligned} 1\boldsymbol{J} &= 6.2415097510^{18}\boldsymbol{eV} \\ 1\boldsymbol{eV} &= 0.12510^{8}\boldsymbol{mm/s} \end{aligned}$$

we end with:

$$\Delta E_{QI} = 6.073349\ Q\,V_{\mathrm{CC}}(1+\frac{\eta^2}{3})^{1/2}\ [\boldsymbol{mm/s}] \tag{14.35}$$

where Q has to be expressed in *barn* and V_{CC} and η in atomic unit.

14.20 Topological analysis

The *TOPOND*[115] public code, written by C. Gatti and interfaced to previous public versions of CRYSTAL[43], has presently been embedded in the code itself. By mean of the keyword *TOPO*, after the evaluation of the wave function, it is now possible to perform a topological analysis of the electron density, according to the Quantum Theory of Atoms in Molecules (QTAIM) developed by Bader and coworkers[11].

The merge of TOPOND into CRYSTAL allows to exploit parallel computing. Starting from CRYSTAL22, the TOPOND module has been extended to also work in terms of f- and g-type basis functions.

Users of this module are kindly reminded to cite the following references:

V. R. Saunders, C. Gatti and C. Roetti, *J. Chem. Phys.*, **101**, 10686 (1994) Crystal field effects on the topological properties of the electron density in molecular crystals. The case of urea.

S. Casassa, A. Erba, J. Baima and R. Orlando, *J. Comput. Chem.*, **36**, 1940-1946 (2015) Electron Density Analysis of Large (Molecular and Periodic) Systems: A Parallel Implementation.

A. Cossard, J. K. Desmarais, S. Casassa, C. Gatti and A. Erba, *J. Phys. Chem. Lett.*, **12**, 1862–1868 (2021) Charge Density Analysis of Actinide Compounds from the Quantum Theory of Atoms in Molecules and Crystals

A. Cossard, S. Casassa, C. Gatti, J. K. Desmarais and A. Erba, *Molecules*, **26**, 4227 (2021) Topology of the Electron Density and of its Laplacian from Periodic LCAO Calculations on f-Electron Materials: The Case of Cesium Uranyl Chloride

The QTAIM allows to perform a detailed study of the electron density through different steps, ruled by various keywords the use of which is fully explained and documented in the *TOPOND* Manual. A brief summary is here presented.
The first step in the study of the electron density is the search of its critical points (CP) i.e. the points where its gradient, $\nabla\rho(r)$, vanishes. CPs can be classified in terms of their type and a two-way correspondence with chemically recognizable structures, namely atoms, bonds, ring and cages, can be performed providing lighting information on the bond nature[116]. A second step concerns the topological analysis of the Laplacian of the electron density, $\nabla^2\rho(r)$. This analysis can reveal the atomic shell structure and the degree of sharing of paired electrons among neighboring atoms. A comparison with the corresponding properties in the case of isolated molecules, or atoms, enable to evaluate the effects of the crystal packing on the bonding structure. A third step deals with the determination of the atomic basins and their local and integrated properties. Electronic population, Lagrangian and Hamiltonian kinetic energy, virial density, Becke electron localization function[21] can be defined and calculated in terms of atomic contributions[117]. Finally, for sake of completeness, a certain number of quantities can be plotted and visualize in 2 and 3 dimensions: in figure 14.5 the electron density, its Laplacian and the gradient trajectories of the Urea crystal are reported.

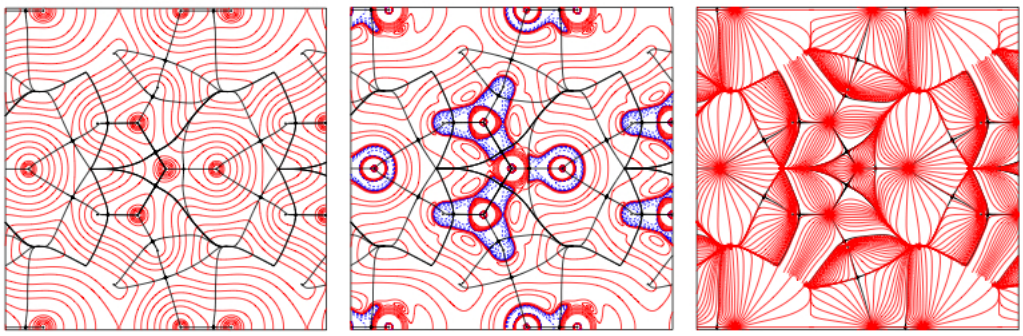


Figure 14.5: From left to right: electron density, ρ, its Laplacian, $\nabla^2\rho$ and the $\nabla\rho$ trajectories for Urea bulk. Bond paths are indicated in heavy black lines and the bond critical points are denoted by filled circles. Dotted blue lines indicate negative contour levels.

Chapter 15

Running CRYSTAL in parallel

The CRYSTAL package contains the following parallel binaries:

- *Pcrystal*
- *MPPcrystal*
- *crystalOMP*
- *PcrystalOMP*
- *MPPcrystalOMP*

It also contains the following packages for the parallel calculation of one-electron properties:

- *Pproperties*

15.1 *Pcrystal* and *MPPcrystal*

Pcrystal (parallel crystal) and *Pproperties* (parallel properties) are replicated-data versions of *crystal* and *properties*, whereas *MPPcrystal* (massively parallel crystal) distributes data and tasks to processes more efficiently than *Pcrystal* or *Pproperties*, and is particularly suitable for large unit cell cases with large memory requirements; the distribution of objects amongst the processes in the MPP versions means in principle much larger systems can be solved, especially if many processes are used. *Pcrystal* and *MPPcrystal* differ essentially in the way they handle data in reciprocal space and for the algorithms used to diagonalize the Fock matrix and process the eigenvectors. On the other hand, data in real space, such as the calculation of one and two-electron integrals, are processed by *Pcrystal* and *MPPcrystal* through the same algorithms and parallelization strategy, and as such have the same memory requirements in the two codes; it is this section of the code that can benefit from the OpenMP approach mentioned below when memory usage becomes prohibitive.

Parallelism in *Pcrystal*, *Pproperties*, *MPPcrystal* is process based, and implemented solely through MPI. *Pcrystal* and *Pproperties* link no other libraries. *MPPcrystal* relies on the use of ScaLAPACK and associated libraries.

The *crystalOMP*, *PcrystalOMP* and *MPPcrystalOMP* binaries employ OpenMP to exploit thread based parallelism. In the cases of *PcrystalOMP* and *MPPcrystalOMP* this is over and above the use of MPI mentioned in the above paragraph. The recommendation is that initial use of parallelism in a project using CRYSTAL should be with the purely MPI based binaries, but should they not prove suitable, the OMP binaries should be investigated; see the section describing the use of the OMP binaries for a description of when this might be appropriate.

15.2 Running *Pcrystal* and *Pproperties*

Pcrystal is fairly efficient for small up to medium sized systems, particularly with high symmetry. Matrices in reciprocal space are distributed to processes over the irreducible representations of the group of the system. In fact, if represented in the basis of the symmetry adapted atomic orbitals (**SYMADAPT**, default; see page 124), such matrices are decomposed into a block-diagonal form, each block (or sub-matrix) corresponding to a row of an irreducible representation. In the case of no symmetry adaptation (**NOSYMADA**, page 117) or P1 symmetry, matrices are distributed over the k points of the Pack-Mokhorst net (or the k points along a selected path in *Pproperties*, for example, to plot energy bands). Thus, the number of processes to be employed to run *Pcrystal* efficiently should not exceed the total number of the irreducible representations for all k points in the Pack-Mokhorst net (or the number of k points when P1 or **NOSYMADA**), unless the diagonalization step is negligible in terms of computational time with respect to the other steps of a calculation. Double-counting must be considered for open-shell systems.

Both *Pcrystal* and *Pproperties* expect to read input data from a file called INPUT in the directory which the executable is running. Output data is written to the standard error in an attempt to avoid buffering by any batch system.

Use of option **CHUNKS** (page 142) is recommended to improve load balancing in DFT calculations. Option **LOWMEM** can be used to reduce memory requirements.

As stated above parallelism in *Pcrystal* and *Pproperties* is enabled via the MPI library. Communication speed is generally not an issue in this case as inter-process communication is limited to a relatively small amount of data at few steps throughout a calculation. The limit on the available parallelism is the number of irreducible representations that can be simultaneously diagonalised, as noted above.

15.3 Running *MPPcrystal*

15.3.1 *MPPcrystal*

MPPcrystal is designed to compute the total energy and wavefunction of large unit cell systems with large memory requirements, as matrices in the reciprocal space are fully distributed over the processors. [198] In particular, MPP running is recommended when n_{irr}/n_{procs} (or n_k/n_{procs}), *i.e.* the ratio of the number of irreducible representations at all k points (or the number of k points sampled in the reciprocal space) to the number of processors used is small. Note, unlike *Pcrystal*, the amount of parallelism that can be exploited is not limited by the number of irreducible representations.

The advantages of *MPPcrystal* include:

- Matrix diagonalization is well balanced because many processors are involved in the diagonalization of one matrix;
- Memory requirement per processor decreases with n_{procs} because data are distributed to processors
- I/O is limited to reading the user's input deck and writing the output files with the results of a job.

Such features make it particularly suitable to run on High Performance computers, but it can also run on smaller clusters. Since communications are more intensive than with *Pcrystal*, performance is improved by a fast interconnect.

To obtain the best performance from *MPPcrystal* the user should give careful consideration when choosing n_{procs}, the number of processes on which they intend to run the job. The reasons for this are two-fold:

- *MPPcrystal* makes much heavier use of the interconnect than *Pcrystal*, and especially on slow interconnects at high process counts the calculation may become inefficient due to the time to solution being dominated by the time to communicate data between the processes.
- When using more than one 1 kpoint *MPPcrystal* is not especially intelligent about how it divides the work up amongst the available cores. As such a careful choice of the number of cores can result in significant performance benefits.

As a very rough guide to the number of processes to use in order to avoid slow down due to communication the number of processes should be chosen such that $n_s * n_k * n_{AO}/n_{procs} \gtrapprox 50$ where n_{AO} is the number of atomic orbitals (basis functions) per unit cell, n_s is the number of spin states, 1 for restricted calculations, 2 for unrestricted, and n_k is the number of irreducible k points. Note this relationship is very much a summary of experience gained on a number of computational systems over the years and has no theoretical underpinning; if the very best performance is required the code will need to be benchmarked on the appropriate computational system. The speed of the interconnect is crucial, fast communications may allow more processes to be used, while a slow one may mean that the limit is much lower.

The second consideration is to choose a number of processes that enables *MPPcrystal* to use dual-level parallelism. In this *MPPcrystal* attempts to split the set of available processes into $n_s * n_k$ subsets, each of which will deal with a single k point at a given spin. To enable this the user should examine the set of k points and their type, and run on a number of processes given by $n_{procs} = n_r + n_c *$ WEIGHT where n_r and n_c denote the number of k points in the Pack-Monkhorst net where the Fock matrix elements are real or complex numbers, respectively, and WEIGHT (see **CMPLXFAC** on page 96) accounts for the extra computational expense required diagonalize a complex matrix relative to a real one. The default value for WEIGHT is 2. Thus if using the default and a calculation has 1 real k point and 7 complex ones, for best efficiency the calculation should be run on a multiple of 15 processes. If this is not the case the current implementation of *MPPcrystal* does not take advantage of dual level parallelism, and the performance penalty can be quite severe. Note if an inefficient choice is made for the number of processes a warning is printed similar to the one below

```
*****************************************************************************
  Please, consider that for this calculation a double level of parallelism
  can be exploited by using a number of processors given by:

      N_proc =     44 x a_non_prime_integer (as 4, or 6, or 8, or 9, ...)

  For instance, N_proc =      176, or      264, or      352, or      396, or ...
                N_proc =     2816, or     5632, or    11264, or    22528, or ...
                N_proc =    45056, or    90112, or ...
*****************************************************************************
```

MPPcrystal sets the defaults:

- "Divide and Conquer" method for matrix diagonalization (disabled by **STDIAG**);
- Reduced memory storage is enabled as with **LOWMEM** (page 114); disabled by **NOLOWMEM**
- two- and one-electron integrals are computed in real time within a SCF cycle as enabled by **SCFDIR** (page 118);
- multipole moments for the evaluation of those two-electron integrals which are approximated through bipolar expansions and exceed the storage memory limit set by **BIPOSIZE** and **EXCHSIZE** are computed in real time (they are stored to disc by *Pcrystal*);

- Data relative to the grid of points used to integrate the exchange-correlation functionals are distributed over n_{procs} cores as enabled by **DISTGRID** (page 142); disabled by **REPLGRID**;
- Symmetry adaptation of the crystalline orbitals is inhibited as with **NOSYMADA** (page 117);
- Eigenvectors are stored to memory and distributed over all cores; they cannot be printed and are not stored to disc at the end of a calculation.

The following options are not available:

	keyword	block
CPHF	**CPHF**	1
Infrared intensities in phonon spectra	**FREQCALC/INTENS**	1
Spin-Orbit Coupling and related	**TWOCOMPON/SOC**	3
Raman intensities in phonon spectra	**FREQCALC/INTRAMAN**	1
Piezoelectric constants	**PIEZOCON** and **ELAPIEZO**	1
Photoelastic constants	**PHOTOELA**	1
Bloch Functions Symmetry Adapted	**SYMADAPT**	3
Printing of eigenvalues of overlap matrix in k space	**EIGS**	3

SCF convergence may proceed along slightly different paths with *Pcrystal* and *MPPcrystal*, particularly in those cases where convergence is difficult to achieve, so that SCF convergence acceleration methods may exhibit different behaviours. A particular setting of convergence tool parameters, for example a given mixing rate of matrices along several SCF steps, may be effective with *Pcrystal* and ineffective with *MPPcrystal*. However, a proper choice of the parameters will always result in achievement of the SCF cycle convergence in both cases.

Like *Pcrystal*, *MPPcrystal* expects to read input data from a file called INPUT in the directory in which the executable is running. Output data is again written to standard error.

MPPcrystal runs as *Pcrystal* if keyword **REPLDATA** (page 118) is added to the input deck.

15.4 Running with MPI+OpenMP parallelism in CRYSTAL

The primary purpose of adding shared memory OpenMP parallelism to *MPPcrystal* was to address situations where the calculation requires more memory per core than is available. In previous release of CRYSTAL this would require running with some of the cores on the node unused. The addition of OpenMP allows the unused cores to be exploited.

In the simplest case if there is sufficient memory per core to run the calculation on all cores with MPI, then just use the MPI-only versions and ignore the OpenMP options.

If each MPI process needs more memory per core than is available then the OpenMP version can be employed to exploit the unused cores and recover some of the lost performance. The number of threads is specified by including

```
export OMP_NUM_THREADS=n
```

(for sh/bash users) in the job script before the command that runs CRYSTAL, where `n` is the number of threads per MPI process. For example, when running 24 MPI processes on a 48 core node use `OMP_NUM_THREADS=2`.

Accordingly possible combinations one might use when running on 8 nodes, each with 64 cores would be

- 512 MPI processes without OpenMP
- 256 MPI processes, each with 2 OpenMP threads
- 128 MPI processes, each with 4 OpenMP threads
- 64 MPI processes, each with 8 OpenMP threads

As to picking the best number of processes and threads similar observations as for *Pcrystal* and *MPPcrystal* apply, and if the very best perofrmance is required the code will need to be benchmarked.

On a dual socket system one should always run an even number of MPI tasks so that each processor chip gets a equal number of those tasks, e.g. on a 48 core node 3 MPI processes with 16 threads would not be recommended. The exception to this would be when there is only 1 MPI process on each node because it needs all the memory.

There are a few other situations where use of the OpenMP enabled versions of the CRYSTAL is beneficial

- *MPPcrystal* reports that increasing the bipolar buffer can improve performance (see 95), but runs out of memory when using the recommended value. The threaded version may provide access to the extra memory required.
- When using a feature only currently available in *Pcrystal*, e.g. spin-orbit coupling, but the calculation requires the use of more cores. The OpenMP version of *Pcrystal* may be able to scale to 2 or 4 cores per k point unlike the MPI-only version which is limited to 1.
- Running simple calculations on a muli-core desktop it is possible to avoid the complication of using MPI in *Pcrystal*. The OpenMP version of *crystal* may be simpler to use.

The current OpenMP distribution is built using Intel's MKL library. On non-Intel processors you will probably need to use

```
export MKL_DEBUG_CPU_TYPE=5
```

in order to get the best performance. Alternatively you can rebuild the executable from the object file package using an appropriate threaded LAPACK library for your machine. We also recommend that you set

```
export OMP_STACKSIZE=16M
```

(or larger) when using the OpenMP version in this release.

Chapter 16

Input examples

16.1 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, *P4₂32*, 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, *P4₂/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, *P6₃/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), α°
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), α°
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), β°
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, $Fm\bar{3}m$, 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), β°
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, $Fd\bar{3}m$, 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 66), 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 20, 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*

16.2 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 0		nuclear charge; number of terms in eq. 3.18 and 3.19
344.84100	-18.00000	-1	eq. 3.18, 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. 3.19, 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. 3.19, 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. 3.19, 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.		

16.3 SCF options

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 27)
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 (FRACTIONAL 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
================================================================================
```

16.4 Geometry optimization

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:

```
 OPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPTOPT
 ******************************************************************************
 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 FRACTIONA LUNITS
 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 16).

▷ 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
`FRAGMENT`	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 16). 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 17

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.

17.1 Molecular BSs performance in periodic systems

Two sets of all electron basis sets are included in CRYSTAL (see Chapter 2.2):

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) [146, 145, 218, 217].

   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 [26, 123]. 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.

17.2 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.

17.3 Valence functions

Molecular crystals

Molecular BSs, minimal and split-valence, are perfectly adequate. Tests have been performed on bulk urea [72] 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 17.1. The last entry of Table 17.1 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 [139] 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 [197, 49].

Table 17.1: 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 [79, 48].
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) [48].
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 [79]. 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 17.1; 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 [244]).

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 17.2: Example of BS for metallic lithium and beryllium derived from the standard 6-21G BS

.

Lithium				Beryllium			
shell	Exp.	Coeff.		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 17.2. Finally, the most diffuse gaussian can be optimized; in the two cases the minimum has not been found owing to numerical instabilities.
See [73] for a more extensive discussion of the metallic lithium case.

17.4 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 17.1 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 17.1 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 2.3 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 17.2), 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).

17.5 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 98). 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 [160] requires basis functions linearly independent. A symptom of numerical dependence may produce an error message in RHOLSK or CHOLSK while running **scf**.

Chapter 18

Theoretical framework

18.1 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).

$$\psi_i(\mathbf{r};\mathbf{k}) = \sum_\mu a_{\mu,i}(\mathbf{k})\phi_\mu(\mathbf{r};\mathbf{k}) \tag{18.1}$$

$$\phi_\mu(\mathbf{r};\mathbf{k}) = \sum_{\mathbf{g}} \varphi_\mu(\mathbf{r}-\mathbf{A}_\mu-\mathbf{g})\; e^{i\mathbf{k}\cdot\mathbf{g}} \tag{18.2}$$

$\mathbf{A}_\mu$ 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:

$$\varphi_\mu(\mathbf{r}-\mathbf{A}_\mu-\mathbf{g}) = \sum_j^{n_G} d_j\; G(\alpha_j;\mathbf{r}-\mathbf{A}_\mu-\mathbf{g}) \tag{18.3}$$

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, $\mathbf{k}$:

$$\mathbf{F}(\mathbf{k})\mathbf{A}(\mathbf{k}) = \mathbf{S}(\mathbf{k})\mathbf{A}(\mathbf{k})\mathbf{E}(\mathbf{k}) \tag{18.4}$$

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:

$$\mathbf{F}(\mathbf{k}) = \sum_{\mathbf{g}} \mathbf{F}^{\mathbf{g}}\; e^{i\mathbf{k}\cdot\mathbf{g}} \tag{18.5}$$

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:

$$F_{12}^{\mathbf{g}} = H_{12}^{\mathbf{g}} + B_{12}^{\mathbf{g}} \tag{18.6}$$

The one electron contribution is the sum of the kinetic and nuclear attraction terms:

$$H_{12}^{\mathbf{g}} = T_{12}^{\mathbf{g}} + Z_{12}^{\mathbf{g}} = \langle \varphi_1^{\mathbf{0}} \mid \widehat{T} \mid \varphi_2^{\mathbf{g}} \rangle + \langle \varphi_1^{\mathbf{0}} \mid \widehat{Z} \mid \varphi_2^{\mathbf{g}} \rangle \tag{18.7}$$

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:

$$B_{12}^{\mathbf{g}} = C_{12}^{\mathbf{g}} + X_{12}^{\mathbf{g}} =$$

$$\sum_{3,4} \sum_{\mathbf{n}} P_{3,4}^{\mathbf{n}} \sum_{\mathbf{h}} [(\varphi_1^{\mathbf{0}} \varphi_2^{\mathbf{g}} \mid \varphi_3^{\mathbf{h}} \varphi_4^{\mathbf{h+n}}) - \frac{1}{2} (\varphi_1^{\mathbf{0}} \varphi_3^{\mathbf{h}} \mid \varphi_2^{\mathbf{g}} \varphi_4^{\mathbf{h+n}})] \tag{18.8}$$

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 $\mathrm{P}^{\mathbf{n}}$ density matrix elements in the AOs basis set are computed by integration over the volume of the Brillouin zone (BZ),

$$P_{3,4}^{\mathbf{n}} = 2 \int_{BZ} d\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{n}} \sum_j a_{3j}^*(\mathbf{k}) a_{4j}(\mathbf{k}) \theta(\epsilon_F - \epsilon_j(\mathbf{k})) \tag{18.9}$$

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:

$$E^{elec} = \frac{1}{2} \sum_{1,2} \sum_{\mathbf{g}} P_{12}^{\mathbf{g}} (H_{12}^{\mathbf{g}} + F_{12}^{\mathbf{g}}) \tag{18.10}$$

A discussion of the different contributions to the total energy is presented in [246, 250] and in Chapter 11 of reference [220].

$$E^{coul} = \frac{1}{2} \sum_{1,2} \sum_{\mathbf{g}} P_{12}^{\mathbf{g}} \sum_{3,4} \sum_{\mathbf{n}} P_{3,4}^{\mathbf{n}} \sum_{\mathbf{h}} [(\varphi_1^{\mathbf{0}} \varphi_2^{\mathbf{g}} \mid \varphi_3^{\mathbf{h}} \varphi_4^{\mathbf{h+n}})] \tag{18.11}$$

$$E^{exch} = -\frac{1}{4} \sum_{1,2} \sum_{\mathbf{g}} P_{12}^{\mathbf{g}} \sum_{34} \sum_{\mathbf{n}} P_{34}^{\mathbf{n}} \sum_{\mathbf{h}} [(\varphi_1^{\mathbf{0}} \varphi_3^{\mathbf{h}} \mid \varphi_2^{\mathbf{g}} \varphi_4^{\mathbf{h+n}})] \tag{18.12}$$

18.2 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).

18.3 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 [78] and by Saunders et al. [250]. It may be summarized as follows.

Consider the Coulomb bielectronic contribution to the Fock matrix ($C_{12}^{\mathbf{g}}$) and to the total energy :

$$E_{ee}^{coul} = \frac{1}{2}\sum_{1,2}\sum_{\mathbf{g}} P_{12}^{\mathbf{g}} \sum_{3,4}\sum_{\mathbf{n}} P_{3,4}^{\mathbf{n}} \sum_{\mathbf{h}} [(\varphi_1^{\mathbf{0}}\varphi_2^{\mathbf{g}} \mid \varphi_3^{\mathbf{h}}\varphi_4^{\mathbf{h+n}}) \tag{18.13}$$

Seven indices are involved in equation 18.13; 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, $\varphi_2^{\mathbf{g}}(\mathbf{r})$ 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: $\{G_{\mathbf{10}}G_{2\mathbf{g}}\}$ and $\{G_{3\mathbf{h}}G_{4\mathbf{h+n}}\}$, 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 18.13 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 ($C_{12}^{\mathbf{g}}$ and $Z_{12}^{\mathbf{g}}$) can be rearranged as follows:

1. Mulliken shell *net* charge distributions are defined as :

$$\rho_\lambda(\mathbf{r}-\mathbf{h}) \equiv \{\lambda\}' \equiv \{\lambda\} - Z_\lambda = \sum_{3\in\lambda}\sum_{4\mathbf{n}} P_{34}^{\mathbf{n}}\ \varphi_3(\mathbf{r}-\mathbf{h})\ \varphi_4(\mathbf{r}-\mathbf{h}-\mathbf{n}) - Z_\lambda \tag{18.14}$$

   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:

$$C_{12}^{\mathbf{g}} + Z_{12}^{\mathbf{g}} = \sum_\lambda \sum_{\mathbf{h}} \int d\mathbf{r}\, d\mathbf{r}' \varphi_1^{\mathbf{0}}(\mathbf{r})\ \varphi_2^{\mathbf{g}}(\mathbf{r})\ |\mathbf{r}-\mathbf{r}'-\mathbf{h}|^{-1}\ \rho_\lambda(\mathbf{r}'-\mathbf{h}) \tag{18.15}$$

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 [250].

The resulting expression for the Coulomb contribution to the Fock matrix is:

$$\begin{aligned} C_{12}^{\mathbf{g}} + Z_{12}^{\mathbf{g}} = \quad & \textstyle\sum_\lambda \{ \sum_{\mathbf{h}}^{B_\lambda} [\sum_{3\in\lambda} \sum_4 \sum_{\mathbf{n}} P_{34}^{\mathbf{n}} (\varphi_1^{\mathbf{0}}\varphi_2^{\mathbf{g}} \mid \varphi_3^{\mathbf{h}}\varphi_4^{\mathbf{h+n}}) + \\ - \quad & \textstyle\sum_{\ell,m} \gamma_\ell^m(\mathbf{A}_\lambda;\{\lambda\}) \Phi_\ell^m(12\mathbf{g};A_\lambda+\mathbf{h})] + \\ + \quad & \textstyle\sum_{\mathbf{h}} \sum_{\ell,m} \gamma_\ell^m(A_\lambda;\{\lambda\}') \Phi_\ell^m(12\mathbf{g};\mathbf{A}_\lambda+\mathbf{h})\} \end{aligned} \tag{18.16}$$

where:

$$\gamma_\ell^m(\mathbf{A}_\lambda;\{\lambda\}) = \int d\mathbf{r}\ \rho_\lambda(\mathbf{r}-\mathbf{A}_\lambda) N_\ell^m X_\ell^m(\mathbf{r}-\mathbf{A}_\lambda) \tag{18.17}$$

$$\Phi_\ell^m(12\mathbf{g};\mathbf{A}_\lambda+\mathbf{h}) = \int d\mathbf{r} \varphi_1^{\mathbf{0}}(\mathbf{r}) \varphi_2^{\mathbf{g}}(\mathbf{r}) X_\ell^m(\mathbf{r}-\mathbf{A}_\lambda-\mathbf{h})\ |\mathbf{r}-\mathbf{A}_\lambda-\mathbf{h}|^{-2\ell-1} \tag{18.18}$$

The Ewald term in eq. 18.16 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 18.16 is that many four-centre (long-range) integrals can be replaced by fewer three-centre integrals.
The attribution of the interaction between $\rho_1 = \{\mathbf{10},\ \mathbf{2g}\}$ 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 117).

18.4 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 [47]. The exchange contribution to the total energy can be written as follows:

$$E^{ex} = \frac{1}{2}\sum_{12}\sum_{\mathbf{g}} P_{12}^{\mathbf{g}}\ [-\frac{1}{2}\sum_{34}\sum_{\mathbf{n}} P_{34}^{\mathbf{n}} \sum_{\mathbf{h}} (\varphi_1^{\mathbf{0}}\varphi_3^{\mathbf{h}} \mid \varphi_2^{\mathbf{g}}\varphi_4^{\mathbf{h+n}})] \tag{18.19}$$

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 $\mathbf{n} = \mathbf{g}$:

$$e_{12}^{\mathbf{g}} = -1/4(P_{12}^{\mathbf{g}})^2(\mathbf{10}\,\mathbf{10}|\mathbf{2g}\,\mathbf{2g}) = -(p^{\mathbf{g}})^2/(4|\mathbf{g}|) \tag{18.20}$$

(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 [222], 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 [224] for a discussion of this point). In order to overcome this problem, the threshold, $P_{ex}^{\mathbf{n}}$ (**ITOL5**) for **n** summation must be more severe than that for **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.

18.5 Bipolar expansion approximation of Coulomb and exchange integrals

We may now return to the partition of the **h** summation in the Coulomb series shown in equation 18.13. 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 **h** (equation 18.14). 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 $\varphi_3^{\mathbf{h}}\varphi_4^{\mathbf{h+n}}$ 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 [224] 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 $G_3^{\mathbf{h}}G_4^{\mathbf{h+n}}$ 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 **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 [78, 250]. 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.

18.6 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 [71] 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:

$$D_{1,2T;3,4Q} = \sum_{\mathbf{n}} [(\varphi_1^{\mathbf{0}} \varphi_2^{\mathbf{g}} \mid \varphi_3^{\mathbf{h}} \varphi_4^{\mathbf{h+n}}) - 1/2(\varphi_1^{\mathbf{0}} \varphi_3^{\mathbf{h}} \mid \varphi_2^{\mathbf{g}} \varphi_4^{\mathbf{h+n}})] \tag{18.21}$$

- 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. M. 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. M. 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

18.7 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 18.9. The technique adopted to compute ϵ_F and the P matrix in the SCF step is described in reference [100]. The Fourier-Legendre technique presented in Chapter II.6 of reference [224] is adopted in the calculation of total and projected DOS. The Fermi energy and the integral in equation 18.9 are evaluated starting from the knowledge of the eigenvalues, $\epsilon_n(\mathrm{k})$ 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*, [191]) 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 18.9 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* [120, 119] 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.

18.8 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, $\pi(\mathbf{p})$ or EMD; the Compton profile function, $J(\mathbf{p})$ or CP; the autocorrelation function, or reciprocal space form factor, or internally-folded density, $B(\mathbf{r})$ 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:

$$\pi(\mathbf{p}) = \sum_{n}^{occ} \sum_{\mu\nu} e^{i\mathbf{p}\cdot(\mathbf{s}_\mu - \mathbf{s}_\nu)} C_{\mu n}(\mathbf{p}_0) C^*_{\nu n}(\mathbf{p}_0)\ \chi_\mu(\mathbf{p}) \chi^*_\nu(\mathbf{p})\ \ \theta(\epsilon_F - \epsilon_n(\mathbf{p}_0)) \quad (18.22)$$

$$\pi(\mathbf{p}) = \sum_{\mu\nu} \sum_{\mathbf{g}} P^{\mathbf{g}}_{\mu\nu}\ e^{i\mathbf{p}(\mathbf{S}_\mu - \mathbf{g} - \mathbf{s}_\nu)}\ \chi_\mu(\mathbf{p})\, \chi^*_\nu(\mathbf{p}) \quad (18.23)$$

In the above equations $\mathbf{p}_0$ is the value of momentum in the Brillouin zone (BZ), which is related to $\mathbf{p}$ by a reciprocal lattice vector $\mathbf{K}$, $\mathbf{s}_\mu$ is the fractional coordinate of the χ_μ centre, and $\chi_\mu(\mathbf{p})$ is the Fourier transform of $\chi_\mu(\mathbf{r})$, calculated analytically:

$$\chi_\mu(\mathbf{p}) = \int d\mathbf{r}\ \chi_\mu(\mathbf{r})\ e^{-\imath\mathbf{p}\cdot\mathbf{r}} \quad (18.24)$$

The two expressions (18.22) and (18.23) are implemented in CRYSTAL and can be used via the **EMDL** (**EMDP**) and **EMDLDM** (**EMDPDM**) keywords, respectively for EMD along a line (in a plane). Let us note, however, that the default choice for keywords **EMDL** and **EMDP** is that of computing the core band contribution via equation (18.23) and the valence band contribution via equation (18.22).

At variance with the electron charge density $\rho(\mathbf{r})$, it is generally difficult to fully exploit the information content of the EMD $\pi(\mathbf{p})$ due to its "collapsed" character about the origin $\mathbf{p} = \mathbf{0}$ in momentum space. A relatively simple way of extracting information from the total EMD of a system, is analyzing its anisotropy $\Delta\pi(\mathbf{p})$ with respect to its Spherical Average (SA) function $\pi_{SA}(p)$ which is, of course, a "radial" function of a single variable:

$$\Delta\pi(\mathbf{p}) = \pi(\mathbf{p}) - \pi_{SA}(p) \qquad \text{where} \qquad p = |\mathbf{p}| \quad (18.25)$$

The average value $\pi_{SA}(p)$ of the function $\pi(\mathbf{p})$ over the surface of a sphere with radius p, can be computed as the average of the function values at the 60 points $\mathbf{p}_n$ (with $n = 1, \ldots, 60$) lying on that surface which form an orbit under the icosahedral group. It is possible to exploit the smoothed character of the spherically averaged EMD $\pi_{SA}(p)$ by explicitly evaluating it on a coarse set of values of p and then by interpolating them (we use a cubic spline). Spherically averaged EMD and EMD-anisotropies can be evaluated via the **EMDLDM** and **EMDPDM** keywords.

A *directional* Compton profile $J_{hkl}(p)$ is related to a particular crystallographic directions $[hkl]$, identified by the unit vector $\mathbf{e}_{hkl}$ where hkl are the *Miller indices*. The directional CP $J_{hkl}(p)$ contains information about the distribution of the electron's momenta along the $[hkl]$ direction of the crystal and can be defined as the 2D integration of $\pi(\mathbf{p})$ over a plane perpendicular to $\mathbf{e}_{hkl}$ through $p\mathbf{e}_{hkl}$:

$$J_{hkl}(p) = \int \pi(\mathbf{p}) \delta(\mathbf{p}\cdot\mathbf{e}_{hkl} - p) d\mathbf{p} \quad (18.26)$$

The weighted average of the directional CPs over all directions is the average CP. Both directional and average CPs can be computed following the expression (18.26) via the **PROF** keyword.

The one-dimensional Fourier transform of a directional CP, gives the so-called directional auto-correlation function $B_{hkl}(r)$:

$$B_{hkl}(r) = \int J_{hkl}(p)\ e^{-\imath p\, r} dp \quad (18.27)$$

$$J_{hkl}(p) = \frac{1}{2\pi} \int B_{hkl}(r)\ e^{\imath p\, r} dr \quad (18.28)$$

Directional CPs can be computed from the corresponding directional auto-correlation function via the keyword **BIDIERD**.

Within the sudden-impulse approximation, $J_{hkl}(p)$ is directly comparable to the outcome of Compton scattering experiments, after correcting the latter for limited resolution and multiple scattering effects. In particular, the effect of limited resolution can be expressed as a convolution of the "infinite resolution" data with a normalized Gaussian function $g(p; \sigma_{cp})$

characterized by a given standard deviation σ_{cp} (or, equivalently, by the fwhm parameter $w_{cp} = \sigma_{cp} \cdot 2\sqrt{2log2}$) which quantifies the experimental resolution:

$$\begin{aligned} J^{\sigma}_{hkl}(p) &= \int_{-\infty}^{+\infty} J_{hkl}(p')\; g(p-p';\,\sigma_{cp})\; dp' \;= \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} B^{\sigma}_{hkl}(r)\;\; e^{-\imath\, p\, r}\, dr \end{aligned} \tag{18.29}$$

In the last integrand a "finite resolution AF" appears, which is simply the *product* of the AF from equation (18.27) by a Gaussian function $g(r, \sigma_{br})$, with $\sigma_{br} = 1/\sigma_{cp}$, and can be extracted from the experimental CP simply by Fourier back-transformation.
The expectation value of the kinetic energy operator $\hat{T}$ (*i.e* the kinetic energy of the system E_{kin}) can be expressed in terms of the second moment of the electron momentum distribution $\pi(\mathbf{p})$ as follows:

$$E_{kin} = \langle \hat{T} \rangle = \frac{1}{2}\int \pi(\mathbf{p})\mathbf{p}^2 d\mathbf{p} \overset{virial}{=} -E_0 \tag{18.30}$$

where the last passage recalls that, if the *virial theorem* is satisfied, the kinetic energy equals (apart from the sign that is opposite) the total energy of the system. Let us introduce the kinetic tensor $\mathbf{T}$ whose elements T_{uv} can be defined as:

$$T_{uv} = \frac{1}{2}\int \pi(\mathbf{p}) p_u\, p_v d\mathbf{p} \quad \text{with} \quad u, v = x, y, z \tag{18.31}$$

The kinetic energy of the system is then the trace of the kinetic tensor:

$$E_{kin} = \langle \hat{T} \rangle = Tr(\mathbf{T}) \tag{18.32}$$

The kinetic tensor can be computed via the **KINETEMD** keyword of PROPERTIES.
A possible strategy for partitioning the electron momentum density (EMD) of a periodic system into chemically meaningful contributions is that of exploiting the spatially localized character of Wannier functions (WF).
Let us introduce the so-called Wannier functions that are real-valued, well localized functions of $\mathbf{r}$, which span altogether the same space as the occupied COs and are translationally equivalent and mutually orthonormal:

$$W_{i,\mathbf{0}}(\mathbf{r}) = W_{i,\mathbf{g}}(\mathbf{r}+\mathbf{g}) \qquad ; \tag{18.33}$$

$$\int W_{i,\mathbf{g}}(\mathbf{r})\, W_{i',\mathbf{g}'}(\mathbf{r})\; d\mathbf{r} = \delta_{ii'}\delta_{\mathbf{g}\mathbf{g}'} \tag{18.34}$$

Such functions are in turn expressed as LCAO:

$$W_{i,\mathbf{g}}(\mathbf{r}) = \sum_{\mu}\sum_{\mathbf{g}'} w_{i,\mathbf{g};\mu\mathbf{g}'}\; \chi_{\mu}(\mathbf{r} - \mathbf{s}_{\mu} - \mathbf{g}') \tag{18.35}$$

The WFs can be obtained from the set of occupied COs via a unitary transformation that imposes spatial localization. Let us express the EMD in terms of WFs:

$$\begin{aligned} \pi(\mathbf{p}) &= \frac{2}{L}\sum_{i=1}^{N_0/2}\sum_{\mathbf{g}}^{L} W_{i,\mathbf{g}}(\mathbf{p})\; W_{i,\mathbf{g}}(\mathbf{p})^* && (18.36)\\ &= 2\sum_{i=1}^{N_0/2} W_{i,\mathbf{0}}(\mathbf{p})\; W_{i,\mathbf{0}}(\mathbf{p})^* && (18.37)\\ &= \sum_{i=1}^{N_0/2} \pi_i(\mathbf{p}) \qquad \text{with} \qquad \pi_i(\mathbf{p}) = 2\; W_{i,\mathbf{0}}(\mathbf{p})\; W_{i,\mathbf{0}}(\mathbf{p})^* && (18.38) \end{aligned}$$

where N_0 is the number of electrons per cell, L the number of cells in the cyclic cluster; in the second passage we have exploited the translational equivalence of the WFs and in the last

passage we make explicit the partition of the total EMD in contributions $\pi_i(\mathbf{p})$ coming from the different WFs.
Let us define a *momentum* WF $W_{i,\mathbf{g}}(\mathbf{p})$ as:

$$W_{i,\mathbf{g}}(\mathbf{p}) = \int d\mathbf{r}\, W_{i,\mathbf{0}}(\mathbf{r}-\mathbf{g})\, e^{-\imath\mathbf{p}\cdot\mathbf{r}} \tag{18.39}$$

$$= e^{-\imath\mathbf{p}\cdot\mathbf{g}} \int d\mathbf{r}\, W_{i,\mathbf{0}}(\mathbf{r})\, e^{-\imath\mathbf{p}\cdot\mathbf{r}} \tag{18.40}$$

$$= e^{-\imath\mathbf{p}\cdot\mathbf{g}} \sum_{\mu} \sum_{\mathbf{g}'} w_{i,\mathbf{0};\mu\mathbf{g}'}\, e^{-\imath\mathbf{p}\cdot(\mathbf{s}_\mu+\mathbf{g}')} \chi_\mu(\mathbf{p}) \tag{18.41}$$

Substitution of equation (18.41) into equation (18.37) gives, for the total EMD:

$$\pi(\mathbf{p}) = 2 \sum_{i=1}^{N_0/2} \sum_{\mu} \sum_{\nu} \sum_{\mathbf{g}} \sum_{\mathbf{g}'} w_{i,\mathbf{0};\mu\mathbf{g}}\, w_{i,\mathbf{0};\nu\mathbf{g}'}\, e^{-\imath\mathbf{p}\cdot(\mathbf{s}_\mu-\mathbf{s}_\nu+\mathbf{g}-\mathbf{g}')} \chi_\mu(\mathbf{p}) \chi^*_\nu(\mathbf{p}) \tag{18.42}$$

Also the total anisotropy $\Delta\pi(\mathbf{p})$ can be partitioned into contributions coming from the different WFs:

$$\Delta\pi(\mathbf{p}) = \sum_{i}^{N_0/2} \Delta\pi_i(\mathbf{p}) \qquad \text{with} \qquad \Delta\pi_i(\mathbf{p}) = \pi_i(\mathbf{p}) - \frac{\pi_{SA}(p)}{N_0/2} \tag{18.43}$$

If one considers as negligible the contribution of core electrons to the total anisotropy of the EMD, can rewrite the previous partitioning as follows:

$$\Delta\pi(\mathbf{p}) = \sum_{i}^{N_v/2} \Delta\pi_i(\mathbf{p}) \qquad \text{with} \qquad \Delta\pi_i(\mathbf{p}) = \pi_i(\mathbf{p}) - \frac{\pi_{SA}(p)}{N_v/2} \tag{18.44}$$

where N_v is the number of valence electrons. Both EMD and EMD-anisotropies from selected WFs can be computed via the **EMDWF** sub-keyword of **LOCALWF** that activates the localization of crystalline orbitals into WFs.
Given the contribution $\pi_i(\mathbf{p})$ of WF i to the EMD of the system, a kinetic tensor $\mathbf{T}^i$ can be introduced whose elements T^i_{uv} can be defined as:

$$T^i_{uv} = \frac{1}{2} \int \pi_i(\mathbf{p}) p_u\, p_v d\mathbf{p} \quad \text{with} \quad u,v = x,y,z \tag{18.45}$$

The computation of such a tensor can be activated via the **EMDWFKIN** sub-keyword of **LOCALWF**.

18.9 Elastic Moduli of Periodic Systems

The elastic constants are *second* derivatives of the energy density with respect to strain components:

$$C_{ij} = 1/V \cdot \frac{\partial^2 E}{\partial\epsilon_i \partial\epsilon_j} \tag{18.46}$$

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 104; **OPTGEOM**, page 182).

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 38) is useful in finding the atomic coordinates to be relaxed. Examples of deformation strategies are discussed in references [79, 46].
In a crystalline system a point $\mathbf{r}$ is usually defined in terms of its fractional components:

$$\mathbf{r} = \mathbf{h}\ \mathbf{L}_p$$

where :

$$\mathbf{L}_p = \begin{bmatrix} \mathbf{l}_1 \\ \mathbf{l}_2 \\ \mathbf{l}_3 \end{bmatrix} = \begin{bmatrix} l_{1x} & l_{1y} & l_{1z} \\ l_{2x} & l_{2y} & l_{2z} \\ l_{3x} & l_{3y} & l_{3z} \end{bmatrix} \tag{18.47}$$

$$V = det(\mathbf{L}_p)$$

$\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3$ 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:

$$\mathbf{r}' = \mathbf{h}\ \mathbf{L}'_p$$

$$\mathbf{L}'_p = (\mathbf{I} + \epsilon)\mathbf{L}_p \tag{18.48}$$

or:

$$\mathbf{L}'_p = \mathbf{L}_p + \mathbf{Z} \tag{18.49}$$

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:

$$\mathbf{L}_p\text{"} = \mathbf{L}'_p\dot{(}V/V')^{1/3} \tag{18.50}$$

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:

$$E(\epsilon) = E(\mathbf{0}) + \sum_i^6 \frac{\partial E}{\partial \epsilon_i}\epsilon_i + 1/2 \sum_{i,j}^6 \frac{\partial^2 E}{\partial \epsilon_i \partial \epsilon_j}\epsilon_i \epsilon_j \tag{18.51}$$

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:

$$\begin{array}{lll} C_{11} & = C_{22} & = C_{33}; \\ C_{44} & = C_{55} & = C_{66}; \\ C_{12} & = C_{13} & = C_{23}; \\ C_{ij} & = 0 & \text{for } i = 1,6, \quad j = 4,6 \quad \text{and } i \neq j. \end{array}$$

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 18.51)

$$E(\epsilon_4) = E(0) + 1/2 \frac{\partial^2 E}{\partial \epsilon_4^2} \delta^2 + \cdots = E(0) + 2 \frac{\partial^2 E}{\partial \epsilon_4^2} x^2 + \cdots = a + bx^2 + \cdots$$

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:

$$E(\epsilon) = E(\mathbf{0}) + 3/2\frac{\partial^2 E}{\partial \epsilon_1^2}\delta^2 + 3\frac{\partial^2 E}{\partial \epsilon_1 \partial \epsilon_2}\delta^2 = \quad (18.52)$$

$$= E(\mathbf{0}) + \frac{3V}{2}[C_{11} + 2C_{12}]\delta^2 \quad (18.53)$$

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.

18.10 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:

$$\varphi_{el}^{(\lambda,1)} = \frac{1}{s2s3} \sum_{j2,j3} \sum_{j_1} \Delta\varphi_{j_1,j_2,j_3}^{(\lambda,1)}(\mathbf{k}) \tag{18.54}$$

The electronic contribution to the polarization of a system λ can be written as :

$$\mathbf{P}_{el}^{(\lambda)} = \frac{1}{\Omega^{(\lambda)}} \left(B^{(\lambda)}\right)^{-1} \varphi_{el}^{(\lambda)} \tag{18.55}$$

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:

$$\mathbf{P}_{nuc}^{(\lambda)} = \frac{1}{\Omega^{(\lambda)}} \sum_{A} \mathbf{R}_A^{(\lambda)} \cdot Z_A \tag{18.56}$$

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

$$\mathbf{P}_{tot}^{(\lambda)} = \mathbf{P}_{nuc}^{(\lambda)} + \mathbf{P}_{el}^{(\lambda)} \tag{18.57}$$

The spontaneous polarization is the difference between the systems $\lambda = 1$ and $\lambda = 0$

$$\mathbf{P} = \mathbf{P}_{tot}^{(\lambda)} - \mathbf{P}_{tot}^{(\lambda)} \tag{18.58}$$

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

$$\mathbf{P}_{el}^{(\lambda)} = \frac{e}{\Omega^{(\lambda)}} \sum_{\mu} \langle \mathbf{r}_\mu \rangle \tag{18.59}$$

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

$$\mathbf{P}_{nuc}^{(\lambda)} = \frac{1}{\Omega^{(\lambda)}} \sum_{A} \mathbf{R}_A \cdot Z_A \tag{18.60}$$

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

$$\mathbf{P}_{tot}^{(\lambda)} = \mathbf{P}_{nuc}^{(\lambda)} + \mathbf{P}_{el}^{(\lambda)} \tag{18.61}$$

The spontaneous polarization is the difference between the both systems $\lambda = 1$ and $\lambda = 0$:

$$\mathbf{P} = \mathbf{P}_{tot}^{(1)} - \mathbf{P}_{tot}^{(2)} \tag{18.62}$$

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 $\lambda = 0$.

18.11 Piezoelectricity through the Berry phase approach

The electronic phase vector of a system λ, is given by (3.1). The nuclear phase vector of a system λ, $\varphi_{nuc}^{(\lambda)}$, can be written as

$$\varphi_{nuc}^{(\lambda)} = \Omega^{(\lambda)}\ B^{(\lambda)}\mathbf{P}_{nuc}^{(\lambda)} \tag{18.63}$$

Where $B^{(\lambda)}$ reciprocal lattice vectors components matrix.The last equation can be simplified thanks to (18.56):

$$\varphi_{nuc}^{(\lambda)} = B^{(\lambda)} \sum_{A} \mathbf{R}_{A}^{(\lambda)} \cdot Z_A \tag{18.64}$$

So the phase vector of a system λ, $\varphi^{(\lambda)}$ is:

$$\varphi^{(\lambda)} = \varphi_{nuc}^{(\lambda)} + \varphi_{el}^{(\lambda)} \tag{18.65}$$

The proper piezoelectric constants can be obtained by:

$$\tilde{e}_{ijk} = -\frac{1}{2\pi}\frac{1}{\Omega}\sum_{\alpha}\frac{d\varphi_\alpha}{d\epsilon_{jk}}\ a_{\alpha,i} \tag{18.66}$$

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:

$$e_{ijk} = \tilde{e}_{ijk} + \delta_{ij}P_k - \delta_{jk}P_i \tag{18.67}$$

In the piezoelectric constants calculations the $\frac{d\varphi_\alpha}{d\epsilon_{jk}}$ term is evaluated numerically. The calculated term is:

$$\frac{d\varphi_\alpha}{d\epsilon_{jk}} \simeq \frac{\Delta\varphi_\alpha}{\Delta\epsilon_{jk}} = \frac{\varphi_\alpha^{(1)} - \varphi_\alpha^{(0)}}{\epsilon_{jk}^{(1)} - \epsilon_{jk}^{(0)}} \tag{18.68}$$

Piezoelectricity through the localized crystalline orbitals approach

The electronic phase vector of a system λ, is given by:

$$\varphi_{el}^{(\lambda)} = \Omega^{(\lambda)}\ B^{(\lambda)}\mathbf{P}_{el}^{(\lambda)} \tag{18.69}$$

Where $B^{(\lambda)}$ reciprocal lattice vectors components matrix. The nuclear phase vector of a system λ, $\varphi_{nuc}^{(\lambda)}$, can be written as

$$\varphi_{nuc}^{(\lambda)} = \Omega^{(\lambda)}\ B^{(\lambda)}\mathbf{P}_{nuc}^{(\lambda)} \tag{18.70}$$

The last equation can be simplified thanks to 18.56:

$$\varphi_{nuc}^{(\lambda)} = B^{(\lambda)} \sum_{A} \mathbf{R}_{A}^{(\lambda)} \cdot Z_A \tag{18.71}$$

So the phase vector of a system λ, $\varphi^{(\lambda)}$ is:

$$\varphi^{(\lambda)} = \varphi_{nuc}^{(\lambda)} + \varphi_{el}^{(\lambda)} \tag{18.72}$$

The proper piezoelectric constants can be obtained by:

$$\tilde{e}_{ijk} = -\frac{1}{2\pi}\frac{1}{\Omega}\sum_{\alpha}\frac{d\varphi_\alpha}{d\epsilon_{jk}}\ a_{\alpha,i} \tag{18.73}$$

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:

$$e_{ijk} = \tilde{e}_{ijk} + \delta_{ij}P_k - \delta_{jk}P_i \tag{18.74}$$

In the piezoelectric constants calculations the $\frac{d\varphi_\alpha}{d\epsilon_{jk}}$ term is evaluated numerically. The calculated term is:

$$\frac{d\varphi_\alpha}{d\epsilon_{jk}} \simeq \frac{\Delta\varphi_\alpha}{\Delta\epsilon_{jk}} = \frac{\varphi_\alpha^{(1)} - \varphi_\alpha^{(0)}}{\epsilon_{jk}^{(1)} - \epsilon_{jk}^{(0)}} \tag{18.75}$$

18.12 Eckart Conditions to the Hessian (Purifying Rotational and Translational Degrees of Freedom)

As the calculation of the Hessian matrix is a numerical procedure, eigenvalues and eigenvectors might be affected by a certain degree of numerical noise, which should be reduced as much as possible, especially if anharmonic calculations are going to be performed on top of the harmonic solution. From an operational point of view, it is important to carefully optimize the structure and to use accurate computational parameters. Some tools are also available to enhance the numerical quality of the Hessian matrix. Sometimes translational and rotational eigensolutions can mix with low frequency modes. In such cases, it is important to project translations and rotations out of the Hessian. This is equivalent to imposing Eckart's conditions to the nuclear motion problem (applied by default; the **NOECKART** option of the **FREQCALC** block switches this option off). We have to generate a projection matrix $\mathbf{P}$ so that translations and rotations (for molecules and polymers) are separated out. Being N_a the number of atoms of the system (in the unit cell for periodic systems), there are $3N_a - N_e$ internal degrees of freedom, where N_c is the number of conditions we have to impose, which is 3 for 3D systems, 6 for non-linear molecules, 5 for linear molecules and 1 for polymers. We define a $\mathbf{D}$ matrix of dimensions $N_c \times 3N_a$, which is the row representation of translations and rotations in the Cartesian frame. The first three vectors (rows) of $\mathbf{D}$ correspond to the three Cartesian translations along $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ and are just the list of the square root of the isotopic mass of the atoms times the corresponding Cartesian unit vector ($\hat{x}$, $\hat{y}$ and $\hat{z}$, for the first, second and third rows, respectively). For example, in the case of water, assuming $m_H = 1$ and $m_O = 16$, we would have

$$\begin{aligned} \mathbf{D}_1 &= (1,0,0,4,0,0,1,0,0) \\ \mathbf{D}_2 &= (0,1,0,0,4,0,0,1,0) \\ \mathbf{D}_3 &= (0,0,1,0,0,4,0,0,1) \end{aligned} \tag{18.76}$$

In order to build a representation for the rotational degrees of freedom, we have first to translate the system so that the origin corresponds to the center of mass (CoM), whose position is given by:

$$\mathbf{r}_{CoM} = \frac{\sum_{\alpha=1}^{N_a} m_\alpha \mathbf{r}_\alpha}{\sum_{\alpha=1}^{N_a} m_\alpha} \tag{18.77}$$

Then, we build the inertia axes tensor:

$$\mathbf{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} = \begin{pmatrix} \sum_\alpha m_\alpha \left(y_\alpha^2 + z_\alpha^2\right) & -\sum_\alpha m_\alpha x_\alpha y_\alpha & -\sum_\alpha m_\alpha x_\alpha z_\alpha \\ -\sum_\alpha m_\alpha y_\alpha x_\alpha & \sum_\alpha m_\alpha \left(x_\alpha^2 + z_\alpha^2\right) & -\sum_\alpha m_\alpha y_\alpha z_\alpha \\ -\sum_\alpha m_\alpha z_\alpha x_\alpha & -\sum_\alpha m_\alpha z_\alpha y_\alpha & \sum_\alpha m_\alpha \left(x_\alpha^2 + y_\alpha^2\right) \end{pmatrix} \tag{18.78}$$

Matrix $\mathbf{I}$ is diagonalized to obtain the eigenvector matrix $\mathbf{X}$ (eigenvectors are organized in columns). For non-linear molecules, the elements of the three additional rotational constraints, $\mathbf{D}_4$, $\mathbf{D}_5$ and $\mathbf{D}_6$, are obtained as:

$$\begin{aligned} D_{4,\alpha i} &= \frac{1}{\sqrt{m_\alpha}} \left((\mathbf{S}_\alpha)_2 X_{i,3} - (\mathbf{S}_\alpha)_3 X_{i,2} \right) \\ D_{5,\alpha i} &= \frac{1}{\sqrt{m_\alpha}} \left((\mathbf{S}_\alpha)_3 X_{i,1} - (\mathbf{S}_\alpha)_1 X_{i,3} \right) \\ D_{6,\alpha i} &= \frac{1}{\sqrt{m_\alpha}} \left((\mathbf{S}_\alpha)_1 X_{i,2} - (\mathbf{S}_\alpha)_2 X_{i,1} \right) \end{aligned} \tag{18.79}$$

where $\mathbf{S}_\alpha$ is an atomic vector obtained as the scalar product of the transpose of $\mathbf{X}$ with the vector of the Cartesian coordinates of atom α:

$$\mathbf{S}_\alpha = \mathbf{X}^T \mathbf{r}_\alpha \tag{18.80}$$

For linear molecules, one of the above conditions is skipped (the one related to the null eigenvalue). Once all the rows of $\mathbf{D}$ are computed, they are ortho-normalized via a Gram-Schmidt procedure. We then build the projector matrix $\mathbf{P}$ (whose dimensions are $3N_a \times 3N_a$) as

$$\mathbf{P} = \mathbf{D}^T \mathbf{D} \tag{18.81}$$

and we define the complementary projector matrix, $\mathbf{P'} = \mathbf{1} - \mathbf{P}$, so that we project out the N_c degrees of freedom from the Hessian matrix $\mathbf{H}$:

$$\mathbf{H'} = \mathbf{P'}^T \mathbf{H} \mathbf{P'} \tag{18.82}$$

By diagonalizing the $\mathbf{H'}$ matrix, the translational and rotational eigenvalues are exactly zero.

Appendix A

Symmetry groups

A.1 Labels and symbols of the space groups

The labels are according to the International Tables for Crystallography [133]. The symbols are derived by the standard SHORT symbols, as shown in the following examples:

Symbol		Input to CRYSTAL
P $\bar{6}$ 2 m	$\rightarrow$	`P␣-6␣2␣M` ;
P 6_3 m	$\rightarrow$	`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. $\rightarrow$ P m 3 m ; 1982 ed. $\rightarrow$ 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	$Cmc2_1$
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	$Imma$
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	$P3_221$

IGR	symbol
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	$P6_3/mmc$

IGR	symbol
	Cubic lattices
195	$P23$
196	$F23$
197	$I23$
198	$P2_13$
199	$I2_13$
200	$Pm\bar{3}$
201	$Pn\bar{3}$
202	$Fm\bar{3}$
203	$Fd\bar{3}$
204	$Im\bar{3}$
205	$Pa\bar{3}$
206	$Ia\bar{3}$
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
	Oblique lattices (P)		
1	$P1$	C_1^1	1
2	$P\bar{1}$	C_i^1	2
3	$P112$	C_2^1	(3)
4	$P11m$	C_s^1	(6)
5	$P11a$	C_s^2	(7)
6	$P112/m$	C_{2h}^1	(10)
7	$P112/a$	C_{2h}^4	(13)
	Rectangular lattices (P or C)		
8	$P211$	C_2^1	(3)
9	$P2_111$	C_2^2	(4)
10	$C211$	C_2^3	(5)
11	$Pm11$	C_s^1	(6)
12	$Pb11$	C_s^2	(7)
13	$Cm11$	C_s^3	(8)
14	$P2/m11$	C_{2h}^1	(10)
15	$P2_1/m11$	C_{2h}^2	(11)
16	$C2/m11$	C_{2h}^3	(12)
17	$P2/b11$	C_{2h}^4	(13)
18	$P2/b11$	C_{2h}^5	(14)
19	$P222$	D_2^1	16
20	$P22_12$	D_2^2	(17)
21	$P2_12_12$	D_2^3	18
22	$C222$	D_2^6	21
23	$Pmm2$	C_{2v}^1	25
24	$Pma2$	C_{2v}^4	28
25	$Pba2$	C_{2v}^8	32
26	$Cmm2$	C_{2v}^1	35
27	$P2mm$	C_{2v}^1	(25)
28	$P2_1am$	C_{2v}^2	(26)
29	$P2_1ma$	C_{2v}^2	(26)
30	$P2mb$	C_{2v}^4	(28)
31	$P2_1mn$	C_{2v}^7	(31)
32	$P2aa$	C_{2v}^3	(27)
33	$P2_1ab$	C_{2v}^5	(29)
34	$P2an$	C_{2v}^6	(30)
35	$C2mm$	C_{2v}^1	(38)
36	$C2mb$	C_{2v}^5	(39)
37	$Pmmm$	D_{2h}^1	47
38	$Pmam$	D_{2h}^5	(51)
39	$Pmma$	D_{2h}^5	51
40	$Pmmn$	D_{2h}^3	59
41	$Pbam$	D_{2h}^9	55
42	$Pmaa$	D_{2h}^3	(49)
43	$Pman$	D_{2h}^7	(53)
44	$Pbma$	D_{2h}^1	(57)
45	$Pbaa$	D_{2h}^8	(54)
46	$Pban$	D_{2h}^4	50
47	$Cmmm$	D_{2h}^9	65
48	$Cmma$	D_{2h}^2	67
	Square lattices (P)		
49	$P4$	C_4^1	75
50	$P\bar{4}$	S_4^1	81
51	$P4/m$	C_{4h}^1	83
52	$P4/n$	C_{4h}^3	85
53	$P422$	D_4^1	89
54	$P42_12$	D_4^2	90
55	$P4mm$	C_{4v}^1	99
56	$P4bm$	C_{4v}^2	100
57	$P\bar{4}2m$	D_{2d}^1	111
58	$P\bar{4}2_1m$	D_{2d}^3	113
59	$P\bar{4}m2$	D_{2d}^5	115
60	$P\bar{4}b2$	D_{2d}^7	117
61	$P4/mmm$	D_{4h}^1	123
62	$P4/nbm$	D_{4h}^3	125
63	$P4/mbm$	D_{4h}^5	127
64	$P4/nmm$	D_{4h}^7	129
	Hexagonal lattices (P)		
65	$P3$	C_3^1	143
66	$P\bar{3}$	C_{3i}^1	147
67	$P312$	D_3^1	149
68	$P321$	D_3^2	150
69	$P3m1$	C_{3v}^1	156
70	$P31m$	C_{3v}^2	157
71	$P\bar{3}1m$	D_{3d}^1	162
72	$P\bar{3}m1$	D_{3d}^3	164
73	$P6$	C_6^1	168
74	$P\bar{6}$	C_{3h}^1	174
75	$P6/m$	C_{6h}^1	175
76	$P622$	D_6^1	177
77	$P6mm$	C_{6v}^1	183
78	$P\bar{6}m2$	D_{3h}^1	187
79	$P\bar{6}2m$	D_{3h}^3	189
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}$m	O_h		24
46	235	I		60
47	m$\bar{3}\bar{5}$	I_h		60

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).

$$P \to A \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \bar{\frac{1}{2}} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \qquad P \to B \quad \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \bar{\frac{1}{2}} & 0 & \frac{1}{2} \end{pmatrix} \qquad A \to P \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix} \qquad B \to P \quad \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$P \to C \quad \begin{pmatrix} \frac{1}{2} & \bar{\frac{1}{2}} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad P \to F \quad \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \qquad C \to P \quad \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad F \to P \quad \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

$$P \to I \quad \begin{pmatrix} \bar{\frac{1}{2}} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \bar{\frac{1}{2}} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \bar{\frac{1}{2}} \end{pmatrix} \qquad R \to H \quad \begin{pmatrix} \frac{2}{3} & \bar{\frac{1}{3}} & \bar{\frac{1}{3}} \\ \frac{1}{3} & \frac{1}{3} & \bar{\frac{2}{3}} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \qquad I \to P \quad \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad H \to R \quad \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}$$

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. [133]).

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	38	–
MAKESAED	printing of symmetry allowed elastic distortions (SAED)	49	–
PRSYMDIR	printing of displacement directions allowed by symmetry.	64	–
SYMMDIR	printing of symmetry allowed geom opt directions	73	–
SYMMOPS	printing of point symmetry operators	73	–
TENSOR	print tensor of physical properties up to order 4	73	I

Symmetry information and control			
BREAKELAS	symmetry breaking according to a general distortion	39	I
BREAKSYM	allow symmetry reduction following geometry modifications	39	–
KEEPSYMM	maintain symmetry following geometry modifications	49	–
MODISYMM	removal of selected symmetry operators	50	I
PURIFY	cleans atomic positions so that they are fully consistent with the group	64	–
SYMMREMO	removal of all symmetry operators	73	–
TRASREMO	removal of symmetry operators with translational components	73	–

Modifications without reduction of symmetry			
ATOMORDE	reordering of atoms in molecular crystals	36	–
NOSHIFT	no shift of the origin to minimize the number of symmops with translational components before generating supercell	59	–
ORIGIN	shift of the origin to minimize the number of symmetry operators with translational components	59	–
PRIMITIV	crystallographic cell forced to be the primitive cell	63	–
ROTCRY	rotation of the crystal with respect to the reference system cell	65	I

Atoms and cell manipulation - possible symmetry reduction (BREAKSYMM)

ATOMDISP	displacement of atoms	36	I
ATOMINSE	addition of atoms	36	I
ATOMREMO	removal of atoms	37	I
ATOMROT	rotation of groups of atoms	37	I
ATOMSUBS	substitution of atoms	38	I
ELASTIC	distortion of the lattice	42	I
POINTCHG	point charges input	63	I
SCELCONF	generation of supercell for configuration counting	68	I
SCELPHONO	generation of supercell for phonon dispersion	68	I
SUPERCEL	generation of supercell - input refers to primitive cell	70	I
SUPERCON	generation of supercell - input refers to conventional cell	70	I
USESAED	given symmetry allowed elastic distortions, reads δ	73	I

From crystals to slabs (3D→2D)

SLABINFO	definition of a new cell, with $xy \parallel$ to a given plane	67	I
SLABCUT	generation of a slab parallel to a given plane (3D→2D)	66	I

From slabs to single and multi-wall nanotubes (2D→1D)

NANOTUBE	building a nanotube from a slab	54	I
SWCNT	building a nanotube from an hexagonal slab	72	I
NANOMULTI	building a multi-wall nanotube from a slab	57	I

From periodic structures to clusters

CLUSTER	cutting of a cluster from a periodic structure (3D→0D)	40	I
CLUSTSIZE	maximum number of atoms in a cluster	41	I
FULLE	building a fullerene from an hexagonal slab (2D→0D)	47	I
HYDROSUB	border atoms substituted with hydrogens (0D→0D)	49	I

Molecular crystals

MOLECULE	extraction of a set of molecules from a molecular crystal (3D→0D)	51	I
MOLEXP	variation of lattice parameters at constant symmetry and molecular geometry (3D→3D)	51	I
MOLSPLIT	periodic structure of non interacting molecules (3D→3D)	52	–
RAYCOV	modification of atomic covalent radii	64	I

BSSE correction

MOLEBSSE	counterpoise method for molecules (molecular crystals only) (3D→0D)	50	I
ATOMBSSE	counterpoise method for atoms (3D→0D)	36	I

Systematic analysis of crystal planes

PLANES	Prints the possible crystal planes	63	I

Gibbs-Wulff construction

WULFF	Building the Gibbs-Wulff polihedron	73	I

From crystals to nanorods (3D→1D)

NANORODS	Building a nanorod from a crystal	53	I

From crystals to nanocrystals (3D→0D)

NANOCRYSTAL	building a nanocrystal from a crystal	52	I

Auxiliary and control keywords			
ANGSTROM	sets input units to Ångstrom	35	–
BOHR	sets input units to bohr	38	–
BOHRANGS	input bohr to Å conversion factor (0.5291772083 default value)	38	I
BOHRCR98	bohr to Å conversion factor is set to 0.529177 (CRY98 value)	–	
END/ENDG	terminate processing of geometry input		–
FRACTION	sets input units to fractional	47	–
LATVEC	maximum number of classified lattice vectors	49	I
MAXNEIGHB	maximum number of equidistant neighbours from an atom	49	I
NEIGHBOR	number of neighbours in geometry analysis	59	I
PRINTCHG	printing of point charges coordinates in geometry output	63	
PRINTOUT	setting of printing options by keywords	64	–
SETINF	setting of inf array options	66	I
SETPRINT	setting of printing options	66	I
STOP	execution stops immediately	68	–
TESTGEOM	stop after checking the geometry input	73	–

Output of data on external units			
COORPRT	coordinates of all the atoms in the cell	42	–
EXTPRT	write file in CRYSTAL geometry input format	44	–
FINDSYM	write file in FINDSYM input format	47	–
STRUCPRT	cell parameters and coordinates of all the atoms in the cell	68	–

External electric field - modified Hamiltonian			
FIELD	electric field applied along a periodic direction	44	I
FIELDCON	electric field applied along a non periodic direction	46	I

Geometry optimization - see index for keywords full list

OPTGEOM	Geometry optimization input block - closed by **END**	182	I
	Type of optimization (default: atom coordinates)		
FULLOPTG	full geometry optimization		–
CELLONLY	cell parameters optimization		–
INTREDUN	optimization in redundant internal coordinates	192	–
ITATOCEL	iterative optimization (atom/cell)		–
CVOLOPT	full geometry optimization at constant volume	197	–
	Initial Hessian		
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		–
HESSNUM	initial guess for the Hessian - numerical estimate		–
	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	200	I
RESTART	data from previous run		–
FINALRUN	Wf single point with optimized geometry		I
	Gradient calculation control		
NUMGRATO	numerical atoms first derivatives	191	–
NUMGRCEL	numerical cell first derivatives	191	–
NUMGRALL	numerical atoms and cell first derivatives	191	–
	External stress		
EXTPRESS	apply external hydrostatic pressure	202	I
	Printing options		
PRINTFORCES	atomic gradients		–
PRINTHESS	Hessian		–
PRINTOPT	optimization procedure		–
PRINT	verbose printing		–

Vibrational Frequencies - see index for keywords full list

FREQCALC	Harmonic Γ-frequencies calculation input - closed by **END**	207	I
	Normal modes analysis		
ANALYSIS		209	–
COMBMODE	TO combination modes and overtones	210	I
MODES	printing eigenvectors [default]	212	–
SCANMODE	scan geometry along selected modes	220	I
	LO/TO splitting		
DIELISO	isotropic dielectric tensor	210	I
DIELTENS	anisotropic dielectric tensor	210	I
	Vibrational spectrum simulation		
INTENS	intensities calculation active	215	–
INTCPHF	IR (and Raman) intensities via CPHF	217	I
INTLOC	IR intensities through Wannier functions	216	–
INTPOL	IR intensities through Berry phase [default]	215	–
INTRAMAN	Raman intensities calculation	217	I
IRSPEC	IR spectrum production	223	I
RAMSPEC	Raman spectrum production	225	I
	Calculation control		
ECKART	Hessian freed by translations and rotations [default]	210	I
FRAGMENT	partial frequency calculation	211	I
ISOTOPES	isotopic substitution	211	I
NORMBORN	normalized Born tensor	212	–
NUMDERIV	technique to compute numerical 2nd derivatives	213	I
PRINT	verbose printing		–
RESTART	data from previous run		–
STEPSIZE	set size of cartesian displacements [0.003 Å]	213	I
TEST[FREQ]	frequency test run		–

USESYMM	full-symmetry exploitation at each point [default]		–
Phonon dispersion			
DISPERSION	frequencies calculated at $\vec{k} \neq \Gamma$ points	227	–
Thermodynamics			
ADP	anisotropic displacement parameters	213	I
PRESSURE	set pressure range	213	I
TEMPERAT	set temperature range		I
ANHARM	Anharmonic frequencies calculation input block - closed by **END**	232	I
ISOTOPES	isotopic substitution	233	I
KEEPSYMM	displace all symmetry equivalent atoms	**??**	–
NOGUESS		233	–
POINTS26	X-H distance varied 26 times around the equilibrium	233	–
PRINT	verbose printing		–
TEST[ANHA]	test run		–

Configurations counting and characterization

CONFCNT	configurations counting and cluster expansion	272	I
CONFRAND	symmetry-adapted uniform at random Monte Carlo	274	I
RUNCONFS	single-point calculations and geometry optimizations	275	I

CPHF - Coupled Perturbed Hartree-Fock 265

ELASTCON - Second order elastic constants 283

EOS - Equation of state 277

Basis set input (Input block 2)

Symmetry control

ATOMSYMM	printing of point symmetry at the atomic positions	38	–

Basis set modification

CHEMOD	modification of the electronic configuration	75	I
GHOSTS	eliminates nuclei and electrons, leaving BS	77	I

Auxiliary and control keywords

CHARGED	allows non-neutral cell	75	–
NOPRINT	printing of basis set removed	77	–
PRINTOUT	setting of printing options	64	I
SETINF	setting of inf array options	66	I
SETPRINT	setting of printing options	66	I
STOP	execution stops immediately	68	–
SYMMOPS	printing of point symmetry operators	73	–
END/ENDB	terminate processing of basis set definition keywords		–

Output of data on external units

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 Hartree Fock 118 –
UHF Unrestricted Hartree-Fock 126 –
ROHF Restricted Open Shell Hartree-Fock 126 –
DFT DFT Hamiltonian 126 –
SPIN spin-polarized solution 135 –

Choice of the exchange-correlation functionals

EXCHANGE exchange functional 129 I
LDA functionals
LDA Dirac-Slater [68] (LDA)
VBH von Barth-Hedin [281] (LDA)
GGA functionals
BECKE Becke 1988[18] (GGA)
PBE Perdew-Becke-Ernzerhof 1996 [204] (GGA)
PBESOL GGA. PBE functional revised for solids [206]
mPW91 modified Perdew-Wang 91 (GGA)
PWGGA Perdew-Wang 91 (GGA)
SOGGA second order GGA. [302]
WCGGA GGA - Wu-Cohen [293]
CORRELAT correlation functional 129 I
LDA functionals
PZ Perdew-Zunger [210] (LDA)
VBH von Barth-Hedin [281] (LDA)
VWN Vosko,-Wilk-Nusair [282] (LDA)
GGA functionals
LYP Lee-Yang-Parr [167] (GGA)
P86 Perdew 86 [202] (GGA)
PBE Perdew-Becke-Ernzerhof [204] (GGA)
PBESOL GGA. PBE functional revised for solids [206]
PWGGA Perdew-Wang 91 (GGA)
PWLSD Perdew-Wang 92 [208, 209, 207] (GGA)
WL GGA - Wilson-Levy [291]
meta-GGA functionals
B95 mGGA - Becke 95 [22]
Standalone keywords: exchange+correlation
SVWN see [68, 282] 129
BLYP see [18, 167] 129
PBEXC see [204] 129
PBESOLXC see [206] 129
SOGGAXC see [302] 129
SOGGA11 see [216] 129

Global Hybrid functionals

Standalone keywords

B3PW	B3PW parameterization	130	–
B3LYP	B3LYP parameterization	130	–
PBE0	Adamo and Barone [5]	130	
PBESOL0	Derived from PBE0	130	
B1WC	see [25]	130	
WC1LYP	see [60]	130	
B97H	see [4, 101]	130	
PBE0-13	see [42]	130	
SOGGA11X	see [215]	130	
mPW1PW91	see [6]	130	
mPW1K	see [175]	130	

User defined global hybrids

HYBRID	hybrid mixing	130	I
NONLOCAL	local term parameterization	131	I

Range-Separated Hybrid functionals

Short-range Corrected RSH functionals

HSE06	Screened-Coulonb PBE XC functional [2, 204]	132	–
HSESOL	Screened-Coulomb PBESOL XC functional [162, 206]	132	–
SC-BLYP	SC-RSH BLYP functional based on ITYH scheme [152, 271]	132	–

Middle-range Corrected RSH functionals

HISS	MC based on PBE XC functional [267, 268]	132	–

Long-range Corrected RSH functionals

RSHXLDA	LC LDA XC functional [1, 151]	132	–
wB97	Chai/Head-Gordon LC functional [153, 4]	132	–
wB97X	Chai/Head-Gordon SC/LC functional [153, 4]	132	–
LC-wPBE	LC hybrid based on PBE XC functional [85]	132	–
LC-wPBESOL	LC hybrid based on PBESOL XC functional [85]	132	–
LC-wBLYP	LC hybrid based on BLYP XC functional [85]	132	–
LC-BLYP	LC-RSH BLYP functional based on ITYH scheme [152, 271]	132	–
CAM-B3LYP	Coulomb-Attenuating method)[270] based on the BLYP XC functional	132	–
LC-PBE	LC-RSH PBE functional based on ITYH scheme [152, 164]	132	–

User defined range separated hybrids

SR-OMEGA	setting of ω_{SR} for SC hybrids	133	–
MR-OMEGA	setting of ω_{SR} and ω_{LR} for MC hybrids	133	–
LR-OMEGA	setting of ω_{LR} for LC hybrids	133	–
SR-HYB_WB97X	Option to change the amount of SR-HF exchange in the ωB97-X functional	133	–
LSRSH-PBE	User-controllable RSH x-functional based on the PBE functional	133	–

meta-GGA functionals

Pure mGGA functionals

M06L	pure mGGA M06-type functional [296]	134	–
revM06L	revised version of pure mGGA M06L functional [284]	134	–
MN15L	58 parameters pure mGGA by Yu, He, and Truhlar [301]	134	–
SCAN	non-empirical mGGA functional based on SCAN XC [264]	134	–
r^{2}SCAN	modified version of regularized SCAN [114]	134	–

Global hybrid mGGA functionals

B1B95	Becke88 exchange combined with Becke95 correlation and 28% of HF exchange [23]	134	–
mPW1B95	mPW91 exchange combined with Becke95 correlation and 31% of HF exchange [303]	134	–
mPW1B1K	mPW91 exchange combined with Becke95 correlation and 44% of HF exchange [303]	134	–
PW6B95	6 parameters mPW91 exchange combined with re-optimized Becke95 correlation and 28% of HF exchange [304]	134	–
PWB6K	6 parameters mPW91 GGA combined with re-optimized Becke95 correlation and 46% of HF exchange [304]	134	–
M05	Minnesota 2005 functional [300]	134	–
M052x	M05-2X functional [299]	134	–
M06	Minnesota 2006 functional [298]	134	–
M062X	M06-2X functional [298]	134	–
M06HF	M06-type functional with 100% HF [296]	134	–
MN15	59 parameters global hybrid mGGA with 44% of HF exchange [138]	134	–
revM06	revised version of global hybrid mGGA M06 with 40.41% of HF exchange [285]	134	–
SCAN0	global hybrid mGGA based on SCAN XC with 25% of HF exchange [149]	134	–
r^{2}SCANh	global hybrid mGGA based on r^{2}SCAN XC with 10% of HF exchange [40]	134	–
r^{2}SCAN0	global hybrid mGGA based on r^{2}SCAN XC with 25% of HF exchange [40]	134	–
r^{2}SCAN50	global hybrid mGGA based on r^{2}SCAN XC with 50% of HF exchange [40]	134	–

Numerical accuracy control

ANGULAR	definition of angular grid	137	I
RADIAL	definition of radial grid	137	I
[BECKE]	selection of Becke weights (default)	137	–
SAVIN	selection of Savin weights	137	–
OLDGRID	"old" default grid	140	
LGRID	"large" predefined grid	140	
[XLGRID]	"extra large" predefined grid (new default)		
XXLGRID	"extra extra large" predefined grid	140	
XXXLGRID	"extra extra extra large" predefined grid (default grid for mGGA based functionals)	141	
HUGEGRID	"ultra extra large grid" predefined grid (default for SCAN based DFAs)	141	
RADSAFE	safety radius for grid point screening		I
TOLLDENS	density contribution screening 6	141	I
TOLLGRID	grid points screening 14	141	I
[BATCHPNT]	grid point grouping for integration	141	I
CHUNKS	max n. points in a batch for numerical int.	142	I
DISTGRID	distribution of DFT grid across nodes	142	
LIMBEK	size of local arrays for integration weights 400	142	I

Atomic parameters control

RADIUS	customized atomic radius	143	I
FCHARGE	customized formal atomic charge	143	I

Auxiliary

PRINTEXC Enable printing of the Exchange and Correlation energy at each SCF cycle. By default in CRYSTAL22 the Exchange and Correlation energy is printed only when the SCF convergence is reached.

END close DFT input block

Numerical accuracy and computational parameters control

BIPOLAR	Bipolar expansion of bielectronic integrals	95	I
BIPOSIZE	size of coulomb bipolar expansion buffer	95	I
EXCHSIZE	size of exchange bipolar expansion buffer	95	I
EXCHPERM	use permutation of centers in exchange integrals	101	–
ILASIZE	Maximum size of array ILA for 2-electron integral calculation 6000	112	I
INTGPACK	Choice of integrals package 0	112	I
MADELIND	reciprocal lattice vector indices for Madelung sums 50	114	I
NOBIPCOU	Coulomb bielectronic integrals computed exactly	116	–
NOBIPEXCH	Exchange bielectronic integrals computed exactly	116	–
NOBIPOLA	All bielectronic integrals computed exactly	116	–
POLEORDR	Maximum order of multipolar expansion 4	117	I
TOLINTEG	Truncation criteria for bielectronic integrals 6 6 6 6 12	125	I
TOLPSEUD	Pseudopotential tolerance 6	125	I

Type of run

ATOMHF	Atomic wave functions	93	I
SCFDIR	SCF direct (mono+biel int computed)	118	–
EIGS	S(**k**) eigenvalues - basis set linear dependence check	98	–
FIXINDEX	Reference geometry to classify integrals	104	–

Basis set - AO occupancy

FDAOSYM	f and d degeneracies analysis	101	I
FDAOCCUP	f and d orbital occupation guess	102	I
GUESDUAL	Density matrix guess - different Basis set	108	I

Integral file distribution

BIESPLIT	writing of bielectronic integrals in n files [$n = 1$],max=10	95	I
MONSPLIT	writing of mono-electronic integrals in n file [$n = 1$], max=10	115	I

Numerical accuracy control and convergence tools

ANDERSON	Fock matrix mixing	93	I
BROYDEN	Fock matrix mixing	95	I
DIIS	Fock matrix mixing/convergence accelerator	96	I
FMIXING	Fock/KS matrix (cycle i and i-1) mixing [0]	105	I
LEVSHIFT	level shifter [no]	113	I
MAXCYCLE	maximum number of cycles [50]	114	I
SMEAR	Finite temperature smearing of the Fermi surface [no]	121	I
TOLDEE	convergence on total energy [6]	125	I

Initial guess

EIGSHIFT	alteration of orbital occupation before SCF [no]	99	I
EIGSHROT	rotation of the reference frame [no]	99	I
GUESSP	density matrix from a previous run	109	–
GUESSPAT	superposition of atomic densities	110	–

Spin-polarized system

ATOMSPIN	setting of atomic spin to compute atomic densities	94	I
BETALOCK	beta electrons locking	94	I
SPINLOCK	spin difference locking	124	I
SPINEDIT	editing of the spin density matrix used as SCF guess	123	I

Auxiliary and control keywords

END	terminate processing of block3 input		–
FULLTIME	detailed report on running time	106	–
KSYMMPRT	printing of Bloch functions symmetry analysis	112	–
LOWMEM	inhibits allocation of large arrays	114	–
NOLOWMEM	allows allocation of large arrays	114	–
MAXNEIGHB	maximum number of equidistant neighbours from an atom	49	I
NEIGHBOR	number of neighbours to analyse in **PPAN**	59	I
MEMOPRT	Synthetic report about dynamic memory usage	114	–
MEMOPRT2	Detailed report about dynamic memory usage	114	–
PRINTOUT	setting of printing options	64	I
QVRSGDIM	maximum size of mutipole moment gradient array [90000000]	118	I
NOSYMADA	No Symmetry Adapted Bloch Functions	117	–
SYMADAPT	Symmetry Adapted Bloch Functions (default)	124	–
SETINF	setting of inf array options	66	I
SETPRINT	setting of printing options	66	I
STOP	execution stops immediately	68	–
TESTPDIM	stop after symmetry analysis	124	–
TEST[RUN]	stop after integrals classification and disk storage estimate	125	–

Restricted to *MPPcrystal*

CMPLXFAC	Overloading in handling matrices at "complex" k points with respect to "real" k points $\boxed{2.\overline{3}}$	96	I
REPLDATA	to run *MPPcrystal* as *Pcrystal*	118	–
STDIAG	Enable standard diagonalization method (D&C method disabled)	124	–

Output of data on external units

NOFMWF	wave function formatted output not written in file fort.98.	116	–
SAVEWF	wave function data written every two SCF cycles	119	–

Post SCF calculations

POSTSCF	post-scf calculations when convergence criteria not satisfied	117	–
EXCHGENE	exchange energy evaluation (spin polarized only)	101	–
GRADCAL	analytical gradient of the energy	106	–
PPAN	population analysis at the end of the SCF $\boxed{\text{no}}$	117	

Properties

RDFMWF wave function data conversion formatted-binary (fort.98 → fort.9)

Preliminary calculations

NEWK	Eigenvectors calculation	341	I
COMMENS	Density Matrix commensurate to the Monchorst net	311	I
NOSYMADA	No symmetry Adapted Bloch Functions	117	–
PATO	Density matrix as superposition of atomic (ionic) densities	343	I
PBAN	Band(s) projected density matrix (preliminary **NEWK**)	343	I
PGEOMW	Density matrix from geometrical weights (preliminary **NEWK**)	344	I
PDIDE	Energy range projected density matrix (preliminary **NEWK**)	344	I
PSCF	Restore SCF density matrix	350	–

Properties computed from the density matrix

ADFT	Atomic density functional correlation energy	302	I
BAND	Band structure	304	I
BIDIERD	Reciprocal form factors	305	I
CLAS	Electrostatic potential maps (point multipoles approximation)	311	I
ECHG	Charge density and charge density gradient - 2D grid	319	I
ECH3	Charge density - 3D grid	317	I
EDFT	Density functional correlation energy (HF wave function only)	320	I
EMDLDM	Electron momentum distribution (along a line)	321	I
EMDPDM	Electron momentum distribution (in a plane)	322	I
HIRSHCHG	Hirshfeld population analysis	111	I
KINETEMD	Kinetic tensor from electron momentum density	326	I
PMP2	MP2 correction to the Valence Density Matrix	344	
POLI	Atom and shell multipoles evaluation	345	I
POTM	Electrostatic potential - 2D grid	348	I
POT3	Electrostatic potential - 3D grid	346	I
POTC	Electrostatic properties	347	I
PPAN	Mulliken population analysis	117	
XFAC	X-ray structure factors	352	I

Properties computed from the density matrix (spin-polarized systems)

ANISOTRO	Hyperfine electron-nuclear spin tensor	303	I
HIRSHCHG	Hirshfeld spin population analysis	111	I
ISOTROPIC	Hyperfine electron-nuclear spin interaction - Fermi contact	325	I
POLSPIN	Atomic spin density multipoles	346	I

Properties computed from eigenvectors (after keyword **NEWK**)

ANBD	Printing of principal AO component of selected CO	302	I
BWIDTH	Printing of bandwidth	311	I
DOSS	Density of states	315	I
EMDL	Electron momentum distribution - line	321	I
EMDP	Electron momentum distribution - plane maps	324	I
PROF	Compton profiles and related quantities	349	I
BOLTZTRA	Transport Properties (electron conductivity, Seebeck) within the semiclassical Boltzmann theory	309	I
SPINCNTM	Spin Contamination	351	

New properties

SPOLBP	Spontaneous polarization (Berry phase approach)	357	–
SPOLWF	Spontaneous polarization (localized CO approach)	358	–
LOCALWF	Localization of Wannier functions	327	I
DIEL	Optical dielectric constant	313	I
ISO+POTC	Mössbauer isomer shift and quadrupolar effects	358	I
TOPO	Topological analysis of the electron density	361	I

Auxiliary and control keywords

ANGSTROM	Set input unit of measure to Ångstrom	35	–
BASISSET	Printing of basis set, Fock/KS, overlap and density matrices	305	–
BOHR	Set input unit of measure to bohr	38	–
CHARGED	Non-neutral cell allowed (**PATO**)	75	–
END	Terminate processing of properties input keywords		–
FRACTION	Set input unit of measure to fractional	47	–
MAPNET	Generation of coordinates of grid points on a plane	339	I
MAXNEIGHB	maximum number of equidistant neighbours from an atom	49	I
NEIGHBOR	Number of neighbours to analyse in **PPAN**	59	I
PRINTOUT	Setting of printing options	64	I
RAYCOV	Modification of atomic covalent radii	64	I
SETINF	Setting of inf array options	66	I
SETPRINT	Setting of printing options	66	I
STOP	Execution stops immediately	68	–
SYMMOPS	Printing of point symmetry operators	73	–

Info - Output of data on external units

ATOMIRR	Coordinates of the irreducible atoms in the cell	303	–
ATOMSYMM	Printing of point symmetry at the atomic positions	38	–
COORPRT	Coordinates of all the atoms in the cell	42	–
CRYAPI_OUT	geometry, BS, direct lattice information	312	–
CRYAPI_OUT	Reciprocal lattice information + eigenvalues	312	–
EXTPRT	Explicit structural/symmetry information	44	–
FMWF	Wave function formatted outputi in file fort.98. Section 14.10	324	–
INFOGUI	Generation of file with wf information for visualization	325	–
XML	generation of XML file for electron transport with WanT	**??**	–

Appendix C

Printing options

Extended printing can be obtained by entering the keywords **PRINTOUT** (page 64) or **SETPRINT** (page 66).
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 *EIGENALL* –	*EIGENVAL*	N
• 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	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					
! 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 D

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**	44
	OPTGEOM	34	optcxxx	Geometry input - opt cell [atoms] - See **EXTPRT**	44
		66	HESSOPT.DAT	Hessian matrix	
		68	OPTINFO.DAT	Information to restart optimization	
crystal	GAUSS98	92	GAUSSIAN.DAT	Input for GAUSS98	77
	FINDSYM	26	FINDSYM.DAT	data in crystallographic format - read by program findsym(IUCR)	42
	STRUCPRT	33	STRUC.INCOOR	Cell parameters, coordinates of atoms	68
	COORPRT	33	fort.33	Coordinates of the atoms in the cell	42
crystal	EXTPRT	34	fort.34	Geometry input	44
properties	PPAN	24	PPAN.DAT	Mulliken population analysis	117
properties	BAND	25	fort.25	Bands (Crgra2006)	304
		24	BAND.DAT	Bands data	304
	CLAS	25	fort.25	Classical potential	311
	DIEL	24	DIEL.DAT	Dielectric constant	313
	DOSS	25	fort.25	Density of states (IPLOT=1)	315
		24	DOSS.DAT	Density of states (IPLOT=2)	315
	ECHG	25	fort.25	Electronic charge density - 2D grid	319
		25	RHOLINE.DAT	Electronic charge density - 1 grid	319
	ECH3	31	—	Electronic charge density - 3D grid	317
			DENS_CUBE.DAT	Electronic charge density - 3D grid CUBE format	317
			SPIN_CUBE.DAT	Spin density - 3D grid CUBE format	317
	POT3		POT_CUBE.DAT	Electostatic potential - 3D grid CUBE format	346
	EMDLDM	25	fort.25	EMD along a line	321
		24	EMDLDM.DAT	EMD along a line	321
	EMDPDM	25	fort.25	EMD map on a plane	322
		94	3DEMDTOTAL.DAT	EMD map on a plane (3D format)	322
		65	3DEMDANISO.DAT	EMD-anisotropy map on a plane (3D format)	322
	EMDL	25	fort.25	EMD line (IPLOT=1)	321
		24	EMDL.DAT	EMD line(IPLOT=2)	321
	EMDP	25		EMD - 2D grid	443
	INFOGUI	32		Data for the graphical user interface	325
	IRSPEC		—	Infrared spectra	223
			IRSPEC.DAT	IR Absorbance and Reflectance	223
			IRREFR.DAT	IR Refractive index	223
			IRDIEL.DAT	IR Dielectric function	223
	RAMSPEC		—	Raman spectra	225
			RAMSPEC.DAT	Raman spectra for polycrystalline powder and single crystal	225
	POTC	24	POTC.DAT	Electrostatic potential V, Electric field, Electric field gradient	347
	POTM	25	fort.25	Electrostatic potential - 2D grid	348
	PROF	25	fort.25	Compton profile and related quantities (IPLOT=1)	349
		24	PROF.DAT	Compton profile and related quantities (IPLOT=2)	349

Please refer to the standard script for running CRYSTAL09 as to handle input/output file names. See:

http://www.crystal.unito.it/tutorials =¿ How to run

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 - EMDPDM - Isovalue maps

The value of the function chosen (classic electrostatic potential (CLAS), charge(+spin) density (ECHG), electrostatic potential (POTM), electron momentum density (EMDPDM)) 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,A4,2I5,3E12.5
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 - Fermi level can be drawn
           3 : open shell,   conducting system - Fermi level can be drawn

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** 14.9 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,A4,2I5,3E12.5
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.
NPUNTI records are written. Each records contains (FORMAT: 10E12.4)

```
p  (emdl(p,ipro), ipro=1,nprojections))
```

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** [253] which can be downloaded from:
`www.cmbi.kun.nl/ schaft/molden/molden.html`

STRUCPRT

The file STRUC.INCOOR 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        )
4        3F        ) cartesian components of the direct lattice vectors
5        3F        )
6        1I        } number of symmetry operators

                    For each symmetry operator 4 records:
7        3F        )
8        3F        ) symmetry operators matrices in cartesian coordinates
9        3F        )
10       3F        } cartesian components of the translation

n        1I        } number of atoms in the primitive cell (up to CRYSTAL14, irreducible atoms
only)
                    For each atom, 1 record:
n+1      I,3F      } conventional atomic number, cartesian coordinates of the atoms
2        3I        } space group number and initial number of symmetry operators if symmetry was not
modified, otherwise 0
```

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

Functions values computed at 3D grid of points by the keywords **ECH3** (page 317), **POT3** (page346), **GRID3D** (page325) are written according to two formats: . All data in atomic units.

1. Fortran unit 31 is written According to the following format. 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

2. Function data computed at 3D grid points are written according to GAUSSIAN CUBE format in files:

DENS_CUBE.DAT	charge density
SPIN_CUBE.DAT	spin density
POT_CUBE.DAT	electrostatic potential

INFOGUI

Fortran unit 32 is written through the keyword **INFOGUI** (page 325). 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
```

IRSPEC

IRSPEC in **FREQCALC** input block writes the files IRSPEC.DAT, IRREFR.DAT and IRDIEL.DAT.

IRSPEC.DAT contains: 1 column with frequency ν in cm^{-1}, 1 column with wavelength λ in nm, 4 columns for the 4 different models of absorbance A, 1 column for reflectance R along each inequivalent polarization direction.

IRREFR.DAT contains: 1 column with frequency ν in cm^{-1}, 1 column with wavelength λ in nm, 2 columns for Re(n) and Im(n) along each direction (n being the refractive index).

IRDIEL.DAT contains: 1 column with frequency ν in cm^{-1}, 1 column with wavelength λ in nm, 3 columns for Re(ϵ), Im(ϵ) and Im($1/\epsilon$) along each direction (ϵ being the dielectric function).

Suppose we have the following input block, for a compound with three inequivalent polarization directions:

```
. . . .
FREQCALC
INTENS
[options for INTENS]
DIELTENS or DIELISO
. . . .
[optional FREQCALC keywords]
. . . .
IRSPEC
END
ENDFREQ
```

The first two columns in the generated IRSPEC.DAT contain frequencies and wavelengths, columns from 3 to 6 the raw absorbance, the classical absorbance, the two Rayleigh scattering absorbances, and column 7-9 the reflectance curves for the three directions.

Suppose we want to plot the raw absorbance with respect to frequency. Once *gnuplot* is opened on the terminal (type gnuplot), it is sufficient to type

```
plot 'IRSPEC.DAT' using 1:3
```

where 1:3 stands for "first column assigned to x axis and third column to y axis". The plot of the raw absorbance appears on the screen and can be saved with the command

```
save 'name_plot'
```

If we want to save the plot as a Post-Script

```
set size 1.0, 0.6
set terminal postscript portrait enhanced mono dashed lw 1 "Helvetica" 14
set output "my-plot.ps"
replot
```

In a similar way, we can type

```
plot 'IRSPEC.DAT' using 1:7
```

in order to obtain the reflectance along the first polarization direction, and

```
plot 'IRSPEC.DAT' using 1:8
```

for the reflectance along the second direction.

Further details about these commands and manipulation of files at http://www.duke.edu/ hp-gavin/gnuplot.html and http://www.gnuplot.info/documentation.html.

RAMSPEC

RAMSPEC in **FREQCALC** input block writes the file RAMSPEC.DAT.

RAMSPEC.DAT contains: 1 column with frequency ν in cm^{-1}, 3 columns for intensities of polycrystalline powders (total intensity, parallel polarization, perpendicular polarization), 6 columns for spectra of single crystals (1 for each inequivalent polarization direction: xx, xy, xz, yy, yz, zz).

Suppose we want to plot the total Raman intensity of a polycrystalline powder with respect to frequency. Once *gnuplot* is opened on the terminal (type gnuplot), it is sufficient to type

```
plot 'RAMSPEC.DAT' using 1:2
```

where 1:2 stands for "first column assigned to x axis and second column to y axis". The plot of the total polycrystalline Raman intensity appears on the screen and can be saved with the command

```
save 'name_plot'
```

If we want to save the plot as a Post-Script

```
set size 1.0, 0.6
set terminal postscript portrait enhanced mono dashed lw 1 "Helvetica" 14
set output "my-plot.ps"
replot
```

In a similar way, we can type

```
plot 'RAMSPEC.DAT' using 1:5
```

in order to obtain the Raman spectrum of a single crystal along the first polarization direction, i.e. xx.

Further details about these commands and manipulation of files at http://www.duke.edu/ hp-gavin/gnuplot.html and http://www.gnuplot.info/documentation.html.

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** 14.11) 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 E

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:

$$\int d\mathbf{r}\ |f(\mathbf{r})|^2 \neq 1; \tag{E.1}$$

however, we can always define a related $f'(\mathbf{r})$, multiplying $f(\mathbf{r})$ by a constant N:

$$f'(\mathbf{r}) = Nf(\mathbf{r}), \tag{E.2}$$

such that:

$$\int d\mathbf{r}\ |f'(\mathbf{r})|^2 = 1. \tag{E.3}$$

$f'(\mathbf{r})$ is said to be a *normalized* function and N is its *Normalization Coefficient* (NC). Substituting eq. E.2 in E.3, we have:

$$N = \left(\int d\mathbf{r}\ |f(\mathbf{r})|^2 \right)^{-1/2}. \tag{E.4}$$

Gaussians: Product Theorem and Normalization

Let us define Gaussian functions as:

$$G(\alpha_i; \mathbf{r} - \mathbf{A}) = \exp(-\alpha_i(\mathbf{r} - \mathbf{A})^2), \tag{E.5}$$

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]

$$G(\alpha; \mathbf{r} - \mathbf{A})G(\beta; \mathbf{r} - \mathbf{B}) = \exp\left(-\frac{\alpha\beta}{\xi}|\mathbf{R}|^2\right) G(\xi; \mathbf{r} - \mathbf{P}); \tag{E.8}$$

[1] Let us prove the Gaussian product theorem:

$$\begin{aligned} G(\alpha; \mathbf{r} - \mathbf{A})G(\beta; \mathbf{r} - \mathbf{B}) &= \exp(-\alpha_i(\mathbf{r} - \mathbf{A})^2) \exp(-\alpha_j(\mathbf{r} - \mathbf{B})^2 \\ &= \exp\left(-\alpha(\mathbf{r}^2 + \mathbf{A}^2 + 2\mathbf{r}\mathbf{A}) - \beta(\mathbf{r}^2 + \mathbf{B}^2 + 2\mathbf{r}\mathbf{B})\right) \\ &= \exp\left[-\xi\left((\mathbf{r} - \mathbf{P})^2 + \mathbf{P}^2 - \frac{\alpha\mathbf{A}^2 + \beta\mathbf{B}^2}{\xi}\right)\right]. \end{aligned} \tag{E.6}$$

with:

$$\xi = \alpha + \beta, \tag{E.9}$$

$$\mathbf{P} = \frac{\alpha\mathbf{A} + \beta\mathbf{B}}{\xi}, \tag{E.10}$$

$$\mathbf{R} = \mathbf{A} - \mathbf{B}. \tag{E.11}$$

From eq. E.4, the NC of Gaussian functions, g_i, can be written as:

$$\begin{aligned} g_i &= \left(\int d\mathbf{r} \ (G(\alpha_i; \mathbf{r}))^2 \right)^{-1/2} \\ &= \left(\int d\mathbf{r} \, G(2\alpha_i; \mathbf{r}) \right)^{-1/2} \\ &= \left(\frac{\pi}{2\alpha_i} \right)^{-3/4}, \end{aligned} \tag{E.12}$$

where the Gaussian product theorem and the Gaussian integral [287] have been used. $G'(\alpha_i; \mathbf{r})$, defined as:

$$G'(\alpha_i; \mathbf{r}) = g_i \, G(\alpha_i; \mathbf{r}), \tag{E.13}$$

is a normalized function.

Harmonic Gaussians

The Definition

The *Solid Harmonic Functions*, Y_ℓ^m, [245] are defined as:

$$Y_\ell^m(\mathbf{r}) = r^\ell P_\ell^{|m|}(cos\vartheta) e^{\mathrm{i}m\phi}, \tag{E.14}$$

where P_ℓ^m is the *Legendre Polynomial Function* characterized by the integers ℓ and m, such that: $\ell \geq 0$ and $-\ell \leq m \leq \ell$. [3]
Starting from Y_ℓ^m, the *Real Solid Harmonic*, X_ℓ^m, can be defined:

$$X_\ell^{|m|}(\mathbf{r}) \quad = \Re(Y_\ell^{|m|}) = \quad \frac{Y_\ell^{|m|}(\mathbf{r}) + Y_\ell^{-|m|}(\mathbf{r})}{2}, \tag{E.15}$$

$$X_\ell^{-|m|}(\mathbf{r}) \quad = \Im(Y_\ell^{|m|}) = \quad \frac{Y_\ell^{|m|}(\mathbf{r}) - Y_\ell^{-|m|}(\mathbf{r})}{2i}. \tag{E.16}$$

We report some examples of X functions.
$\ell = 0$:

$$X_0^0(\mathbf{r}) = 1; \tag{E.17}$$

$\ell = 1$:

$$X_1^0(\mathbf{r}) = z, \quad X_1^1(\mathbf{r}) = x, \quad X_1^{-1}(\mathbf{r}) = y; \tag{E.18}$$

$\ell = 2$:

$$X_2^0(\mathbf{r}) = z^2 - 0.5(x^2 - y^2), \quad X_2^1(\mathbf{r}) = 3zx, \quad X_2^{-1}(\mathbf{r}) = 3zy, \tag{E.19}$$

$$X_2^2(\mathbf{r}) = 3(x^2 + y^2), \quad X_2^{-2}(\mathbf{r}) = 3xy. \tag{E.20}$$

Using eqs. E.9, E.10 and E.11, eq. E.6 can be rewritten as:

$$G(\alpha; \mathbf{r} - \mathbf{A}) G(\beta; \mathbf{r} - \mathbf{B}) = \exp\left(-\frac{\alpha\beta}{\xi} |\mathbf{R}|^2 \right) G(\xi; \mathbf{r} - \mathbf{P}). \tag{E.7}$$

We have now the tools required to define the *Solid Harmonic Gaussian*, [245] ξ:

$$\xi^{n\ell m}(\alpha_i;\mathbf{r}) = |\mathbf{r}|^{2n} Y_\ell^m(\mathbf{r}) G_i(\alpha_i;\mathbf{r}), \tag{E.21}$$

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):

$$\xi^{\ell m}(\alpha_i;\mathbf{r}) = Y_\ell^m(\mathbf{r}) G(\alpha_i;\mathbf{r}). \tag{E.22}$$

Substituting Y with X (eqs. E.15 and E.16) in eq. E.22, *Real Harmonic Gaussians*, γ, can be defined:

$$\gamma^{\ell m}(\alpha_i;\mathbf{r}) = X_\ell^m(\mathbf{r}) G(\alpha_i;\mathbf{r}). \tag{E.23}$$

γ are used as basis functions in the CRYSTAL program and are related to the ξ ones by followings relations:

$$\gamma^{\ell|m|} = \frac{\xi^{\ell|m|} + \xi^{\ell\,-|m|}}{2}, \tag{E.24}$$

$$\gamma^{\ell-|m|} = \frac{\xi^{\ell|m|} - \xi^{\ell\,-|m|}}{2i}, \tag{E.25}$$

where eqs. E.15 and E.16 have been used.
Note that, when ℓ is equal to 0, ξ and γ functions degenerate to simple Gaussians:

$$\xi^{00} = \gamma^{00} = G, \tag{E.26}$$

where eq. E.17 has been used and ξ degenerates to γ when $m = 0$:

$$\xi^{\ell 0} = \gamma^{\ell 0}, \tag{E.27}$$

where eqs. E.24 and E.25 have been used.

The Normalization Coefficient

Let us consider now ξ and γ's normalization coefficients (b and c, respectively), from eq. E.4, follows:

$$b_i^{\ell m} = (\Xi)^{-1/2} \tag{E.28}$$

$$c_i^{\ell m} = (\Upsilon)^{-1/2}, \tag{E.29}$$

where

$$\Xi = \int d\mathbf{r} \; \left|\xi^{\ell m}(\alpha_i;\mathbf{r})\right|^2 \tag{E.30}$$

$$\Upsilon = \int d\mathbf{r} \; \left(\gamma^{\ell m}(\alpha_i;\mathbf{r})\right)^2. \tag{E.31}$$

Using eqs. E.5, E.8, E.14, E.22 and a spherical polar coordinate system,[2] the Ξ integral can be factorized as:

$$\begin{aligned}\Xi &= \int d\mathbf{r} \; [Y_\ell^m(\mathbf{r}) G(\alpha_i;\mathbf{r})]^* \, Y_\ell^m(\mathbf{r}) G(\alpha_i;\mathbf{r}) \\ &= \int d\mathbf{r} \, Y_\ell^{-m}(\mathbf{r}) \, Y_\ell^m(\mathbf{r}) \, G(2\alpha_i;\mathbf{r}) \\ &= \Xi_r \; \Xi_\vartheta \; \Xi_\phi,\end{aligned} \tag{E.32}$$

[2] $d\mathbf{r} = r^2 sin\vartheta \, dr \, d\vartheta \, d\phi$

with:

$$\begin{aligned}\Xi_r &= \int_0^\infty dr\ \exp(-2\alpha_i r^2) r^{2\ell+2} \\ &= \frac{\Gamma(\ell+3/2)}{2(2\alpha_i)^{\ell+3/2}} \\ &= \frac{\pi^{1/2}(2\ell+1)!!}{2^{\ell+2}(2\alpha_i)^{\ell+3/2}},\end{aligned} \tag{E.33}$$

where we have used the Γ function's properties; [3]

$$\begin{aligned}\Xi_\vartheta &= \int_0^\pi d\vartheta\ \left(P_\ell^{|m|}(\cos\vartheta)\right)^2 \sin\vartheta \\ &= \frac{2(\ell+|m|)!}{(2\ell+1)(\ell-|m|)!},\end{aligned} \tag{E.34}$$

where the Legendre polynomials' properties have been used, [3] and

$$\Xi_\phi = \int_0^{2\pi} d\phi = 2\pi. \tag{E.35}$$

Substituting eqs. E.32, E.33, E.34 and E.35 in the b definition (eq. E.28) we obtain:

$$\begin{aligned}b_i^{\ell m} &= \frac{\pi^{1/2}(2\ell+1)!!}{2^{\ell+2}(2\alpha_i)^{\ell+3/2}} \frac{2(\ell+|m|)!}{(2\ell+1)(\ell-|m|)!} 2\pi \\ &= \left(\frac{\pi^{3/2}\,(2\ell-1)!!\,(\ell+|m|)!}{2^{2\ell+3/2}\,\alpha_i^{\ell+3/2}\,(\ell-|m|)!}\right)^{-1/2}.\end{aligned} \tag{E.36}$$

Note that b is independent from the sign of m (as Ξ is), that is:

$$b_i^{\ell|m|} = b_i^{\ell\,-|m|}. \tag{E.37}$$

In order to deduce the explicit expression for c, we are interested now in solving the integral of eq. E.31:

$$\Upsilon = \int d\mathbf{r}\ \left(X_\ell^m(\mathbf{r})\right)^2\ G(2\alpha_i;\mathbf{r}), \tag{E.38}$$

where eqs. E.8 and E.23 have been used. Substituting eq. E.24 (γ functions with $m \geq 0$) in previous equation, we have:

$$\begin{aligned}\Upsilon^{m\geq 0} = \frac{1}{4}\Bigg(&\int d\mathbf{r}\ \left|\xi^{\ell|m|}(\alpha_i;\mathbf{r})\right|^2 + \int d\mathbf{r}\ \left|\xi^{\ell\,-|m|}(\alpha_i;\mathbf{r})\right|^2 + \\ &+ 2\int d\mathbf{r}\,\xi^{\ell|m|}(\alpha_i;\mathbf{r})\xi^{\ell\,-|m|}(\alpha_i;\mathbf{r})\Bigg).\end{aligned} \tag{E.39}$$

The first two integrals in eq. E.39 can be recognized as Ξ (eq. E.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, [3] therefore:

$$\Upsilon^{m>0} = \frac{\Xi}{2}. \tag{E.40}$$

The same result is found for negative m, substituting eq. E.25 (instead of eq. E.24, as done) in eq. E.31:

$$\Upsilon^{m<0} = \frac{\Xi}{2}, \tag{E.41}$$

so Υ (as Ξ is) is independent from the m sign. If $m = 0$, the last integral in eq. E.39 is equal to Ξ, as the first two ones:

$$\Upsilon^{m=0} = \Xi; \tag{E.42}$$

the previous equation can be deduced also from eq. E.27.
Summarizing, from eqs. E.40, E.41 and E.42, we get:

$$\Upsilon = \frac{\Xi}{2 - \delta_{m0}} \tag{E.43}$$

and, finally, substituting eqs. E.32 and E.43 in eq. E.29, we obtain:

$$c_i^{\ell m} = \left(\frac{\pi^{3/2}\,(2\ell-1)!!\,(\ell+|m|)!}{2^{2\ell+3/2}(2-\delta_{m0})\,\alpha_i^{\ell+3/2}\,(\ell-|m|)!} \right)^{-1/2}. \tag{E.44}$$

The c expression (eq. E.44) can be reorganized in a two factors formula:

$$c_i^{\ell m} = a_i^{\ell}\, f^{\ell m}, \tag{E.45}$$

with:

$$a_i^{\ell} = \left(\frac{\pi^{3/2}}{(2\alpha_i)^{\ell+3/2}} \right)^{-1/2}, \tag{E.46}$$

the α-dependent term, and

$$f^{\ell m} = \left(\frac{(2\ell-1)!!\,(\ell+|m|)!}{2^{\ell}(2-\delta_{m,0})(\ell-|m|)!} \right)^{-1/2}, \tag{E.47}$$

the m dependent term.
Note that,

- If $\ell = 0$, γ degenerates in a simple Gaussian (eq. E.26),

$$f^{00} = 1 \quad \text{and} \quad c_i^{00} = a_i^0 = g_i, \tag{E.48}$$

  where g_i is the G's NC (eq. E.12).

- If $\ell = 1$, $f^{1m} = 1/2$ for the three m-values:

$$f^{1m} = 1/2 \quad \text{and} \quad c_i^{1m} = \frac{a_i^1}{2} = \frac{\alpha_i^{5/4} 2^{7/4}}{\pi^{3/4}},\ \forall\ m = -1, 0, 1. \tag{E.49}$$

- If $\ell = 2$, we have:

$$c_i^{20} = \frac{\alpha_i^{7/4} 2^{11/4}}{\pi^{3/4}\sqrt{3}}; \quad c_i^{21} = c_i^{2-1} = \frac{\alpha_i^{7/4} 2^{11/4}}{\pi^{3/4} 3}; \quad c_i^{22} = c_i^{2-2} = \frac{\alpha_i^{7/4} 2^{7/4}}{\pi^{3/4} 3}. \tag{E.50}$$

Let us verify, for two examples, that

$$\gamma' = c\,\gamma \tag{E.51}$$

is a normalized function, proving that the following integral, I, is equal to 1,

$$I_i^{\ell m} = \int d\mathbf{r}\ \left(c_i^{\ell m} \gamma^{\ell m}(\alpha_i; \mathbf{r}) \right)^2. \tag{E.52}$$

The s Case ($\ell = 0$, $m = 0$)

$$\begin{aligned} I_i^{00} &= \left(c_i^{00}\right)^2 \int d\mathbf{r} \, \left(\gamma^{00}(\alpha_i;\mathbf{r})\right)^2 \\ &= (g_i)^2 \int d\mathbf{r} \, \left(G(\alpha_i;\mathbf{r})\right)^2 \\ &= \left(\int \left(G(\alpha_i;\mathbf{r})\right)^2 d\mathbf{r}\right)^{-1} \int \left(G(\alpha_i;\mathbf{r})\right)^2 d\mathbf{r} = 1, \end{aligned} \tag{E.53}$$

where eqs. E.48, E.27 and E.12 have been used.

A d Case ($\ell = 2$, $m = 1$)

$$I_i^{21} = \int d\mathbf{r} \, \left(c_i^{21}\gamma^{21}(\alpha_i;\mathbf{r})\right)^2 = \left(c_i^{21}\right)^2 J, \tag{E.54}$$

with:

$$J = \int d\mathbf{r} \, \left(3zxG(2\alpha_i;\mathbf{r})\right)^2, \tag{E.55}$$

where eqs. E.23 and E.19 have been used.
Gaussians are separable functions, that is:

$$G(\alpha_i;\mathbf{r}) = G_x(\alpha_i;x)\, G_y(\alpha_i;y)\, G_z(\alpha_i;z), \tag{E.56}$$

with:

$$G_x(\alpha_i;x) = \exp(-\alpha_i x^2) \tag{E.57}$$

and similarly for y and z. Substituting eq. E.56 in eq. E.55, we have:

$$J = 9 J_x J_y J_z, \tag{E.58}$$

with:

$$J_x = \int x^2 G_x(2\alpha_i;x)dx = \frac{\sqrt{\pi}}{2}(2\alpha_i)^{-3/2}, \tag{E.59}$$

$$J_y = \int G_y(2\alpha_i;y)dy = \left(\frac{\pi}{2\alpha_i}\right)^{1/2}, \tag{E.60}$$

$$J_z = \int z^2 G_z(2\alpha_i;z)dz = \frac{\sqrt{\pi}}{2}(2\alpha_i)^{-3/2}, \tag{E.61}$$

where ref. [287] has been used in solving the integrals. Substituting now eqs. E.49 and E.58 in eq. E.54, we obtain:

$$I_i^{21} = \frac{\alpha_i^{7/2}\, 2^{11/2}}{\pi^{3/2}\, 9} \; 9 \left(\frac{\sqrt{\pi}}{2}(2\alpha_i)^{-3/2}\right)^2 \left(\frac{\pi}{2\alpha_i}\right)^{1/2} = 1. \tag{E.62}$$

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. E.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:

$$\varphi_\lambda^{\ell m} = N_\lambda \sum_i d_i^\lambda \, c_i^{\ell m} \, \gamma^{\ell m}(\alpha_i^\lambda;\mathbf{r}) = N_\lambda \sum_i d_i^\lambda \, c_i^{\ell m} \, X_\ell^m(\mathbf{r})\, G(\alpha_i^\lambda;\mathbf{r}). \tag{E.63}$$

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. E.29), respectively. N_λ is the NC of AOs belonging to λ and is defined as:

$$N_\lambda = \left(\int d\mathbf{r} \left(\sum_i d_i^\lambda \, c_i^{\ell m} \, \gamma^{\ell m}(\alpha_i^\lambda; \mathbf{r}) \right)^2 \right)^{-1/2}, \tag{E.64}$$

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.

$$p_z = \varphi_4^{10} = N^4 \left(d_1^4 \, c_1^{10} \, \gamma^{10}(\alpha_1^4; \mathbf{r}) + d_2^4 \, c_2^{10} \, \gamma^{10}(\alpha_2^4; \mathbf{r}) \right), \tag{E.65}$$

$$p_x = \varphi_4^{11} = N^4 \left(d_1^4 \, c_1^{11} \, \gamma^{11}(\alpha_1^4; \mathbf{r}) + d_2^4 \, c_2^{11} \, \gamma^{11}(\alpha_2^4; \mathbf{r}) \right), \tag{E.66}$$

$$p_y = \varphi_4^{1-1} = N^4 \left(d_1^4 \, c_1^{1-1} \, \gamma^{1-1}(\alpha_1^4; \mathbf{r}) + d_2^4 \, c_2^{1-1} \, \gamma^{1-1}(\alpha_2^4; \mathbf{r}) \right). \tag{E.67}$$

Let us put our attention on N_λ. Eq. E.64 can be rewritten as:

$$N^\lambda = \left(\sum_{i,j} d_i^\lambda \, d_j^\lambda \, c_i^{\ell m} c_j^{\ell m} \Upsilon' \right)^{-1/2}, \tag{E.68}$$

with:

$$\Upsilon' = \int d\mathbf{r} \, \gamma^{\ell m}(\alpha_i; \mathbf{r}) \, \gamma^{\ell m}(\alpha_j; \mathbf{r}), \tag{E.69}$$

where the shell index on α has been omitted for simplicity. Substituting eq. E.23 in eq. E.69, we have:

$$\Upsilon' = \int X_\ell^m(\mathbf{r}) \, G(\alpha_i; \mathbf{r}) \, X_\ell^m(\mathbf{r}) \, G(\alpha_j; \mathbf{r}) d\mathbf{r} = \int \left(X_\ell^m(\mathbf{r}) \right)^2 G[(\alpha_i + \alpha_j); \mathbf{r}] d\mathbf{r}, \tag{E.70}$$

where the Gaussian product theorem (eq. E.8) has been used.
From eq. E.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. E.43, E.32, E.34 and E.35, Υ' is rewritten as:

$$\Upsilon' = \frac{\Upsilon'_r \, \Xi_\vartheta \, \Xi_\varphi}{2 - \delta_{m0}}, \tag{E.71}$$

with:

$$\begin{aligned} \Upsilon'_r &= \int_0^\infty dr \, \exp[-(\alpha_i + \alpha_j) r^2] r^{2\ell+2} \\ &= \frac{\Gamma(\ell + 3/2)}{2(\alpha_i + \alpha_j)^{\ell+3/2}} \\ &= \frac{\pi^{1/2} (2\ell + 1)!!}{2^{\ell+2} (\alpha_i + \alpha_j)^{\ell+3/2}}. \end{aligned} \tag{E.72}$$

Substituting eqs. E.44, E.71 and E.72 in eq. E.68, we obtain:

$$N_\lambda = \left(\sum_{i,j} d_i^\lambda \, d_j^\lambda \, \left(\frac{2\sqrt{\alpha_i^\lambda \alpha_j^\lambda}}{\alpha_i^\lambda + \alpha_j^\lambda} \right)^{\ell+3/2} \right)^{-1/2}, \tag{E.73}$$

where it is clear that N depends only on λ.

The Code

In order to explain easily the organization of NCs in CRYSTAL, eq. E.63 is reorganized as follows:

$$\varphi_\lambda^{\ell m} = \sum_i n_{\lambda,i}^{\ell m}\, \gamma^{\ell m}(\alpha_i^\lambda; \mathbf{r}), \tag{E.74}$$

with:

$$n_{\lambda,i}^{\ell m} = N_\lambda\, d_i^\lambda\, c_i^{\ell m}. \tag{E.75}$$

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 $\pi^{5/8}\, 2^{1/4}$; [224] therefore, in the code, pre-Gaussian factors are not stored, but the following quantities, that we shall call *code pre-Gaussian constants*:

$$\mathsf{S}_i^\lambda = \pi^{5/8}\, 2^{1/4}\, n_{\lambda,i}^{00} \tag{E.76}$$

$$\mathsf{P}_i^\lambda = \pi^{5/8}\, 2^{1/4}\, n_{\lambda,i}^{1m} \;\forall\; m = 0, 1, -1 \tag{E.77}$$

$$\mathsf{D}_i^\lambda = \pi^{5/8}\, 2^{1/4} \sqrt{\frac{(2+|m|)!}{(2-\delta_{m0})(2-|m|)!}}\, n_{\lambda,i}^{2m} \;\forall\; m = 0, 1, -1, 2, -2. \tag{E.78}$$

Note that the square root in eq. E.78 (the inverse of the m-dependent part of c, eq. E.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 P'^λ_i and D'^λ_i coefficients, respectively:

$$\mathsf{P}'^\lambda_i = \frac{\pi^{5/8}\, 2^{1/4}}{2\alpha_i}\, n_{\lambda,i}^{1m} \;\forall\; m = 0, 1, -1 \tag{E.79}$$

$$\mathsf{D}'^\lambda_i = \frac{\pi^{5/8}\, 2^{1/4}}{(2\alpha_i)^2} \sqrt{\frac{(2+|m|)!}{(2-\delta_{m0})(2-|m|)!}}\, n_{\lambda,i}^{2m} \;\forall\; m = 0, 1, -1, 2, -2. \tag{E.80}$$

[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:

$$S_{\mu\nu} = \int d\mathbf{r}\, \varphi^\lambda_{00}(\mathbf{r})\varphi^\sigma_{20}(\mathbf{r}). \tag{E.81}$$

Substituting eq. E.74 in the previous equation, we have:

$$S_{\mu\nu} = \sum_{ij} n^{00}_{\lambda,i}\, n^{20}_{\sigma,j} \int d\mathbf{r}\, \gamma^{00}(\alpha^\lambda_i;\mathbf{r})\gamma^{20}(\alpha^\sigma_j;\mathbf{r}). \tag{E.82}$$

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. E.76 and E.78, and we rewrite the overlap integral as:

$$S_{\mu\nu} = \left(\pi^{5/8}\, 2^{1/4}\right)^{-2} \sqrt{\frac{(2-\delta_{m0})(2-|m|)!}{(2+|m|)!}} \sum_{ij} \mathsf{S}^\lambda_i \mathsf{D}^\sigma_i \int d\mathbf{r}\, \gamma^{00}(\alpha^\lambda_i;\mathbf{r})\gamma^{20}(\alpha^\sigma_j;\mathbf{r}). \tag{E.83}$$

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 F

CRYSTAL09 versus CRYSTAL06

Geometry

- Roto-translational symmetry
  In the case of polymers it can treat helical structures (translation followed by a rotation around the periodic axis). See keyword **HELIX** (page 16) and examples therein. CRYSTAL06 allowed commensurate rotations only, by adopting a suitably large unit cell.

- Nanotubes
  A special input option allows generation of 1D structures (nanotubes) from 2D one. See keyword **NANOTUBE** (page 54).

Geometry optimization

- Default choice modified: TRUSTRADIUS scheme active.

  Use keyword **NOTRUSTR** in **OPTGEOM** input block (page 188) to run geometry optimization with CRYSTAL06 default.

Frequencies calculation

- default value for SCF convergence on total energy is 10^{-9}

- default choice for numerical integration grid (DFT Hamiltonian): **XLGRID**. To run DFT Hamiltonian cases with CRYSTAL06 numerical integration accuracy insert **LGRID** in DFT input block (page 140)

- default choice to compute IR intensities: Berry phase approach. To compute IR intensities with CRYSTAL09 as with CRYSTAL06 insert keyword **INTLOC** in **FREQCALC** input block.

Basis set

- f orbitals
  - f orbitals (local functions basis set) with non-zero occupancy allowed. This new feature allows study of systems with Lanthanides.
  d and f orbitals occupation guess
  - **FDOCCUP** (input block3, page 102) defines the occupation of specific f or d orbitals

in a given shell in the SCF initial guess, according to the local atomic symmetry in the crystal lattice.

- - Effective core pseudo potentials Projector operators up to angular quantum number $\ell = 4$ are allowed. ***Input deck has been changed***

Test cases
crystal09 versus crystal06 total energies (hartree)

	CRYSTAL06	CRYSTAL09	diff
TEST 0	-110.7649354541	-110.7649354541	0.00E+00
TEST 1	-39.7267242374	-39.7267242374	0.00E+00
TEST 2	-223.7874756819	-223.7874756819	0.00E+00
TEST 3	-893.8746580004	-893.8746580004	0.00E+00
TEST 4	-1400.1776585535	-1400.1776585535	0.00E+00
TEST 5	-74.8333583570	-74.8333583570	1.01E-12
TEST 6	-58.4208255980	-58.4208255980	1.00E-12
TEST 7	-2800.7355953744	-2800.7355953744	0.00E+00
TEST 8	-571.3207540595	-571.3207540595	3.00E-11
TEST 9	-29.2566111179	-29.2566111179	0.00E+00
TEST10	-577.8265583253	-577.8265583271	-1.86E-09
TEST11	-274.6641530559	-274.6641530559	-3.00E-11
TEST12	-447.6810664796	-447.6810664796	-6.00E-11
TEST13	-23.9856901143	-23.9856901143	0.00E+00
TEST14	-159.6970601598	-159.6970601598	0.00E+00
TEST15	-5229.8366027793	-5229.8366027783	1.00E-09
TEST16	-2995.2869386583	-2995.2869386582	1.00E-10
TEST17	-2674.3752958019	-2674.3752958111	-9.20E-09
TEST18	-679.2766564082	-679.2766564082	0.00E+00
TEST19	-223.8070777853	-223.8070777853	0.00E+00
TEST20	-89.9552981103	-89.9552981101	1.79E-10
TEST21	-447.5749511978	-447.5749511978	0.00E+00
TEST22	-460.7186326563	-460.7186326563	-3.00E-11
TEST23	-8.0613160317	-8.0613160317	0.00E+00
TEST24	-1400.1776188146	-1400.1776188146	0.00E+00
TEST25	-74.8442039913	-74.8442039913	-9.95E-13
TEST26	-58.4208255860	-58.4208255860	0.00E+00
TEST27	-2800.7355409839	-2800.7355409839	0.00E+00
TEST28	-8.0115274157	-8.0115274157	-9.95E-14
TEST29	-2047.6430862971	-2047.6430862970	9.98E-11
TEST30	-109.0441458665	-109.0441458665	0.00E+00
TEST31	-4095.2867581787	-4095.2867581681	1.06E-08
TEST32	-92.1408103960	-92.1408103960	0.00E+00
TEST33	-92.1416129818	-92.1416129818	0.00E+00
TEST34	-1117.5230436113	-1117.5230436113	0.00E+00
TEST35	-936.5017511475	-936.5017511475	3.00E-11
TEST36	-112.5648952230	-112.5648952230	1.00E-11
TEST37	-3028.3682392877	-3028.3682392877	0.00E+00
TEST38	-2279.1395902376	-2279.1395902366	1.00E-09

Appendix G

CRYSTAL14 versus CRYSTAL09

Old Settings

- Some default computational parameters have changed with respect to the CRYSTAL09 version of the program (see below). The keyword **OLDREF09**, to be inserted in the geometry input block, switches back them all;

Geometry optimization

- Full geometry optimizations (atomic coordinates and lattice parameters) are now performed as a default option when the **OPTGEOM** keyword is used. The sub-keyword **ATOMONLY** switches back to an atomic positions only optimization;
- The **FINALRUN** = 4 option is now set by default (before it was 0). See page 189 for details;

Density Functional Theory

- The size of the default numerical integration grid has changed. Now it corresponds to the **XLGRID** option. The option **OLDGRID** has been added to set back the old grid size;
- By default, an unlocked energy shifting of 0.6 hartree is applied to separate apart occupied from virtual orbitals, which corresponds to option **LEVSHIFT** with parameters 6 0.

Frequencies calculation

- Eckart conditions for cleaning the Hessian matrix as regards translational and rotational vibration modes are now activated by default. See page 210 for details;

SCF Parameters

- A Fock (Kohn-Sham) matrix mixing of 30 % between subsequent SCF cycles is now active by default (see keyword **FMIXING**);
- A full direct approach for the computation of the integrals (keyword **SCFDIR**) is now used as a default. Use keyword **NODIRECT** for switching this option off;
- The tresholds governing the bipolar approximation have changed from 14 10 to 18 14. See keyword **BIPOLA** for details;

Test cases
CRYSTAL14 versus CRYSTAL09 total energies (hartree)

	CRYSTAL09 v.2.0.1 Etot	CRYSTAL14 Old settings Etot	CRYSTAL14 New defaults Etot	CRY14-CRY09 Old settings DEtot
test00	-110.76493545	-110.76493545	-110.76493545	0.00000000
test01	-39.72672424	-39.72672424	-39.72672424	0.00000000
test02	-223.78747568	-223.78747568	-223.78754817	0.00000000
test03	-893.87465800	-893.87465800	-893.87461008	0.00000000
test04	-1400.17765855	-1400.17765855	-1400.17840956	0.00000000
test05	-74.83335836	-74.83335836	-74.83336244	0.00000000
test06	-58.42082558	-58.42082558	-58.42098571	0.00000000
test07	-2800.73559537	-2800.73559537	-2800.73725678	0.00000000
test08	-571.32075406	-571.32075406	-571.32081226	0.00000000
test09	-29.25661159	-29.25661159	-29.25662600	0.00000000
test10	-577.82655833	-577.82655832	-577.82670817	0.00000000
test11	-274.66415306	-274.66415306	-274.66419189	0.00000000
test12	-447.68106648	-447.68106648	-447.68124595	0.00000000
test13	-23.98569011	-23.98569011	-23.98569013	0.00000000
test14	-159.69706016	-159.69706016	-159.69729413	0.00000000
test15	-5229.83660278	-5229.83660278	-5229.83555014	0.00000000
test16	-2995.28693866	-2995.28693866	-2995.28683802	0.00000000
test17	-2674.37529581	-2674.37529581	-2674.37559033	0.00000000
test18	-679.27665641	-679.27665641	-679.27667659	0.00000000
test19	-223.80707779	-223.80707779	-223.80715189	0.00000000
test20	-89.95529811	-89.95529811	-89.95529811	0.00000000
test21	-447.57495120	-447.57495120	-447.57509617	0.00000000
test22	-460.71863266	-460.71863266	-460.71872430	0.00000000
test23	-8.06131603	-8.06131603	-8.06132051	0.00000000
test24	-1400.17761881	-1400.17761881	-1400.17836984	0.00000000
test25	-74.84420399	-74.84420399	-74.84419555	0.00000000
test26	-58.42082557	-58.42082557	-58.42098569	0.00000000
test27	-2800.73554098	-2800.73554098	-2800.73720223	0.00000000
test28	-8.01152742	-8.01152742	-8.01152780	0.00000000
test29	-2047.64308630	-2047.64308630	-2047.64342093	0.00000000
test30	-109.04414587	-109.04414587	-109.04434229	0.00000000
test31	-4095.28675817	-4095.28675818	-4095.28742264	-0.00000001
test32	-92.14081040	-92.14081040	-92.14081024	0.00000000
test33	-92.14161298	-92.14161298	-92.14161286	0.00000000
test34	-1117.52304361	-1117.52304361	-1117.52300498	0.00000000
test35	-936.50175115	-936.50175115	-936.50185912	0.00000000
test36	-112.56489522	-112.56489522	-112.56490316	0.00000000
test37	-3028.36823929	-3028.36823929	-3028.36857687	0.00000000
test38	-2279.13959024	-2279.13959024	-2279.14018803	0.00000000
test43	-1613.25523877	-1613.25523877	-1613.25545994	0.00000000
test47	-4394.62644451	-4394.62644451	-4394.62706583	0.00000000
test01_dft	-40.32096680	-40.32096680	-40.32090900	0.00000000

test02_dft	-224.92502981	-224.92502981	-224.92495090	0.00000000
test03_dft	-895.51814784	-895.51814784	-895.51821309	0.00000000
test04_dft	-1403.86073017	-1403.86073017	-1403.86074899	0.00000000
test05_dft	-75.25763976	-75.25763976	-75.25766593	0.00000000
test08_dft	-573.30069771	-573.30069771	-573.30058383	0.00000000
test11_dft	-275.43123078	-275.43123078	-275.43126854	0.00000000
test39_dft	-485.38270022	-485.38270022	-485.38264752	0.00000000
test40_dft	-337.10879536	-337.10879536	-337.10946296	0.00000000
test41_dft	-1313.33749919	-1313.33749919	-1313.33738975	0.00000000
test42_dft	-7.53085315	-7.53085315	-7.53085381	0.00000000
test44_dft	-1421.65733745	-1421.65733745	-1421.65785507	0.00000000
test45_dft	-5280.86118400	-5280.86118400	-5280.86080258	0.00000000
test46_dft	-5280.43937642	-5280.43937642	-5280.43953416	0.00000000
test47_dft	-4406.89968860	-4406.89968860	-4406.90099322	0.00000000
test48_dft	-37.61966054	-37.61966054	-37.61967775	0.00000000
test49_dft	-37.61966059	-37.61966059	-37.61967828	0.00000000
test50_dft	-37.61966059	-37.61966059	-37.61967828	0.00000000
test51_dft	-1020.29088144	-1020.29088144	-1020.29350751	0.00000000

Appendix H

CRYSTAL17 versus CRYSTAL14

Old Settings

- Some default computational parameters have changed with respect to the CRYSTAL14 version of the program (see below). The keyword **OLDREF14**, to be inserted in the geometry input block, switches back them all;

SCF Parameters

- The DIIS scheme for convergence acceleration is now active by default for all SCF and CPHF calculations. The keyword **NODIIS**, to be inserted in the third input block for SCF and within the CPHF input block for CPHF, switches it off.

Density Functional Theory

- A new default value of 200 is set for the **CHUNKS** option, which enables a better load balance for DFT numerical integration in replicated-data parallel calculations (PCRYSTAL);

Geometry Optimization

- New system-dependent default values for the **FIXDELTE**, **FIXDELTX** and **FIXDELTN** options are set, which control the updating of the indexes for integral evaluation during the optimization process;

Equation-of-State

- The "Trust-radius" strategy for step control during the constant-volume optimizations has been turned off by default;
- The pre-screening for the integrals to be computed is now performed on the most compressed rather than on the equilibrium configuration;

Strain-related Properties

- The "Trust-radius" strategy for step control during the geometry optimizations for the calculation of the nuclear-relaxed term of the elastic, piezoelectric and photo-elastic tensors has been turned off by default;

Test cases
CRYSTAL17 versus CRYSTAL14 total energies (hartree)

	CRYSTAL14 [default] Etot	CRYSTAL17 [Old settings] Etot	CRY17-CRY14 [Old settings] DEtot	CRYSTAL17 [New defaults] Etot
test00	-110.76493545	-110.76493545	0.00000000	-110.76493545
test01	-39.72672424	-39.72672424	0.00000000	-39.72672424
test02	-223.78754817	-223.78754817	0.00000000	-223.78754821
test03	-893.87461008	-893.87461008	0.00000000	-893.87461011
test04	-1400.17840956	-1400.17840956	0.00000000	-1400.17840968
test05	-74.83336244	-74.83336244	0.00000000	-74.83336244
test06	-58.42098571	-58.42098571	0.00000000	-58.42082558
test07	-2800.73725678	-2800.73725678	0.00000000	-2800.73725684
test08	-571.32081226	-571.32081226	0.00000000	-571.32081224
test09	-29.25662600	-29.25662580	0.00000020	-29.25662575
test10	-577.82670817	-577.82670817	0.00000000	-577.82670839
test11	-274.66419189	-274.66419189	0.00000000	-274.66419186
test12	-447.68124595	-447.68124595	0.00000000	-447.68124599
test13	-23.98569013	-23.98569013	0.00000000	-23.98569013
test14	-159.69729413	-159.69729413	0.00000000	-159.69729410
test15	-5229.83555014	-5229.83555014	0.00000000	-5229.83555019
test16	-2995.28683802	-2995.28683802	0.00000000	-2995.28683806
test17	-2674.37559033	-2674.37559033	0.00000000	-2674.37559036
test18	-679.27667659	-679.27667659	0.00000000	-679.27667667
test19	-223.80715189	-223.80715189	0.00000000	-223.80715191
test20	-89.95529811	-89.95529811	0.00000000	-89.95529813
test21	-447.57509617	-447.57509617	0.00000000	-447.57509621
test22	-460.71872430	-460.71872430	0.00000000	-460.71872433
test23	-8.06132051	-8.06132051	0.00000000	-8.06132047
test24	-1400.17836984	-1400.17836984	0.00000000	-1400.17836996
test25	-74.84419555	-74.84419555	0.00000000	-74.84420399
test26	-58.42098569	-58.42098569	0.00000000	-58.42098616
test27	-2800.73720223	-2800.73720223	0.00000000	-2800.73720229
test28	-8.01152780	-8.01152780	0.00000000	-8.01152475
test29	-2047.64342093	-2047.64342093	0.00000000	-2047.64342095
test30	-109.04434229	-109.04434229	0.00000000	-109.04434299
test31	-4095.28742264	-4095.28742263	0.00000001	-4095.28742264
test32	-92.14081024	-92.14081024	0.00000000	-92.14081075
test33	-92.14161286	-92.14161286	0.00000000	-92.14161323
test34	-1117.52300498	-1117.52300498	0.00000000	-1117.52300500
test35	-936.50185912	-936.50185912	0.00000000	-936.50185910
test36	-112.56490316	-112.56490316	0.00000000	-112.56490317
test37	-3028.36857687	-3028.36857687	0.00000000	-3028.36857693
test38	-2279.14018803	-2279.14018803	0.00000000	-2279.14018803
test43	-1613.25545994	-1613.25545994	0.00000000	-1613.25546005
test47	-4394.62706583	-4394.62706583	0.00000000	-4394.62706582

test01_dft	-40.32090900	-40.32090900	0.00000000	-40.32090900
test02_dft	-224.92495090	-224.92495090	0.00000000	-224.92495093
test03_dft	-895.51821309	-895.51821309	0.00000000	-895.51821314
test04_dft	-1403.86074899	-1403.86074899	0.00000000	-1403.86074926
test05_dft	-75.25766593	-75.25766593	0.00000000	-75.25766595
test08_dft	-573.30058383	-573.30058383	0.00000000	-573.30058387
test11_dft	-275.43131232	-275.43131232	0.00000000	-275.43131232
test39_dft	-485.38264752	-485.38264752	0.00000000	-485.38264774
test40_dft	-337.10946296	-337.10946296	0.00000000	-337.10946311
test41_dft	-1313.33738975	-1313.33738975	0.00000000	-1313.33738951
test42_dft	-7.52368726	-7.52368726	0.00000000	-7.52368748
test44_dft	-1421.65785507	-1421.65785507	0.00000000	-1421.65785497
test45_dft	-5280.86080258	-5280.86080258	0.00000000	-5280.86080182
test46_dft	-5280.43953416	-5280.43953416	0.00000000	-5280.43953374
test47_dft	-4406.90099322	-4406.90099322	0.00000000	-4406.90099348
test48_dft	-37.61967832	-37.61967832	0.00000000	-37.61967832
test49_dft	-37.61967828	-37.61967828	0.00000000	-37.61967832
test50_dft	-37.61967828	-37.61967828	0.00000000	-37.61967832
test51_dft	-1020.28687445	-1020.28687453	-0.00000008	-1020.28687390

Appendix I

CRYSTAL23 versus CRYSTAL17

CRYSTAL23 introduces very few changes to the default settings of the program compared to CRYSTAL17:

SCF

- The DIIS scheme is disabled by default when running in massively parallel mode with MPPCRYSTAL; It can be activated by use of the **DIIS** keyword.

Density Functional Theory

- The size of the numerical integration grid has increased for meta-GGA functionals. For most of them, it corresponds to the **XXXLGRID** option, while for SCAN based DFAs is set to **HUGEGRID** option. In both cases the new default integration grid cannot be modified by the user.

Test cases
CRYSTAL23 versus CRYSTAL17 total energies (hartree)

	CRYSTAL17 [default] Etot	CRYSTAL23 [default] Etot	CRY23-CRY17 [default] DEtot
test00	-110.76493545	-110.76493545	0.00000000
test01	-39.72672424	-39.72672424	0.00000000
test02	-223.78754821	-223.78754821	0.00000000
test03	-893.87461011	-893.87461011	0.00000000
test04	-1400.17840968	-1400.17840968	0.00000000
test05	-74.83336244	-74.83336244	0.00000000
test06	-58.42082558	-58.42082558	0.00000000
test07	-2800.73725684	-2800.73725684	0.00000000
test08	-571.32081224	-571.32081222	0.00000002
test09	-29.25662575	-29.25662588	-0.00000013
test10	-577.82670839	-577.82670820	0.00000019
test11	-274.66419186	-274.66419186	0.00000000
test12	-447.68124599	-447.68124600	-0.00000001
test13	-23.98569013	-23.98569013	0.00000000
test14	-159.69729410	-159.69729410	0.00000000
test15	-5229.83555019	-5229.83555019	0.00000000
test16	-2995.28683806	-2995.28683806	0.00000000
test17	-2674.37559036	-2674.37559036	0.00000000
test18	-679.27667667	-679.27667667	0.00000000
test19	-223.80715191	-223.80715191	0.00000000
test20	-89.95529813	-89.95529813	0.00000000
test21	-447.57509621	-447.57509621	0.00000000
test22	-460.71872433	-460.71872433	0.00000000
test23	-8.06132047	-8.06132047	0.00000000
test24	-1400.17836996	-1400.17836996	0.00000000
test25	-74.84420399	-74.84420399	0.00000000
test26	-58.42098616	-58.42098616	0.00000000
test27	-2800.73720229	-2800.73720229	0.00000000
test28	-8.01152727	-8.01152727	0.00000000
test29	-2047.64342095	-2047.64342095	0.00000000
test30	-109.04434299	-109.04434299	0.00000000
test31	-4095.28742264	-4095.28742264	0.00000000
test32	-92.14081075	-92.14081075	0.00000000
test33	-92.14161323	-92.14161323	0.00000000
test34	-1117.52300500	-1117.52300500	0.00000000
test35	-936.50185910	-936.50185910	0.00000000
test36	-112.56490317	-112.56490317	0.00000000
test37	-3028.36857693	-3028.36857693	0.00000000
test38	-2279.14018803	-2279.14018803	0.00000000
test43	-1613.25546005	-1613.25546005	0.00000000
test47	-4394.62706582	-4394.62706582	0.00000000

test01_dft	-40.32090900	-40.32090900	0.00000000
test02_dft	-224.92495093	-224.92495093	0.00000000
test03_dft	-895.51821314	-895.51821314	0.00000000
test04_dft	-1403.86074926	-1403.86074926	0.00000000
test05_dft	-75.25766595	-75.25766593	0.00000002
test08_dft	-573.30058387	-573.30058388	-0.00000001
test11_dft	-275.43131232	-275.43131233	-0.00000001
test39_dft	-485.38264774	-485.38264773	0.00000001
test40_dft	-337.10946311	-337.10946311	0.00000000
test41_dft	-1313.33738951	-1313.33738951	0.00000000
test42_dft	-7.52368748	-7.52368748	0.00000000
test44_dft	-1421.65785497	-1421.65785497	0.00000000
test45_dft	-5280.86080182	-5280.86080182	0.00000000
test46_dft	-5280.43953374	-5280.43953375	-0.00000001
test47_dft	-4406.90099348	-4406.90099350	-0.00000002
test48_dft	-37.61967832	-37.61967832	0.00000000
test49_dft	-37.61967832	-37.61967832	0.00000000
test50_dft	-37.61967832	-37.61967832	0.00000000
test51_dft	-1020.28687390	-1020.28688020	-0.00000630

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] A. Savin, J. Toulouse and H. J. Flad. Short-range exchange-correlation energy of a uniform electron gas with modified electron-electron interaction. *Int. J. Quantum Chem.*, 100(6):1047, 2004.

[2] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov and G. E. Scuseria. . *J. Chem. Phys.*, 125(22):224106, 2006.

[3] M. Abramovitz and I. Segun, editors. *Handbook of Mathematical Functions*. Dover, New York, 1965.

[4] A.D. Becke. *J. Chem. Phys.*, 107:8544, 1997.

[5] C. Adamo and V. Barone. Toward reliable density functional methods without adjustable parameters: the PBE0 model. *J. Chem. Phys.*, 110:6158–6170, 1999.

[6] Carlo Adamo and Vincenzo Barone. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The m pw and m pw1pw models. *The Journal of chemical physics*, 108(2):664–675, 1998.

[7] D. G. Anderson. Iterative procedures for nonlinear integral equations. *J. Assoc. Comput. Mach.*, 12:547, 1965.

[8] E. Aprà. *Metodi quanto-meccanici per lo studio di solidi cristallini*. PhD thesis, University of Torino, Torino, 1993.

[9] N.W. Ashcroft and N.D. Mermin. *Solid state physics*. Saunders College, Philadelphia, 1976.

[10] 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.

[11] R. R. W. Bader. *Atoms in Molecules*. Clarendon Press, Oxford, 1990.

[12] E. B. Barrosa, A. Joriob, G. G. Samsonidzef, R. B. Capazc, A. G. Souza Filhoa, J. Mendes Filhoa, G. Dresselhause, and M. S. Dresselhaus. Review on the symmetry-related properties of carbon nanotubes. *Phys. Rep.*, 431:261–302 (and references therein), 2006.

[13] J. C. Barthelat. Private communication.

[14] J. C. Barthelat and P. Durand. Recent progress of pseudopotentials in quantum chemistry. *Gazz. Chim. Ital.*, 108:225, 1978.

[15] J. C. Barthelat, P. Durand, and A. Serafini. Non-empirical pseusopotentials for molecular calculations. *Molec. Phys.*, 33:159, 1977.

[16] A. D. Becke. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.*, 88:2547, 1988.

[17] A. D. Becke. Correlation energy of an inhomogeneous electron gas: A coordinate-space model. *J. Chem. Phys.*, 88:1053, 1988.

[18] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38:3098–3100, 1988.

[19] A. D. Becke. Density-functional thermochemistry. III The role of exact exchange. *J. Chem. Phys.*, 98:5648, 1993.

[20] A. D. Becke and E. R. Johnson. *J. Chem. Phys.*, 122:154101, 2005.

[21] A.D. Becke and K.E. Edgecombe. A simple measure of the electron localization in atomic and molecular systems. *J. Chem. Phys.*, 92:5397, 1990.

[22] Axel D Becke. Density-functional thermochemistry. iv. a new dynamical correlation functional and implications for exact-exchange mixing. *The Journal of chemical physics*, 104(3):1040–1046, 1996.

[23] Axel D Becke. Density-functional thermochemistry. iv. a new dynamical correlation functional and implications for exact-exchange mixing. *The Journal of chemical physics*, 104(3):1040–1046, 1996.

[24] K Bencheikh. Spin–orbit coupling in the spin-current-density-functional theory. *J. Phys. A*, 36(48):11929, 2003.

[25] D. I. Bilc, R. Orlando, R. Shaltaf, G. M. Rignanese, J. Iniguez, and Ph. Ghosez. Hybrid exchange-correlation functional for accurate prediction of the electronic and structural properties of ferroelectric oxides. *Phys. Rev. B*, 77:165107, 2008.

[26] J. S. Binkley, J. A. Pople, and W. J. Hehre. Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *J. Am. Chem. Soc.*, 102:939, 1980.

[27] D. M. Bishop. Explicit nondivergent formulas for atomic and molecular dynamic hyperpolarizabilities. 100:6535, 1994.

[28] D. M. Bishop and B. Kirtman. A perturbation method for calculating vibrational dynamic dipole polarizabilities and hyperpolarizabilities. 95:2646, 1991.

[29] Filippo Bodo, Jacques K Desmarais, and Alessandro Erba. Spin current density functional theory of weyl semimetals. *Phys. Rev. B*, 105(12):125108, 2022.

[30] C. F. Bohren and D. Huffman. *Absorption and scattering of light by small particles*, page 356. Wiley-Intersience Publication, New York, 1983.

[31] Y. Bouteiller, C. Mijoule, M. Nizam, J. C. Barthelat, J.P. Daudey, M. Pellissier, and B. Silvi. Extended gaussian-type valence basis sets for calculations involving non-empirical core pseudopotentials I. PS-31 G basis for Li to Ca and Ga to Kr atoms. *Mol. Phys.*, 65:295, 1988.

[32] S. F. Boys and F. Bernardi. *Mol. Phys.*, 19:553–566, 1970.

[33] 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.

[34] J. G. Brandenburg, M. Alessio, B. Civalleri, M. F. Peintinger, T. Bredow, and S. Grimme. *J. Phys. Chem. A*, 117:9282–9292, 2013.

[35] J. G. Brandenburg, E. Caldeweyher, and S. Grimme. Screened exchange hybrid density functional for accurate and efficient structures and interaction energies. *Phys. Chem. Chem. Phys.*, 18:15519–15523, 2016.

[36] J. G. Brandenburg and S. Grimme. Dispersion Corrected Hartree-Fock and Density Functional Theory for Organic Crystal Structure Prediction. *Top Curr Chem*, 345:1–23, 2014.

[37] C. G. Broyden. A class of methods for solving nonlinear simultaneous equations. *Math. Comput.*, 19:577, 1965.

[38] C. G. Broyden. The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA J. Appl. Math.*, 6:76–90, 1970.

[39] C.G. Broyden. The Convergence of a Class of Double-rank Minimization Algorithms 1. The New Algorithm. *IMA J. Appl. Math.*, 6:222–231, 1970.

[40] Markus Bursch, Hagen Neugebauer, Sebastian Ehlert, and Stefan Grimme. Dispersion corrected r2scan based global hybrid functionals: r2scanh, r2scan0, and r2scan50. *The Journal of Chemical Physics*, 156(13):134105, 2022.

[41] H. Burzlaff and A. Hountas. Computer program for the derivation of symmetry operations from the space-group symbols. *J. Appl. Cryst.*, 15:464, 1982.

[42] C. A. Guido, E. Bremond, C. Adamo and P. Cortona. One third: A new recipe for the PBE0 paradigm. *J. Chem. Phys.*, 138:021104, 2013.

[43] V. R. Saunders C. Gatti and C. Roetti. Crystal field effects on the topological properties of the electron density in molecular crystals. The case of urea. *J. Chem. Phys.*, 101:10686, 1994.

[44] S. Casassa and R. Demichelis. Relative energy of aluminium hydroxides: the role of electron correlation. *J. Phys. Chem. C*, 116:13313, 2012.

[45] S. Casassa, C. M. Zicovich-Wilson, and C. Pisani. Symmetry-adapted localized Wannier functions suitable for periodic calculations. *Theor. Chem. Acc.*, 116:726–733, 2006.

[46] 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.

[47] 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.

[48] 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.

[49] 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.

[50] B. Civalleri, C. M. Zicovich-Wilson, L. Valenzano, and P. Ugliengo. B3LYP augmented with an empirical dispersion term (B3LYP-D*) as applied to molecular crystal. *Crys. Eng. Comm.*, 10:405–410, 2008.

[51] B. Civalleri, C. M. Zicovich-Wilson, L. Valenzano, and P. Ugliengo. ERRATA CORRIGE TO: "B3LYP augmented with an empirical dispersion term (B3LYP-D*) as applied to molecular crystals. Cryst. Eng. Comm., 10, 405-410 (2008)". *Cryst. Eng. Comm.*, 10:1693–1694, 2008.

[52] E. Clementi. *MOTECC 91*. ESCON, Leiden, 1991.

[53] Enrico Clementi. *Modern techniques in computational chemistry: MOTECC-91*, volume 91. Springer Science & Business Media, 1991.

[54] Beatriz Cordero, Verónica Gómez, Ana E Platero-Prats, Marc Revés, Jorge Echeverría, Eduard Cremades, Flavia Barragán, and Santiago Alvarez. Covalent radii revisited. *Dalton Trans.*, (21):2832–2838, 2008.

[55] M. Cutini, B. Civalleri, M. Corno, R. Orlando, J. G. Brandenburg, L. Maschio, and P. Ugliengoa. Assessment of different quantum mechanical methods for the prediction of structure and cohesive energy of molecular crystals. *J. Chem. Theory Comput.*, 12:3340–3352, 2016.

[56] Loredana Edith Daga, Bartolomeo Civalleri, and Lorenzo Maschio. Gaussian basis sets for crystalline solids: all-purpose basis set libraries vs system-specific optimizations. *Journal of Chemical Theory and Computation*, 16(4):2192–2201, 2020.

[57] S. Dall'Olio, R. Dovesi, and 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.

[58] Ph. D'Arco, Y. Noel, R. Demichelis, and R. Dovesi. Single-layered chrysotile nanotubes: A quantum mechanical ab initio simulation. *J. Chem. Phys.*, 131:204701, 2009.

[59] C. Darrigan, M. Rerat, G. Mallia, and 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.

[60] R. Demichelis, B. Civalleri, M. Ferrabone, and R. Dovesi. On the performance of eleven DFT functionals in the description of the vibrational properties of aluminosilicates. *Int. J. Quantum Chem.*, 110:406–415, 2010.

[61] R. Demichelis, Y. Noel, Ph. D'Arco, L. Maschio, and R. Orlando. Structure and energetics of imogolite: a quantum mechanical ab initio study with B3LYP hybrid functional. *J. Mater. Chem.*, 20:10417–10425, 2010.

[62] Jacques K Desmarais, Jean-Pierre Flament, and Alessandro Erba. Fundamental role of fock exchange in relativistic density functional theory. *J. Phys. Chem. Lett.*, 10(13):3580–3585, 2019.

[63] Jacques K Desmarais, Jean-Pierre Flament, and Alessandro Erba. Spin-orbit coupling from a two-component self-consistent approach. i. generalized hartree-fock theory. *J. Chem. Phys.*, 151(7):074107, 2019.

[64] Jacques K Desmarais, Jean-Pierre Flament, and Alessandro Erba. Adiabatic connection in spin-current density functional theory. *Phys. Rev. B*, 102(23):235118, 2020.

[65] Jacques K Desmarais, Jean-Pierre Flament, and Alessandro Erba. Spin-orbit coupling in periodic systems with broken time-reversal symmetry: Formal and computational aspects. *Phys. Rev. B*, 101(23):235142, 2020.

[66] Jacques K Desmarais, Stanislav Komorovsky, Jean-Pierre Flament, and Alessandro Erba. Spin–orbit coupling from a two-component self-consistent approach. ii. non-collinear density functional theories. *J. Chem. Phys.*, 154(20):204110, 2021.

[67] Jacques Kontak Desmarais, A Erba, and R Dovesi. Generalization of the periodic lcao approach in the crystal code to g-type orbitals. *Theor. Chem. Acc.*, 137(2):1–11, 2018.

[68] P. A. M. Dirac. Note on exchange phenomena in the Thomas-Fermi atom. *Proc. Cambridge Phil. Soc.*, 26:376, 1930.

[69] Michael Dolg and Xiaoyan Cao. Accurate relativistic small-core pseudopotentials for actinides. energy adjustment for uranium and first applications to uranium hydride. *J. Phys. Chem. A*, 113(45):12573–12581, 2009.

[70] L Doná, JG Brandenburg, and B Civalleri. Extending and assessing composite electronic structure methods to the solid state. *J. Chem. Phys.*, 151:121101, 2019.

[71] 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.

[72] 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.

[73] 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.

[74] R. Dovesi and R. Orlando. Convergence properties of the supercell approach in the study of local defects in solids. *Phase Trans.*, 52:151, 1994.

[75] R. Dovesi, R. Orlando, B. Civalleri, C. Roetti, V. R. Saunders, and C. M. Zicovich-Wilson. Crystal: a computational tool for the ab initio study of the electronic properties of crystals. *Z. Kristallogr.*, 220:571–573, 2005.

[76] R. Dovesi, R. Orlando, A. Erba, C. M. Zicovich-Wilson, B. Civalleri, S. Casassa, L. Maschio, M. Ferrabone, M. De La Pierre, Ph. D'Arco, Y. Noël, M. Causá, M. Rérat, and B. Kirtman. CRYSTAL14: A program for the ab initio investigation of crystalline solids. *Int. J. Quantum Chem.*, 114:1287–1317, 2014.

[77] 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.

[78] 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.

[79] 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.

[80] R. Dovesi, C. Roetti, and V. R. Saunders. *CRYSTAL92 User's Manual*. Università di Torino and SERC Daresbury Laboratory, 1992.

[81] 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.

[82] 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, and M. Llunell. *CRYSTAL06 User's Manual*. Università di Torino, Torino,2006.

[83] Roberto Dovesi, Alessandro Erba, Roberto Orlando, Claudio M Zicovich-Wilson, Bartolomeo Civalleri, Lorenzo Maschio, Michel Rérat, Silvia Casassa, Jacopo Baima, Simone Salustro, et al. Quantum-mechanical condensed matter simulations with crystal. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 8(4):e1360, 2018.

[84] P. Durand and J. C. Barthelat. A theoretical method to determine atomic pseudopotentials for electronic structure calculations of molecules and solids. *Theor. Chim. Acta*, 38:283–302, 1975.

[85] E. Weintraub, T. M. Henderson, Scuseria G. E. Long-Range-Corrected Hybrids Based on a New Model Exchange Hole. *J. Chem. Theory Comput.*, 5:754–762, 2009.

[86] U. Ekström, L. Visscher, R. Bast, A. J. Thorvaldsen, and K. Ruud. Arbitrary-order density functional response theory from automatic differentiation. *J. Chem. Theory Comput.*, 6:1971, 2010.

[87] A. Erba, S. Casassa, R. Dovesi, L. Maschio, and C. Pisani. Periodic density functional theory and local-mp2 study of the librational modes of ice xi. *J. Chem. Phys.*, 130:074505, 2009.

[88] A. Erba and R. Dovesi. Photoelasticity of crystals from theoretical simulations. *Phys. Rev. B*, 88:045121, 2013.

[89] A. Erba, Kh. E. El-Kelany, M. Ferrero, I. Baraille, and M. Rérat. Piezoelectricity of srtio$_3$: An ab initio description. *Phys. Rev. B*, 88:035102, 2013.

[90] A. Erba, M. Ferrabone, J. Baima, R. Orlando, M. Rérat, and R. Dovesi. The vibration properties of the (n,0) boron nitride nanotubes from ab initio quantum chemical simulations. *J. Chem. Phys.*, 138:054906, 2013.

[91] A. Erba, M. Ferrabone, R. Orlando, and R. Dovesi. Accurate dynamical structure factors from ab initio lattice dynamics: The case of crystalline silicon. *J. Comput. Chem.*, 34:346, 2013.

[92] A. Erba and M. Halo. *CRYSCOR09 User's Manual*. Università di Torino, Torino, 2009. http://www.cryscor.unito.it.

[93] A. Erba, A. Mahmoud, R. Orlando, and R. Dovesi. Elastic properties of six silicate garnet end-members from accurate ab initio simulations. *Phys. Chem. Miner.*, 41:151–160, 2014.

[94] A. Erba, L. Maschio, S. Salustro, and S. Casassa. A post-hartree-fock study of pressure-induced phase transitions in solid nitrogen: the case of the α, γ and ϵ low-pressure phases. *J. Chem. Phys.*, 134:074502, 2011.

[95] A. Erba and C. Pisani. Evaluation of the electron momentum density of crystalline systems from ab initio linear combination of atomic orbitals calculations. *J. Comput. Chem.*, 33:822, 2012.

[96] A. Erba, C. Pisani, S. Casassa, L. Maschio, M. Schütz, and D. Usvyat. A mp2 versus dft theoretical investigation of the compton profiles of crystalline urea. *Phys.Rev. B*, 81:165108, 2010.

[97] O. Eriksson and A. Svane. Isomer shifts and hyperfine fields in iron compounds. *J. Phys.: Condens. Matter*, 1:1589, 1989.

[98] Walter C. Ermler, Yoon S. Lee, Phillip A. Christiansen, and Kenneth S. Pitzer. Ab initio effective core potentials including relativistic effects. a procedure for the inclusion of spin-orbit coupling in molecular wavefunctions. *Chem. Phys. Lett.*, 81(1):70–74, 1981.

[99] WC Ermler, RB Ross, and PA Christiansen. Ab initio relativistic effective potentials with spin-orbit operators. vi. fr through pu. *Int. J. Quantum Chem.*, 40(6):829–846, 1991.

[100] V. R. Saunders et al. *Electronic structure theory: from molecules to crystals*. private communication, 1999.

[101] F. A. Hamprecht, A. J. Cohen, D. J. Tozer and N. C. Handy. *J. Chem. Phys.*, 109(15):6264, 1998.

[102] Luis Fernandez Pacios and PA Christiansen. Ab initio relativistic effective potentials with spin-orbit operators. i. li through ar. *J. Chem. Phys.*, 82(6):2664–2671, 1985.

[103] M. Ferrero, M. Rérat, B. Kirtman, and R. Dovesi. Calculationoffirst and second static hyperpolarizabilities of one- to three-dimensional periodic compounds. implementationin the crystal code. *J. Chem. Phys.*, 129:244110, 2008.

[104] M. Ferrero, M. Rérat, R. Orlando, and R. Dovesi. Coupled perturbed Hartree-Fock for periodic systems: the role of symmetry and related computational aspects. *J. Chem. Phys.*, 128:014100, 2008.

[105] M. Ferrero, M. Rérat, R. Orlando, and R. Dovesi. The calculation of static polarizabilities in 1-3D periodic compounds. The implementation in the CRYSTAL code. *J. Comput. Chem.*, 29:1450–1459, 2008.

[106] M. Ferrero, M. Rérat, R. Orlando, R. Dovesi, and I. J. Bush. Coupled Perturbed Kohn-Sham calculation of static polarizabilities of periodic compounds. *J. Phys.: Conf. Ser.*, 117:012016, 2008.

[107] Detlev Figgen, Guntram Rauhut, Michael Dolg, and Hermann Stoll. Energy-consistent pseudopotentials for group 11 and 12 atoms: adjustment to multi-configuration dirac–hartree–fock data. *Chem. Phys.*, 311(1-2):227–244, 2005.

[108] M. Filatov. On the calculation of mössbauer isomer shift. *J. Chem. Phys.*, 127:084101, 2007.

[109] R. Fletcher. A new approach to variable metric algorithms. *Comput. J.*, 13:317, 1970.

[110] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr. and T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, , and J. A. Pople. *Gaussian 03, Revision 02*. Gaussian, Inc., Wallingford CT, 2004, 2004.

[111] 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. and 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.

[112] 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.

[113] 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.

[114] James W Furness, Aaron D Kaplan, Jinliang Ning, John P Perdew, and Jianwei Sun. Accurate and numerically efficient r2scan meta-generalized gradient approximation. *The journal of physical chemistry letters*, 11(19):8208–8215, 2020.

[115] C. Gatti. *TOPOND-96 : an electron density topological program for systems periodic in N (N=0-3) dimensions, User's manual*. CNR-CSRSRC, Milano, 1996.

[116] C. Gatti. Chemical bonding in crystals: new direction. *Zeitschrift für Kristallographie*, 220:399, 2005.

[117] C. Gatti. Challenging chemical concepts through charge density of molecules and crystals. *Phys. Scr.*, 87:048102, 2013.

[118] C. Giacovazzo. *Fundamentals of Crystallography*. IUCr texts on crystallography. IUC - Oxford University Press, Oxford, 1992.

[119] G. Gilat. Analysis of methods for calculating spectral properties in solids. *J. Comp. Phys.*, 10:432, 1972.

[120] G. Gilat and L. J. Raubenheimer. Accurate numerical method for calculating frequency-distribution in solids. *Phys. Rev.*, 144:390, 1966.

[121] P. M. W. Gill, B. G. Johnson, and J. A. Pople. A standard grid for density function calculations. *Chem. Phys. Lett.*, 209:506–512, 1993.

[122] D. Goldfarb. A family of variable-metric methods derived by variational means. *Math. Comput.*, 24:23, 1970.

[123] M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, and W. J. Hehre. Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements. *J. Am. Chem. Soc.*, 104:2797, 1982.

[124] S. Grimme. *J. Comp. Chem.*, 27:1787, 2006.

[125] S. Grimme. Semiempirical gga-type density functional constructed with a long range dispersion correction. *J. Comput. Chem.*, 27:1787–, 2006.

[126] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. *J. Chem. Phys.*, 132:154104, 2010.

[127] S. Grimme, J. G. Brandenburg, C. Bannwarth, and A. Hansen. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *J. Chem. Phys.*, 143:054107, 2015.

[128] S. Grimme, J. G. Brandenburg, C. Bannwarth, and A. Hansen. *J. Chem. Phys.*, 148:064104, 2018.

[129] S. Grimme, S. Ehrlich, and L. Goerigk. *J. Comp. Chem.*, 32:1456, 2011.

[130] S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth. Dispersion-corrected mean-field electronic structure methods. *Chem. Rev.*, 116:5105–5154, 2016.

[131] Giuseppe Grosso and Giuseppe Pastori Parravicini. *Solid State Physics*. Academic Press, 2000.

[132] M. F. Guest and V. R. Saunders. On methods for converging open-shell Hartree-Fock wave-functions. *Mol. Phys.*, 28:819–828, 1974.

[133] T. Hahn. Space Group Symmetry. In *International Tables for Crystallography*, volume A. Reidel Publishing Company, 1987.

[134] D. R. Hamann. Semiconductor Charge Densities with Hard-Core and Soft-Core Pseudopotentials. *Phys. Rev. Lett.*, 42:662–665, 1979.

[135] Tim Hangele and Michael Dolg. Accuracy of relativistic energy-consistent pseudopotentials for superheavy elements 111–118: Molecular calibration calculations. *J. Chem. Phys.*, 138(4):044104, 2013.

[136] Tim Hangele, Michael Dolg, Michael Hanrath, Xiaoyan Cao, and Peter Schwerdtfeger. Accurate relativistic energy-consistent pseudopotentials for the superheavy elements 111 to 118 including quantum electrodynamic effects. *J. Chem. Phys.*, 136(21):214105, 2012.

[137] Tim Hangele, Michael Dolg, and Peter Schwerdtfeger. Relativistic energy-consistent pseudopotentials for superheavy elements 119 and 120 including quantum electrodynamic effects. *J. Chem. Phys.*, 138(17):174113, 2013.

[138] S Yu Haoyu, Xiao He, Shaohong L Li, and Donald G Truhlar. Mn15: A kohn–sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and non-covalent interactions. *Chemical science*, 7(8):5032–5051, 2016.

[139] P. C. Hariharan and J. A. Pople. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta*, 28:213, 1973.

[140] 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.

[141] 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.

[142] 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.

[143] 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.

[144] D. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople. *Ab initio molecular orbital theory.* Wiley, New York, 1986.

[145] W. J. Hehre, R. Ditchfield, R. F. Stewart, and J. A. Pople. Self-Consistent Molecular Orbital Methods. IV. Use of Gaussian Expansions of Slater-Type Orbitals. Extension to Second-Row Molecules. *J. Chem. Phys.*, 52:2769, 1970.

[146] W. J. Hehre, R. F. Stewart, and J. A. Pople. Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals. *J. Chem. Phys.*, 51:2657, 1969.

[147] P Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864, 1964.

[148] Bruce S. Hudson. Inelastic neutron scattering: A tool in molecular vibrational spectroscopy and a test of ab initio methods. *J. Phys. Chem. A*, 105(16):3949–3960, 2001.

[149] Kerwin Hui and Jeng-Da Chai. Scan-based hybrid and double-hybrid density functionals from models without fitted parameters. *The Journal of chemical physics*, 144(4):044114, 2016.

[150] MM Hurley, Luis Fernandez Pacios, PA Christiansen, RB Ross, and WC Ermler. A binitio relativistic effective potentials with spin-orbit operators. ii. k through kr. *J. Chem. Phys.*, 84(12):6840–6853, 1986.

[151] I. C. Gerber and J. G. Angyan. London dispersion forces by range-separated hybrid density functional with second order perturbational corrections: The case of rare gas complexes. *J. Chem. Phys.*, 126(4), 2007.

[152] H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao. A long-range correction scheme for generalized-gradient approximation exchange functionals. *J. Chem. Phys.*, 115:3540–3544, 2001.

[153] Jeng-Da Chai and M. Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.*, 128(8):084106, 2008.

[154] D. D. Johnson. Modified Broyden's method for accelerating convergence in self-consistent calculations. *Phys. Rev B*, 38:12807–12813, 1988.

[155] E. R. Johnson and A. D. Becke. *J. Chem. Phys.*, 123:024101, 2005.

[156] E. R. Johnson and A. D. Becke. *J. Chem. Phys.*, 124:174104, 2006.

[157] W. D. Johnston. Nonlinear optical coefficients and the raman scattering efficiency of lo and to phonons in acentric insulating cryst als. *Phys. Rev. B*, 1:3494, 1970.

[158] D.D. Koelling and B.N. Harmon. A technique for relativistic spin-polarised calculations. *J. Phys. C*, 10(16):3107, 1977.

[159] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133, 1965.

[160] L. Kronsjo. *Algorithms, their complexity and efficiency*, volume 2nd ed. Wiley, New York, 1986.

[161] H. Kruse and S. Grimme. A geometrical correction for the inter- and intramolecular basis set superposition error in hartree-fock and density functional theory calculations for large systems. *J. Chem. Phys.*, 136:154101, 2012.

[162] L. Schimka, J. Harl and G. Kresse. Improved hybrid functional for solids: The hsesol functional. *J. Chem. Phys.*, 134(2):024116, 2011.

[163] LA LaJohn, PA Christiansen, RB Ross, T Atashroo, and WC Ermler. A binitio relativistic effective potentials with spin–orbit operators. iii. rb through xe. *J. Chem. Phys.*, 87(5):2812–2824, 1987.

[164] Adrian W Lange, Mary A Rohrdanz, and John M Herbert. Charge-transfer excited states in a π-stacked adenine dimer, as predicted using long-range-corrected time-dependent density functional theory. *The Journal of Physical Chemistry B*, 112(20):6304–6308, 2008.

[165] Joachim Laun and Thomas Bredow. Bsse-corrected consistent gaussian basis sets of triple-zeta valence with polarization quality of the sixth period for solid-state calculations. *Journal of Computational Chemistry*, 42(15):1064–1072, 2021.

[166] Joachim Laun and Thomas Bredow. Bsse-corrected consistent gaussian basis sets of triple-zeta valence with polarization quality of the fifth period for solid-state calculations. *Journal of Computational Chemistry*, 43(12):839–846, 2022.

[167] 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.

[168] 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.

[169] WanZhen Liang and Martin Head-Gordon. Approaching the basis set limit in density functional theory calculations using dual basis sets without diagonalization. *J. Phys. Chem. A*, 108:3206–3210, 2004.

[170] Ivan S Lim, Peter Schwerdtfeger, Bernhard Metz, and Hermann Stoll. All-electron and relativistic pseudopotential studies for the group 1 element polarizabilities from k to element 119. *J. Chem. Phys.*, 122(10):104103, 2005.

[171] Ivan S Lim, Hermann Stoll, and Peter Schwerdtfeger. Relativistic small-core energy-consistent pseudopotentials for the alkaline-earth elements from ca to ra. *J. Chem. Phys.*, 124(3):034107, 2006.

[172] B. Lindberg. A new efficient method for calculation of energy eigenvalues and eigenstates of the one-dimensional Schrödinger equation. *J. Chem. Phys.*, 88:3805, 1988.

[173] R. Lindh, A. Bernhardsson, G. Katlström, and P.A. Malmqvist. On the use of a hessian model function in molecular geometry optimizations. *Chem. Phys. Lett.*, 241:423–428, 1995.

[174] Matthew S. Lucas, M. Kresch, R. Stevens, and B. Fultz. Phonon partial densities of states and entropies of fe and cr in bcc fe-cr from inelastic neutron scattering. *Phys. Rev. B*, 77:184303, 2008.

[175] Benjamin J Lynch, Patton L Fast, Maegan Harris, and Donald G Truhlar. Adiabatic connection for kinetics. *The Journal of Physical Chemistry A*, 104(21):4811–4815, 2000.

[176] T. Bredow M. F. Peintinger, D. V. Oliveira. Consistent gaussian basis sets of triple-zeta valence with polarization quality for solid-state calculations. *J. Comput. Chem.*, 34:451–459, 2013.

[177] V. Magnasco and A. Perico. Uniform Localization of Atomic and Molecular Orbitals. I. *J. Chem. Phys.*, 47:971, 1967.

[178] N.L. Marana, Y. Noel, J.R. Sambrano, C. Ribaldone, and S. Casassa. Ab initio modeling of multiwall: A general algorithm first applied to carbon nanotubes. *J. Phys. Chem. A*, 18:4003–4008, 2021.

[179] R. Martinez-Casado, G. Mallia, D. Usvyat, L. Maschio, S. Casassa, M. Schütz, and N. Harrison. Periodic quantum mechanical simulation of the he-mgo(100) interaction potential. *J. Chem. Phys.*, 134:014706, 2011.

[180] L. Maschio, B. Kirtman, R. Orlando, and M. Rérat. *Ab Initio* analytical infrared intensities for periodic systems through a coupled perturbed hartree-fock/kohn-sham method. 137:204113, 2012.

[181] L. Maschio, B. Kirtman, M. Rérat, R. Orlando, and R. Dovesi. Ab initio analytical Raman intensities for periodic systems through a coupled perturbed Hartree-Fock/Kohn-Sham method in an atomic orbital basis. I. Theory. *J. Chem. Phys.*, 139:164101, 2013.

[182] L. Maschio, B. Kirtman, M. Rérat, R. Orlando, and R. Dovesi. Ab initio analytical Raman intensities for periodic systems through a coupled perturbed Hartree-Fock/Kohn-Sham method in an atomic orbital basis. II. Validation and comparison with experiments. *J. Chem. Phys.*, 139:164102, 2013.

[183] L. Maschio, B. Kirtman, M. Rérat, R. Orlando, and R. Dovesi. Ab initio analytical Raman intensities for periodic systems through a coupled perturbed Hartree-Fock/Kohn-Sham method in an atomic orbital basis. I. Theory. 139:164101, 2013.

[184] L. Maschio, D. Usvyat, and B. Civalleri. Ab initio study of van der waals and hydrogen-bonded molecular crystals with a periodic local-mp2 method. *Crys. Eng. Comm.*, 12:2429, 2010.

[185] Lorenzo Maschio and Bernard Kirtman. Coupled perturbation theory approach to dual basis sets for molecules and solids. 1. general theory and application to molecules. *J. of Chem. Theory Comput.*, 16(1):340–353, 2019.

[186] Larry E McMurchie and Ernest R Davidson. Calculation of integrals over ab initio pseudopotentials. *J. Comput. Phys.*, 44(2):289–301, 1981.

[187] N. D. Mermin. Thermal properties of the inhomogeneous electron gas. *Phys. Rev. A*, 137:1441, 1965.

[188] Bernhard Metz, Marcus Schweizer, Hermann Stoll, Michael Dolg, and Wenjian Liu. A small-core multiconfiguration dirac–hartree–fock-adjusted pseudopotential for tl–application to tlx (x= f, cl, br, i). *Theor. Chem. Acc.*, 104(1):22–28, 2000.

[189] Bernhard Metz, Hermann Stoll, and Michael Dolg. Small-core multiconfiguration-dirac–hartree–fock-adjusted pseudopotentials for post-d main group elements: Application to pbh and pbo. *J. Chem. Phys.*, 113(7):2563–2569, 2000.

[190] D. Moncrieff and V. R. Saunders. *ATMOL manual*. Manchester, 1986.

[191] H. J. Monkhorst and J. D. Pack. Special points for Brillouin-zone integrations. *Phys. Rev. B*, 13:5188, 1976.

[192] R.L. Mössbauer. Kernresonanzfluoreszenz von gammastrahlung in ir^{191}. *Z. Physik*, 151:124, 1958.

[193] Clinton S Nash, Bruce E Bursten, and Walter C Ermler. Ab initio relativistic effective potentials with spin-orbit operators. vii. am through element 118. *J. Chem. Phys.*, 106(12):5133–5142, 1997.

[194] Y. Noel, Ph. D'Arco, R. Demichelis, C. M. Zicovich-Wilson, and R. Dovesi. On the use of symmetry in the ab initio quantum mechanical simulation of nanotubes and related materials. *J. Comput. Chem.*, 31:855–862, 2010.

[195] Y. Noël, M. De La Pierre, C.M. Zicovich-Wilson, R. Orlando, and R. Dovesi. Structural, electronic and energetic properties of giant icosahedral fullerenes up to C6000: insights from an ab initio hybrid DFT study. *Phys. Chem. Chem. Phys.*, 16:13390–13401, 2014.

[196] Y. Noel, C. M. Zicovich-Wilson, B. Civalleri, Ph. D'Arco, and R. Dovesi. Polarization properties of zno and beo: An ab initio study through the berry phase and wannier functions approaches. *Phys. Rev. B*, 65:014111, 2001.

[197] 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.

[198] R. Orlando, I. J. Bush M. Delle Piane, P. Ugliengo, M. Ferrabone, and R. Dovesi. A new massively parallel version of crystal for large systems on high performance computing architectures. *J. Comput. Chem.*, 33:2276–2284, 2012.

[199] B. J. Orr and J. F. Ward. Perturbation theory of the non-linear optical polarization of an isolated system. 20:513, 1971.

[200] R. Osborn, E. A. Goremychkin, A. I. Kolesnikov, and D. G. Hinks. Phonon density of states in mgb_2. *Phys. Rev. Lett.*, 87:017005, 2001.

[201] D. Viterbo P. Ugliengo and G. Chiari. MOLDRAW: Molecular Graphics on a Personal Computer. *Zeitschrift fur Kristallographie*, 207:9, 1993.

[202] J. P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B*, 33:8822, 1986.

[203] J. P. Perdew. Unified Theory of Exchange and Correlation Beyond the Local Density Approximation. In P. Ziesche and H. Eschrig, editors, *Electronic Structure of Solids 1991*, volume 11, pages 11–20, Berlin, 1991. Akademie Verlag.

[204] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, 77:3865–3868, 1996.

[205] 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.

[206] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Phys. Rev. Lett.*, 100:136406, 2008.

[207] J. P. Perdew and Y. Wang. Accurate and simple analytic representation of the electron gas correlation energy. *Phys. Rev. B*, 45:13244, 1992.

[208] 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.

[209] 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.

[210] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23:5048–5079, 1981.

[211] W. F. Perger, J. Criswell, B. Civalleri, and R. Dovesi. Ab-initio calculation of elastic constants of crystalline systems with the crystal code. *Comput. Phys. Commun.*, 180:1753–1759, 2009.

[212] Kirk A Peterson, Detlev Figgen, Michael Dolg, and Hermann Stoll. Energy-consistent relativistic pseudopotentials and correlation consistent basis sets for the 4 d elements y–pd. *J. Chem. Phys.*, 126(12):124101, 2007.

[213] Kirk A Peterson, Detlev Figgen, Erich Goll, Hermann Stoll, and Michael Dolg. Systematically convergent basis sets with relativistic pseudopotentials. ii. small-core pseudopotentials and correlation consistent basis sets for the post-d group 16–18 elements. *J. Chem. Phys.*, 119(21):11113–11123, 2003.

[214] Kirk A Peterson, Benjamin C Shepler, Detlev Figgen, and Hermann Stoll. On the spectroscopic and thermochemical properties of clo, bro, io, and their anions. *J. Phys. Chem. A*, 110(51):13877–13883, 2006.

[215] Roberto Peverati and Donald G Truhlar. Communication: A global hybrid generalized gradient approximation to the exchange-correlation functional that satisfies the second-order density-gradient constraint and has broad applicability in chemistry. *The Journal of Chemical Physics*, 135(19):191102, 2011.

[216] Roberto Peverati, Yan Zhao, and Donald G Truhlar. Generalized gradient approximation that recovers the second-order density-gradient expansion with optimized across-the-board performance. *The Journal of Physical Chemistry Letters*, 2(16):1991–1997, 2011.

[217] W. J. Pietro, E. S. Blurock, R. F. Hout, W. J. Hehre, W. J. DeFrees, and R. F. Stewart. Molecular orbital theory of the properties of inorganic and organometallic compounds. 2. STO-NG basis sets for fourth-row main-group elements. *Inorg. Chem.*, 20:3650, 1981.

[218] W. J. Pietro, B. A. Levi, W. J. Hehre, and R. F. Stewart. Molecular orbital theory of the properties of inorganic and organometallic compounds. 1. STO-NG basis sets for third-row main-group elements. *Inorg. Chem.*, 19:2225, 1980.

[219] J. Pipek and P. G. Mezey. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.*, 90:4916, 1989.

[220] C. Pisani. *Quantum-Mechanical Ab-Initio Calculation of the Properties of Crystalline Materials*, volume 67 of *Lecture Notes in Chemistry*. Springer Verlag, Berlin, 1996.

[221] 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.

[222] 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.

[223] C. Pisani, M. Busso, G. Capecchi, S. Casassa, R. Dovesi, L. Maschio, C. Zicovich-Wilson, and M. Schütz. Local-mp2 electron correlation method for non conducting crystals. *J. Chem. Phys.*, 122:094133, 2005.

[224] C. Pisani, R. Dovesi, and C. Roetti. *Hartree-Fock ab initio Treatment of Crystalline Systems*, volume 48 of *Lecture Notes in Chemistry*. Springer Verlag, Heidelberg, 1988.

[225] C. Pisani, A. Erba, S. Casassa, M. Itou, and Y. Sakurai. The anisotropy of the electron momentum distribution in α-quartz investigated by compton scattering and ab initio simulations. *Phys. Rev. B*, 84:245102, 2011.

[226] C. Pisani, M. Schütz, S. Casassa, D. Usvyat, L. Maschio, M. Lorenz, and A. Erba. Cryscor: a program for the post-hartree-fock treatment of periodic systems. *Phys. Chem. Chem. Phys.*, 14:7615, 2012.

[227] S Pittalis, G Vignale, and FG Eich. U (1)$\times$ su (2) gauge invariance made simple for density functional approximations. *Phys. Rev. B*, 96(3):035141, 2017.

[228] Russell M Pitzer and Nicholas W Winter. Electronic-structure methods for heavy-atom molecules. *J. Phys. Chem.*, 92(11):3061–3063, 1988.

[229] Russell M Pitzer and Nicholas W Winter. Spin-orbit (core) and core potential integrals. *Int. J. Quantum Chem.*, 40(6):773–780, 1991.

[230] J. A Pople and R. K. Nesbet. Self-consistent orbitals for radicals. *J. Chem. Phys.*, 22:571, 1954.

[231] M. J. D. Powell. *Nonlinear optimization*. NATO conference series: Systems science. Academic Press, New York, 1982.

[232] H. Preuss, editor. *Arbeitsbericht*, volume 27. Instituts für Theoretische Chemie, Universität Stuttgart, Stuttgart, 1990.

[233] S.A. Prosandeev, U. Waghmare, I. Levin, and J. Maslar. First-order Raman spectra of AB'$_{1/2}$B"$_{1/2}$O$_3$ double perovskites. *Phys. Rev. B*, 71:214307, 2005.

[234] P. Pulay. Improved scf convergence acceleration. *Journal of Computational Chemistry*, 3(4):556–560, 1982.

[235] P. Pulay and S. Saebø. Orbital-invariant formulation and second-order gradient evaluation in møller-plesset perturbation theory. *Theor. Chim. Acta*, 69:357, 1986.

[236] Péter Pulay. Convergence acceleration of iterative sequences. the case of scf iteration. *Chemical Physics Letters*, 73(2):393 – 398, 1980.

[237] M. Rérat, M. Ferrero, E. Amzallag, I. Baraille, and R. Dovesi. Comparison of the polarizability of periodic systems computed by using the length and velocity operators. *J. Phys.: Conf. Ser.*, 117:012023, 2008.

[238] Michel Rérat, Lorenzo Maschio, Bernard Kirtman, Bartolomeo Civalleri, and Roberto Dovesi. Computation of second harmonic generation for crystalline urea and kdp. an ab initio approach through the coupled perturbed hartree–fock/kohn–sham scheme. *Journal of Chemical Theory and Computation*, 12(1):107–113, 2016.

[239] R. Resta. The geometric phase approach to macroscopic polarization in crystalline dielectrics. *Rev. Mod. Phys.*, 66:809, 1994.

[240] 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.

[241] C. C. J. Roothaan. Self-Consistent Field Theory for Open Shells of Electronic Systems. *Rev. Mod. Phys.*, 32:179–185, 1960.

[242] RB Ross, JM Powers, T Atashroo, WC Ermler, LA LaJohn, and PA Christiansen. A binitio relativistic effective potentials with spin–orbit operators. iv. cs through rn. *J. Chem. Phys.*, 93(9):6654–6670, 1990.

[243] Richard B Ross, Sanjukta Gayen, and Walter C Ermler. Ab initio relativistic effective potentials with spin–orbit operators. v. ce through lu. *J. Chem. Phys.*, 100(11):8145–8155, 1994.

[244] 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.

[245] V. R. Saunders. From molecules to solids. Book in preparation.

[246] V. R. Saunders. *Ab initio* hartree-fock calculations for periodic systems. *Faraday Symp. Chem. Soc.*, 19:79–84, 1984.

[247] 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.

[248] 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, and M. Llunell. *CRYSTAL03 User's Manual*. Università di Torino, Torino,2003.

[249] V. R. Saunders, C. Freyria-Fava, R. Dovesi, and C. Roetti. On the electrostatic potential in linear periodic polymers. *Comp. Phys. Commun.*, 84:156, 1993.

[250] 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.

[251] A. Savin. A combined density functional and configuration interaction method. *Int. J. Quantum Chem.*, S22:59–69, 1988.

[252] Giovanni Scalmani and Michael J Frisch. A new approach to noncollinear spin density functional theory beyond the local density approximation. *J. Chem. Theor. Comput.*, 8(7):2193–2196, 2012.

[253] G. Schaftenaar. *MOLDEN a pre- and post processing program of molecular and electronic structure*. CMBI, the Netherlands.

[254] H. B. Schlegel. Optimization of equilibrium and transition structures. *J. Comp. Chem.*, 3:214–218, 1982.

[255] H. B. Schlegel. Estimating the hessian for gradient-type geometry optimizations. *Theor. Chim. Acta*, 66:333–340, 1984.

[256] Anthony P Scott and Leo Radom. Harmonic vibrational frequencies: an evaluation of hartree- fock, møller- plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *The Journal of Physical Chemistry*, 100(41):16502–16513, 1996.

[257] B. G. Searle. Dl visualize. *Comp. Phys. Commun.*, 137:25–32, 2001.

[258] V.F. Sears. Neutron scattering lengths and cross sections. *Neutron News*, 3(3):29–37, 1992.

[259] D. F. Shanno. Conditioning of quasi-Newton methods for function minimization. *Math. Comput.*, 24:647, 1970.

[260] D.A. Shirley. Isomer shift and quadrupolar splitting. *Rev. Mod. Phys.*, 36:339, 1964.

[261] W. J. Stevens, H. Bash, and 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.

[262] W. J. Stevens, M. Krauss, H. Basch, and 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.

[263] Hermann Stoll, Bernhard Metz, and Michael Dolg. Relativistic energy-consistent pseudopotentials—recent developments. *J. Comput. Chem.*, 23(8):767–778, 2002.

[264] Jianwei Sun, Adrienn Ruzsinszky, and John P Perdew. Strongly constrained and appropriately normed semilocal density functional. *Physical review letters*, 115(3):036402, 2015.

[265] R. Sure, J. G. Brandenburg, and S. Grimme. Small atomic orbital basis set first-principles quantum chemical methods for large molecular and periodic systems. *ChemistryOpen*, 5:94–109, 2016.

[266] R. Sure and S. Grimme. Corrected small basis set hartree-fock method for large systems. *J. Comput. Chem.*, 34:1672–1685, 2013.

[267] T. M. Henderson, A. F. Izmaylov, G. E. Scuseria and A. Savin. The importance of middle-range hartree-fock-type exchange for hybrid density functionals. *J. Chem. Phys.*, 127(22):221103, 2007.

[268] T. M. Henderson, A. F. Izmaylov, G. E. Scuseria and A. Savin. Assessment of a middle-range hybrid functional. *J. Chem. Theory Comput.*, 4(8):1254, 2008.

[269] T. M. Henderson, B. G. Janesko and Scuseria G. E. Generalized gradient approximation model exchange holes for range-separated hybrids. *J. Chem. Phys.*, 128:194105, 2008.

[270] N. Handy T. Yanai, D. Tew. A new hybrid exchangeârrelation functional using the coulomb-attenuating method (cam-b3lyp). *Chem. Phys. Lett.*, 393:51–57, 2004.

[271] Y. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai, and K. Hirao. A long-range-corrected time-dependent density functional theory. *J. Chem. Phys.*, 120:8425–8433, 2004.

[272] J. Tersoff and D. R. Hamann. Theory of scanning tunneling microscope. *Phys. Rev. B*, 31:805, 1985.

[273] 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.

[274] O. Treutler and R. Ahlrichs. Efficient molecular numerical integration schemes. *J. Chem. Phys.*, 102:346–354, 1995.

[275] Egor Trushin and Andreas Görling. Spin-current density-functional theory for a correct treatment of spin-orbit interactions and its application to topological phase transitions. *Phys. Rev. B*, 98(20):205137, 2018.

[276] P. Ugliengo, C. M. Zicovich-Wilson, S. Tosoni, and B. Civalleri. Role of dispersive interactions in layered materials: a periodic B3LYP and B3LYP-D* study of $Mg(OH)_2$, $Ca(OH)_2$ and kaolinite. *J. Mater. Chem.* , 19:2564–2572, 2009.

[277] Valérie Vallet, Laurent Maron, Christian Teichteil, and Jean-Pierre Flament. A two-step uncontracted determinantal effective hamiltonian-based so–ci method. *J. Chem. Phys.*, 113(4):1391–1402, 2000.

[278] M. Veithen, X. Gonze, and Ph. Ghosez. Nonlinear optical susceptibilities, raman efficiencies, and electro-optic tensors from first-principles density functional perturbation theory. *Phys. Rev. B*, 71:125107, 2005.

[279] G Vignale and Mark Rasolt. Current-and spin-density-functional theory for inhomogeneous electronic systems in strong magnetic fields. *Phys. Rev. B*, 37(18):10685, 1988.

[280] Daniel Vilela Oliveira, Joachim Laun, Michael F Peintinger, and Thomas Bredow. Bsse-correction scheme for consistent gaussian basis sets of double- and triple-zeta valence with polarization quality for solid-state calculations. *Journal of Computational Chemistry*, 40(27):2364–2376, 2019.

[281] U. von Barth and L. Hedin. A local exchange-correlation potential for the spin polarized case. *J. Phys. C: Solid State Phys.*, 5:1629, 1972.

[282] 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.

[283] Y. Wang, J. J. Wang, W. Y. Wang, Z. G. Mei, S. L. Shang, L. Q. Chen, and Z. K. Liu. A mixed-space approach to first-principles calculations of phonon frequencies for polar materials. *J. Phys.: Condens. Matter*, 22:202201, 2010.

[284] Ying Wang, Xinsheng Jin, S Yu Haoyu, Donald G Truhlar, and Xiao He. Revised m06-l functional for improved accuracy on chemical reaction barrier heights, noncovalent interactions, and solid-state physics. *Proceedings of the National Academy of Sciences*, 114(32):8487–8492, 2017.

[285] Ying Wang, Pragya Verma, Xinsheng Jin, Donald G Truhlar, and Xiao He. Revised m06 density functional for main-group and transition-metal chemistry. *Proceedings of the National Academy of Sciences*, 115(41):10257–10262, 2018.

[286] U. D. Wdowik and K. Ruebenbauer. Calibration of the isomer shift for the 14.4-kev transition in ^{57}fe using the full-potential linearized augmented plane-wave method. *Phys. Rev. B*, 76:155118, 2007.

[287] R. C. Weast, M. J. Astle, and W. H. Beyer, editors. *CRC Handbook of Chemistry and Physics*, pages A–61 and A–62. CRC Press, Inc., Boca Raton, Florida, 1986-87.

[288] Anna Weigand, Xiaoyan Cao, Tim Hangele, and Michael Dolg. Relativistic small-core pseudopotentials for actinium, thorium, and protactinium. *J. Phys. Chem. A*, 118(13):2519–2530, 2014.

[289] J. A. Weil, J. R. Bolton, and E. Wertz. *Electron ParamagneticResonance - Elementary Theory and Practical Applications*. John Wiley, New York, 1994.

[290] C. T. White, D. H. Robertson, and J. W. Mintmire. Helical and rotational symmetries of nanoscale graphitic tubules. *Phys. Rev. B*, 47:5485–5488, 1993.

[291] L. C. Wilson and M. Levy. Nonlocal Wigner-like correlation-energy density functional through coordinate scaling. *Phys. Rev. B*, 41:12930–12932, 1990.

[292] J. M. Wittbrodt and H. B. Schlegel. Estimating stretching force constants for geometry optimization. *J. Mol. Struct. - Theochem*, 398-399:55–61, 1997.

[293] Z. Wu and R. E. Cohen. More accurate generalized gradient approximation for solids. *Phys. Rev. B*, 73:235116, 2006.

[294] G. Wulff. Zur Frage der Geschwindigkeit des Wachstums und der Auflösung von Krystallflächen. *Z. Kristallogr.*, 34:499, 1901.

[295] E. Oldfield Y. Zhang, J. Mao. Nuclear magnetic resonance shifts in paramagnetic metalloporphyrins and metalloproteins. *J. Am. Chem. Soc.*, 124:7829, 2002.

[296] Y. Zhao and D. G. Truhlar. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.*, 125:194101, 2006.

[297] Y. Zhao and D. G. Truhlar. Density Functional for Spectroscopy: No Long-Range Self-Interaction Error, Good Performance for Rydberg and Charge-Transfer States, and Better Performance on Average than B3LYP for Ground States. *J. Phys. Chem. A*, 110:13126, 2006.

[298] Y. Zhao and D. G. Truhlar. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06- Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.*, 120:215, 2008.

[299] Y. Zhao, N. E. Schultz and D. G. Truhlar. Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.*, 2:364–382, 2005.

[300] Y. Zhao, N. E. Schultz and D. G. Truhlar. Exchange-Correlation Functional with Broad Accuracy for Metallic and Nonmetallic Compounds, Kinetics, and Noncovalent Interactions. *J. Chem. Phys.*, 123:161103, 2005.

[301] Haoyu S Yu, Xiao He, and Donald G Truhlar. Mn15-l: A new local exchange-correlation functional for kohn–sham density functional theory with broad accuracy for atoms, molecules, and solids. *Journal of chemical theory and computation*, 12(3):1280–1293, 2016.

[302] Y. Zhao and D. G. Truhlar. Construction of a generalized gradient approximation by restoring the density-gradient expansion and enforcing a tight Lieb–Oxford bound . *J. Chem. Phys.*, 128:184109, 2008.

[303] Yan Zhao and Donald G Truhlar. Hybrid meta density functional theory methods for thermochemistry, thermochemical kinetics, and noncovalent interactions: the mpw1b95 and mpwb1k models and comparative assessments for hydrogen bonding and van der waals interactions. *The Journal of Physical Chemistry A*, 108(33):6908–6918, 2004.

[304] Yan Zhao and Donald G Truhlar. Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *The Journal of Physical Chemistry A*, 109(25):5656–5667, 2005.

[305] C. M. 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.

[306] C. M. 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.

[307] C. M. Zicovich-Wilson and R. Dovesi. Localized functions in crystalline systems and their variational manifolds. In R. Hernández-Lamoneda, editor, *Beyond standard quantum chemistry: applications from gas to condensed*

phases 2007, pages 140–169, Kerala, India, 2007. Transworld Research Network.

[308] 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.

[309] C. M. Zicovich-Wilson, R. Dovesi, and V. R. Saunders. A general method to obtain well localized wannier functions for composite energy bands in linear combination of atomic orbital periodic calculation. *J. Chem. Phys.*, 115:9708, 2001.

[310] J.M. Ziman. *Principles of the Theory of Solids*. Cambridge University Press, London, England, 1964.

Index